Semiconductors, Vol. 39, No. 6, 2005, pp. 609–614. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 641–646.

ATOMIC STRUCTURE AND NONELECTRONIC PROPERTIES OF SEMICONDUCTORS

Kinetics of Defect Formation in ZnO Subjected to a Flux of Oxygen Radicals

M. B. Kotlyarevsky*, I. V. Rogozin^, and A. V. Marakhovskiĭ****

**Academy of Management and Information Technologies, Berdyansk, 71118 Ukraine*

***Berdyansk State Pedagogical University, Berdyansk, 71100 Ukraine*

^e-mail: rogozin@bdpu.org

Submitted September 20, 2004; accepted for publication October 14, 2004

Abstract—A kinetic analysis of the defect formation in a II–VI compound (ZnO) subjected to a flux of oxygen radicals has made it possible to develop a new method for obtaining single-crystal layers that uses a high-temperature heat treatment of the above single-crystal compound under these conditions. Using this method, a change from *n*- to *p*-type conductivity in ZnO is accomplished. This technology makes it possible to obtain heterostructures. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

In this study, high-energy chemistry methods are used in the synthesis of II–VI film structures with specified properties [1]. In this case, the gaseous phase of the VI component is activated using an external excitation source (photolysis, gaseous discharge, etc.) until the point where it becomes neutral, excited, or ionized radicals (atoms) and enters into the reaction zone. One of the main advantages of the methods of high-energy chemistry is that these methods reduce the synthesis temperature.

We showed that these methods offer potential benefits in relation to the technology of II–VI and III–V semiconductor compounds, in which the nonmetal components (O, N, S, Se, Te, P, and As) exist only in the gaseous phase at the temperatures of the compound's synthesis.

The problem with the aforementioned compounds (ZnS, ZnO, ZnSe, CdS, and GaN), which have band gaps ≥2.4 eV, is related to the high (and, therefore, difficult to control) concentration of intrinsic defects (donors) that are actively generated in the course of their high-temperature synthesis. In turn, this circumstance gives rise to the *n*-type monopolar defect-related conductivity of the compounds under consideration and restricts their application as the main components in semiconductor optoelectronics.

Let us consider the interaction of a II–VI crystal (ZnS, ZnSe, or ZnO) with the vapor phase of a Group VI (S, Se, or O) element. Under equilibrium conditions, a heat treatment of the crystal in the vapors of the Group VI element at high temperatures does not give rise to any superstoichiometric defects related to the chalcogen (for example, acceptor intrinsic defects of an interstitial type (B_i) in the chalcogen sublattice or vacancy defects (V_{Zn}) in the Zn sublattice), since it is thermodynamically advantageous to form a compensating donor center (V_B or A_i) when an acceptor center is formed (here and in what follows, the designations *A* and *B* correspond to the elements of Groups II and VI, respectively). The high-temperature equilibrium between the intrinsic defects in ZnS, ZnSe, CdS, and ZnO was previously thermodynamically analyzed in [2–5]. Calculations make it possible to conclude with confidence that, in these compounds, there is a wide range of chalcogen-vapor pressures p_{B_2} (for example, p = 100–1000 GPa for ZnSe and ZnS) when the prevalent defects are V_A' and $V_B^{\bullet}$, meaning that the defect-related intrinsic conductivity itself is completely compensated.

In our previous publications [6–8], we analyzed the thermodynamic and kinetic aspects of the self-compensation of *p*-type conductivity in detail and established that the main factor controlling the monopolar conductivity type of the compounds under consideration is the fact that the chalcogen atoms are present in the gaseous phase in the form of molecules. It was also shown [6–8] that, for the II–VI compounds, if the synthesis is accomplished in the activated chalcogen vapor, the suppression of acceptor-center compensation can be attained by decreasing the synthesis temperature.

In Fig. 1, we show the results of calculating the dependence of the concentrations of defects in ZnS on the partial pressures of the Zn and S_2 vapors at 1000 K. As can be seen from Fig. 1, there is a relatively wide range of S_2 vapor pressures (10^{-1}–10^6 Pa) for which the conductivity is nearly unchanged and remains very low. The calculations show that the range of vapor pressures p_{B_2} at which the self-compensation factor begins to decrease is within the range of inaccessible sulfur-

1063-7826/05/3906-0609$26.00 © 2005 Pleiades Publishing, Inc.

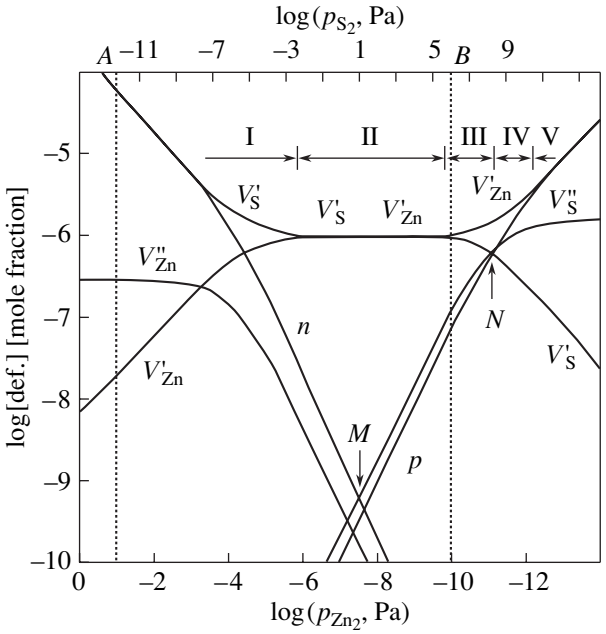

Fig. 1. The concentration equilibrium of intrinsic defects in ZnS at 1000 K. The points *A* and *B* correspond to saturated vapors of Zn and S_2, respectively. See the text for an explanation of the other notation.

vapor pressures. These conclusions are also valid for zinc oxide.

The thermodynamic factors that give rise to the tendency of II–VI compounds towards monopolar conductivity are caused by special features and properties of their crystal–vapor phases. A thermodynamic analysis suggests certain methods for obtaining intrinsic defect-related and extrinsic *p*-type conductivity in II–VI compounds. These methods include, first and foremost, those of high-energy chemistry with respect to the treatment of the crystals, for example, ion-implantation doping with the intrinsic nonmetallic component or an acceptor impurity, an annealing of the crystal after being coated with a material in which the solubilities of the binary-compound components are inadequate, and also an annealing of the crystal in an atmosphere of the atomic chalcogen [9–11]. It is the latter method that we consider below.

2. A KINETIC ANALYSIS OF THE TYPES AND CONCENTRATIONS OF INTRINSIC DEFECTS

The aforementioned method implies an increase in the atomic-chalcogen flux J_B to the point where it exceeds the equilibrium value for a given pressure p_{B_2} and for a given temperature owing to dissociation of the chalcogen molecules under an external effect (here and below, B stands for a specific type of chalcogen). The percentage of dissociating oxygen molecules can be as high as 15% in an RF discharge. If 1% of the oxygen molecules are dissociated at a pressure $p_{O_2} \approx 1$ Pa, the atomic-oxygen pressure is $p_O \approx 10^{-2}$ Pa. The same pressure of activated oxygen can be attained owing to a thermal dissociation of the molecular oxygen at a pressure $p_{O_2} = p_O^2/K_D = 10^2$ GPa at 1000 K (here, K_D is the rate constant of the reaction $O_2 = 2O + D$, where *D* is the energy at which dissociation of an oxygen molecule occurs). Consequently, as a result of heat treatment in the gaseous phase with an RF activation, an atomic-oxygen pressure that is practically unattainable using thermal dissociation under the given conditions of the heat treatment is produced.

A nonequilibrium concentration of the chalcogen atoms in the course of annealing corresponds to an effective (much higher) pressure p_{B_2} that is not accompanied by condensation, which makes it practically possible to realize the states that correspond to the points *M* and *N* (see Fig. 1) and, as a consequence, attain the stoichiometric and superstoichiometric content of chalcogen in the compound.

The kinetic method of analysis of the composition of intrinsic defects has been found very useful in the case of a system consisting of a crystal and an activated gaseous phase of the nonmetal component. Exactly this approach was used in [12] to show that it is possible to appreciably widen the range of the conductivity related to intrinsic defects and to attain the inversion of the conductivity type in the compounds under consideration. An analysis of the defect-formation kinetics made it possible to suggest certain physicochemical methods for controlling the defect-formation processes that depend on the adsorption–desorption–crystallization equilibrium at a crystal's surface (Fig. 2).

Using the model of adsorption–desorption–crystallization equilibrium, we now suggest a model of the defect formation in II–VI compounds. We derived and solved a set of kinetic equations for a system that consists of a II–VI crystal and the chalcogen vapor. We assumed that only the atomic component of the vapor interacts with the crystal's surface. In calculating the adsorption–desorption–crystallization equilibrium at the crystal's surface, we took into account the interaction of neutral and charged defects with the gaseous phase.

The system of kinetic differential equations describing the defect-formation process is written as

$$\frac{dN'_B}{d\tau} = J_B - J_{\mathrm{des}B} - 2J_{\mathrm{des}B_2} - J_{\mathrm{cr}} + J_{\mathrm{rec}(AV_A)} - J_{\mathrm{rec}(BV'_B)}, \quad (1)$$

$$\frac{dV_A}{d\tau} = J_{\mathrm{cr}} - J_{\mathrm{rec}(AV_A)} - g[V_A] + \gamma_h p[V'_A] - J_{\mathrm{des}(BV_A)}, \quad (2)$$

$$\frac{dV'_A}{d\tau} = g[V_A] - \gamma_h p[V'_A] + J_{\mathrm{des}(BV_A)}, \quad (3)$$

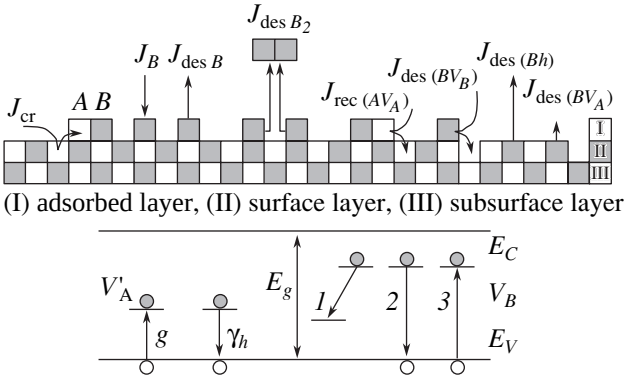

Fig. 2. A schematic representation that illustrates both the adsorption–desorption–crystallization equilibrium at the surface of *AB* crystals and the corresponding electronic transitions: (*1*) a transition as a result of interaction between a *B* atom in the surface layer and V_A ($(J_{\mathrm{des}(BV_A)})$); (*2*) a transition as a result of interaction between a *B* atom in the surface layer and a hole ($J_{\mathrm{des}(Bh)}$); and (*3*) a transition as a result of the recombination of a *B* atom in the adsorbed layer with $V_B^{\cdot}$ ($J_{\mathrm{rec}(BV_B^{\cdot})}$).

$$\frac{dV_B^{\cdot}}{d\tau} = J_{\mathrm{des}(BV_A)} + J_{\mathrm{des}(Bh)} - J_{\mathrm{rec}(BV_B^{\cdot})}, \quad (4)$$

$$\frac{dp}{d\tau} = g[V_A] - \gamma_h p[V'_A] - J_{\mathrm{des}(Bh)} + J_{\mathrm{rec}(BV_B^{\cdot})}, \quad (5)$$

$$p + [V_B^{\cdot}] = [V'_A]. \quad (6)$$

In Fig. 3, we show the dependences of the defect concentrations in a ZnO crystal on the flux of oxygen atoms. As can be seen, an increase in temperature with the J_O flux kept constant does not increase its *p*-type conductivity; rather, only the compensation factor is increased. The value of the critical oxygen flux at which the transition from compensated conductivity to uncompensated conductivity occurs and the electroneutrality equation $[V'_{Zn}] = [V_O^{\cdot}]$ transforms into $[V'_{Zn}] = p$ (i.e., the situation where the appearance of the metal vacancies, rather than compensated chalcogen vacancies, leads to the appearance of free holes in the valence band) depends on the annealing temperature.

For the above critical oxygen flux, there exists a critical temperature above which it is impossible to obtain uncompensated *p*-type conductivity. This critical temperature T_C is defined as

$$T_C = \frac{G_{\mathrm{des}B} - E_g + E_{a_1} + E_{d_1} + G_{\mathrm{dif}A} + G_{\mathrm{dif}B}}{k\{\ln(\nu_A d_B^2 N_B \nu_B^{1/2}) - \ln(J_B^{1/2}\nu_B d_A)\}}. \quad (7)$$

In Fig. 4, we show the dependence of the critical temperature on the flux of oxygen atoms. It can be seen that the critical temperature can be appreciably increased only by increasing J_O. Calculations show that, for ZnO, T_C is equal to 800 K for oxygen-atom fluxes on the order of 10^{16}–10^{17} cm^{-2} s^{-1}, which are quite attainable under experimental conditions.

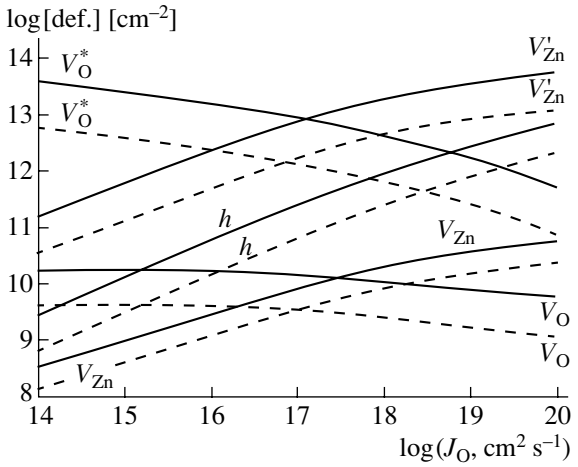

Fig. 3. Concentrations of defects in the surface layer of a ZnO crystal in relation to the flux of oxygen atoms (radicals) at 1200 K (the solid lines) and 1000 K (dashed lines).

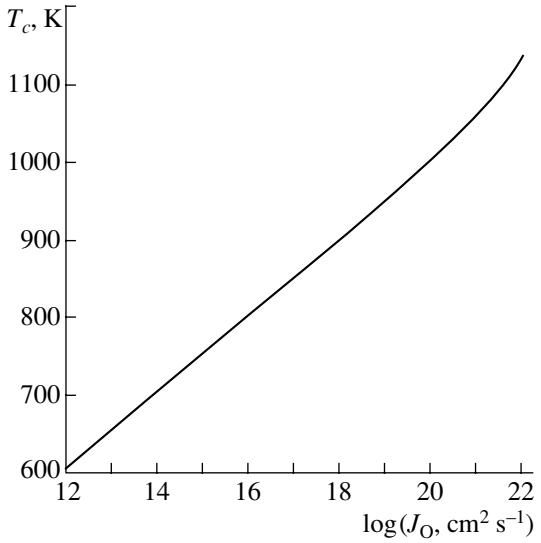

Fig. 4. The critical temperature as a function of a flux of oxygen atoms.

Thus, it follows from the thermodynamic and kinetic analyses that one of the methods for controlling the types and concentration of intrinsic defects in the compounds under consideration, in order to widen the range within which the conductivity type can be controlled, consists in an increase in the nonmetal element's flux above the equilibrium value for a given p_{B_2} and a given temperature as a result of the molecular-nonmetal's activation.

3. RADICAL-BEAM GETTERING EPITAXY

On the basis of our analysis of the adsorption–desorption–crystallization equilibrium in a system consisting of a crystal and radicals (atoms) in the gaseous

phase, we developed a new method for fabricating epitaxial II–VI films in heterophase and homophase systems. In this method, technology based on the heat treatment of a II–VI compound in an atmosphere of chalcogen radicals (atoms) is used for the first time [13]. Generation of the radicals (atoms) is accomplished using an RF discharge or photolysis. We named this method radical-beam gettering epitaxy (RBGE) [14]. In the case of annealing in an atomic-chalcogen vapor (for example, an annealing of ZnO in an atmosphere of O radicals), a quasi-epitaxial mechanism, in which the crystal-lattice layers are grown on a crystalline substrate, is prevalent. In this situation, the chalcogen atoms are supplied from the gaseous phase, while the metal atoms are drawn (gettered) by the chalcogen atoms from the crystal's bulk, with the subsequent formation of acceptors (metal vacancies, Zn vacancies in the case under consideration).

Since zinc vacancies can be found in two charge states, depending on the degree of compensation of these vacancies by oxygen vacancies and on the position of the Fermi level in the *p*-type crystals, one of the two levels corresponding to a singly or doubly charged zinc vacancy (V'_{Zn} or V''_{Zn}) can be observed in electrical conductivity.

We [15] studied the photoluminescence (PL) of ZnO layers with a superstoichiometric content of oxygen. ZnO single crystals (both undoped and doped with Li or Na) grown from the vapor phase were used. The impurity concentration was 10^{16}–10^{17} cm^{-3}, and the electrical conductivity was 10^{-5}–10^{-7} S/cm. All of the samples were subjected to a heat treatment according to the RBGE technology.

The initial ZnO single crystals were annealed in an atmosphere of oxygen radicals (atoms) at 670–970 K. The oxygen pressure in the growth chamber was $1–10^{-3}$ Pa. In order to obtain the oxygen radicals, we used an RF discharge with a power of 80 W. Separation of the charged particles was accomplished using a magnetic filter. The flux of atomic oxygen was $J_O = 10^{16}$–10^{17} cm^{-2} s^{-1}.

The ZnO layers formed when the oxygen-radical flux $J_O \approx 1.5 \times 10^{17}$ cm^{-2} s^{-1} and the heat-treatment temperature was 670–870 K had *p*-type conductivity according to studies of the Hall effect. The *p*-ZnO layers with the lowest resistivity ρ were obtained at a heat-treatment temperature of 710 K and featured $\rho \approx 6 \times 10^2$ Ω cm and a hole mobility $\mu_p \approx 23$ cm^2/(V s). An increase in the heat-treatment temperature T_a to 810 K led to an increase in the resistivity to $\rho \approx 3 \times 10^7$ Ω cm. At $T_a \geq 910$ K, the ZnO layers featured *n*-type conductivity and a resistivity higher than 10^9 Ω cm. The data reported above indicate that the critical heat-treatment temperature for ZnO is exceeded at $T_a \geq 910$ K. These data are in good agreement with the results of theoretical calculations.

In the ultraviolet region, the PL spectrum of *p*-ZnO layers with a superstoichiometric oxygen content exhibits a high-intensity band that is peaked at 369.5 nm and has a half-width of 14 meV. A low-intensity band with a peak at 400 nm is observed in the visible region of the PL spectrum. By comparing the PL spectra of *n*- and *p*-ZnO layers, we can conclude that the corresponding PL spectra exhibit bound-exciton bands with peaks at 369.0 and 369.5 nm, respectively. It was shown by Butkhuzi *et al.* [16] that the band peaked at 369.0 nm is related to excitons bound to neutral donors, whereas the band peaked at 369.5 nm is related to excitons bound to neutral acceptors.

An increase in the duration of annealing at 710 K leads to an increase in the intensity of the band with a peak at 400 nm. A single band peaked at 400 nm is observed in the PL spectra of the samples annealed for 4 h. It can be seen that this band overlaps completely the excitonic region of the spectrum.

By analogy with the 345-nm band in ZnO [8], we assume that a singly positively charged $(V_{Zn}–V_O)^{\bullet}$ complex can represent the center responsible for the band peaked at 400 nm. This assumption is supported by the fact that the oxygen-vacancy concentration decreases as a result of the enrichment of ZnO with oxygen; consequently, the electroneutrality is maintained owing to the recharging of a fraction of the vacancies to the doubly charged state $(V_O^{\bullet\bullet})$.

4. HETEROPHASE EPITAXY OF THE II–VI COMPOUNDS

The kinetics of defect formation in a heterophase system is more complex. In this case, the crystal is represented by a compound that involves a nonmetal B' (for example, selenium), whereas the flux of B" radicals is related to another nonmetal (for example, oxygen).

However, the main laws and mechanisms of defect formation in heterophase epitaxy should be similar to those of homophase epitaxy. In an AB' crystal, V_A vacancies are formed, while the A metal atoms that reach the surface layer form AB" dimers. In system of kinetic Eqs. (1)–(5), we should take into account the fact that atoms of one nonmetal can combine with those of the other nonmetal to form molecules (for example, SeO_2), with the subsequent desorption of these molecules. This kind of desorption leads to the formation of $V_{B'}$ defects in the crystal; as a result, the crystal itself transforms into an AB" crystal owing to diffusion of B" over the $V_{B'}$ vacancies.

Thus, two mechanisms of defect formation compete in a heterophase system in the early stages of annealing. The first mechanism is quasi-epitaxial and is related to formation of V_A vacancies and growth of a AB" layer. The second mechanism is related to the replacement of B' atoms with B" atoms.

In the first approximation, the set of kinetic equations that simulates the processes in the heteroepitaxial system under consideration can be written as (it is assumed that A corresponds to Zn; B', to Se; and B", to O)

$$\frac{dN'_O}{d\tau} = J_O - J_{desO} - 2J_{desO_2} - J_{des(SeO_2)} - J_{cr} - J_{rec(OV_{Se})} - J_{rec(ZnV_{Zn})}, \quad (8)$$

$$\frac{dN_O}{d\tau} = \lambda[V_{Se}]J_O + J_{rec(OV_{Se})} - J_{des(OV_{Zn})}, \quad (9)$$

$$\frac{dN_{Zn}}{d\tau} = J_{cr} - J_{rec(ZnV_{Zn})}, \quad (10)$$

$$\frac{dN_{Se}}{d\tau} = J_{des(SeO_2)} - J_{rec(OV_{Se})} + J_{des(SeV_{Zn})}. \quad (11)$$

The equation for the material balance is written as $[V_{Se}] + [O] + [Se] = \text{const}$.

The notation in Eqs. (8)–(11) is the same as that used in Eqs. (1)–(5); i.e., J_O is the flux of oxygen atoms (radicals) incident on a two-dimensional crystal, N'_O is the concentration of oxygen atoms adsorbed in the layer on the surface, N_O is the concentration of oxygen atoms in the surface layer, and [O] is the oxygen concentration in the adsorbed and surface layers.

In the heterophase case, diffusion to the bulk from the surface layer is very important. It is the diffusion processes that determine, to a great extent, whether the ZnO surface grows and an abrupt interface (Fig. 5) is formed in the ZnO–ZnSe(S, Te) heterojunction or whether the ZnO layers grow owing to the oxygen diffusion into the bulk, thus forming to an interface that is diffuse. After solving Eqs. (8)–(11), we can use the ratio between the rates of generation of the V_A and V_B vacancies to determine the parameters of heterophase quasi-epitaxy for which particular types of heterojunctions are obtained, the conductivity type, and the defect concentration in the ZnO and ZnSe(S, Te) layers.

Thus, the method based on the quasi-epitaxy model that we suggested for analyzing the defect-formation kinetics indicates clearly that the role of the surface is important in the adsorption–desorption–crystallization equilibrium. This method makes it possible to assess the role of the conditions (temperature and J_B flux) in the establishment of transient and steady-state concentration relations between defects.

Simulation of the defect-formation kinetics in a system consisting of a II–VI crystal and nonequilibrium chalcogen vapor showed that the steady-state concentrations of defects in the surface layers are established very rapidly; therefore, as was assumed in the model, the diffusion of defects into the crystal bulk does not affect these concentrations.

Using this technological method, we can obtain virtually any optoelectronic and acoustoelectronic device

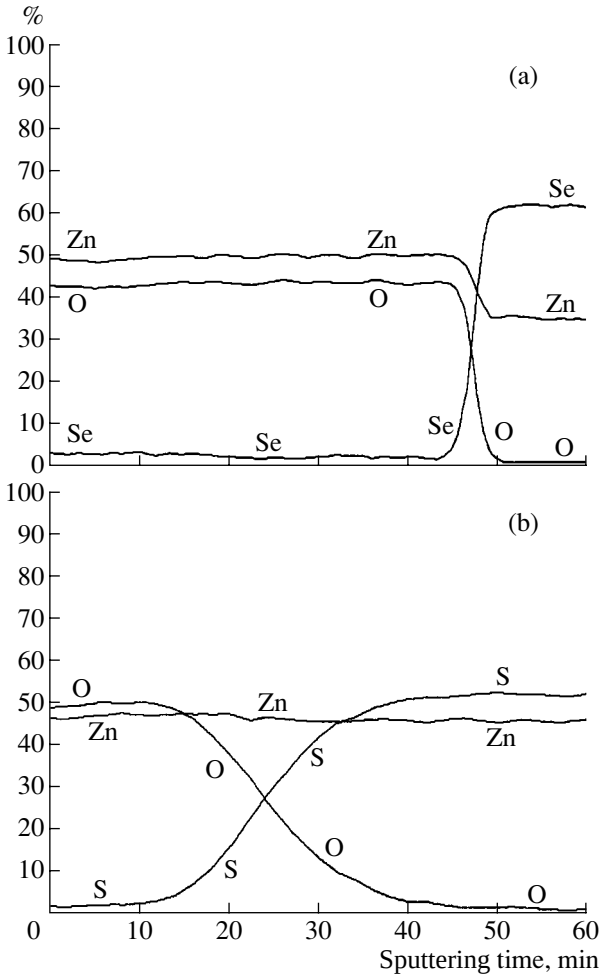

Fig. 5. Concentration profiles in (a) ZnO–ZnSe and (b) ZnO–ZnS heterostructures. The substrate temperature is 800 K.

structures (for example, ZnO–ZnSe, ZnS–ZnSe, CdS–CdSe, ZnO–ZnTe, and other structures). Moreover, the parameters of the synthesized layers can be different. For example, ZnO layers formed on a ZnSe substrate can exhibit either *n*- or *p*-type conductivity.

Using the above-described method, we earlier obtained and studied [17] ZnO–ZnSe emitting structures. The spectral characteristics of electroluminescence (EL) in these structures were measured at 300 K. High-intensity bands peaked at 467.0 and 525.0 nm and a low-intensity band with a peak at 615.0 nm were observed in the EL spectrum. The band peaked at 525.0 nm is caused by the appearance of *p*-ZnSe layers. It is worth noting that the structure begins to exhibit a low-intensity EL at a voltage of ~2.5 V; however, the studies of the spectral characteristics were hampered by the low emission intensity. The following special features of the EL should be noted: low injection voltages and the presence of an EL band, whose intensity increases by a factor of 10 as the voltage increases from 4.5 to 8 V, in the region of the fundamental-absorption

edge. Measurements of the external emission efficiency showed that the fabricated structures exhibited an external emission efficiency of 0.1%.

The current–voltage (*I–V*) characteristic exhibits a linear portion of as high as 2.5 V at forward-bias voltages. This portion corresponds to the Sah–Shockley–Noyce formula derived for the case of the charge-carrier recombination in the region of a *p–n* junction: $I = I_S\exp[qU/(nkT - 1)]$. The fact that the dependence is no longer exponential at bias voltages in excess of 2.5 V can be attributed to the presence of ZnSe bulk resistance; it follows that the barrier height in the *p–n* junction is ~2.5 V.

We studied the *p–n*-junction region using a PHI-6600 secondary-ion mass spectrometer [17]. The concentration profiles for O and Se indicate that the ZnO–ZnSe junction is abrupt. The results of measuring the *I–V* characteristic suggest that the *p–n* junction is located within the ZnSe region; i.e., we can state that a *p*-ZnO–*p*-ZnSe–*n*-ZnSe structure is formed. The surface ZnO layer plays the role of an injecting contact to the *p*-ZnSe layer and also serves as a window for letting the emission out.

REFERENCES

1. A. N. Georgobiani, M. B. Kotlyarevskiĭ, V. N. Mikhalenko, *et al.*, USSR Inventor's Certificate No. 684 810 (1977).
2. N. K. Morozova and O. M. Morozova, Izv. Akad. Nauk SSSR, Neorg. Mater. **17**, 1335 (1981).
3. M. T. Nyges, Candidate's Dissertation (Tallin, 1974).
4. A. M. Gurvich, *Introduction to the Physical Chemistry of Phosphorescent Crystals* (Vysshaya Shkola, Moscow, 1982), p. 132 [in Russian].
5. V. A. Nikitenko, S. A. Stenli, and N. K. Morozova, Izv. Akad. Nauk SSSR, Neorg. Mater. **24**, 1830 (1988).
6. V. N. Mikhalenko, B. P. Dement'ev, M. B. Kotlyarevskiĭ, and A. N. Georgobiani, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 8, 150 (1978).
7. A. N. Georgobiani, M. B. Kotlyarevskiĭ, and V. N. Mikhalenko, Izv. Akad. Nauk SSSR, Neorg. Mater. **17**, 1329 (1981).
8. A. N. Georgobiani, M. B. Kotlyarevskiĭ, and V. N. Mikhalenko, Tr. Fiz. Inst. im. P. N. Lebedeva, Akad. Nauk SSSR **138**, 79 (1983).
9. A. N. Georgobiani and M. B. Kotlyarevsky, Radiat. Eff. **47**, 21 (1980).
10. A. N. Georgobiani, M. B. Kotlyarevskiĭ, and V. N. Mikhalenko, Kratk. Soobshch. Fiz., No. 4, 14 (1977).
11. A. N. Georgobiani, M. B. Kotlyarevskiĭ, V. V. Lastovka, and D. A. Noskov, Kratk. Soobshch. Fiz., No. 6, 30 (1977).
12. V. N. Mikhalenko, M. B. Kotlyarevskiĭ, A. N. Georgobiani, and V. A. Sokolov, Zh. Fiz. Khim. **54**, 345 (1980).
13. A. N. Georgobiani and M. B. Kotlyarevskiĭ, Izv. Akad. Nauk SSSR, Neorg. Mater. **17**, 1153 (1981).
14. A. N. Georgobiani, M. B. Kotlyarevskiĭ, and V. N. Mikhalenko, Izv. Akad. Nauk SSSR, Neorg. Mater. **18**, 12 (1982).
15. M. B. Kotlyarevskiĭ, A. N. Georgobiani, I. V. Rogozin, and A. V. Marakhovskiĭ, Zh. Prikl. Spektrosk. **70** (1–3), 86 (2003).
16. T. V. Butkhuzi, A. V. Bureyev, A. N. Georgobiani, *et al.*, J. Cryst. Growth **117**, 366 (1992).
17. A. N. Georgobiani, M. B. Kotlyarevskiĭ, V. V. Kidalov, and I. V. Rogozin, Neorg. Mater. **33** (2), 232 (1997).

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 615–622. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 647–654.

ATOMIC STRUCTURE AND NONELECTRONIC PROPERTIES OF SEMICONDUCTORS

Specific Features of the Segregation-Related Redistribution of Phosphorus during Thermal Oxidation of Heavily Doped Silicon Layers

O. V. Aleksandrov* and N. N. Afonin**

**St. Petersburg State Electrotechnical University, St. Petersburg, 197376 Russia*
e-mail: Aleksandrov@svs.ru
***Voronezh State Pedagogical University, ul. Lenina 86, Voronezh, 394043 Russia*
e-mail: nafonin@vspu.ac.ru

Submitted May 25, 2004; accepted for publication November 10, 2004

Abstract—A model of the diffusion–segregation redistribution of phosphorus in an SiO_2/Si system during thermal oxidation of highly doped silicon layers is developed taking into account the formation of a peak of surface impurity concentration at the interface. The formation of this surface concentration peak is attributed to a change in the free energy of the impurity atoms near the silicon surface. This process is simulated by a diffusion–segregation equation. It is shown that the developed diffusion–segregation model is quite adequate for describing the phosphorus redistribution occurring during the oxidation of uniformly doped silicon layers. For the oxidation of implanted silicon layers, it was found that the segregation coefficient of the phosphorus at the SiO_2/Si interface is not constant but depends on time in the same way as the efficiency of transient enhanced diffusion in silicon. This phenomenon is explained by the reactivity of the impurity segregation during the thermal oxidation of silicon, when excess point defects in the implanted silicon layer affect both the oxidation process and the capture of impurity atoms by the growing silicon dioxide. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The diffusion doping and ion implantation of silicon with phosphorus are widely used in the fabrication of semiconductor devices and integrated circuits. Therefore, an adequate description of the phosphorus redistribution occurring during a thermal oxidation of silicon is urgently required. When the impurity concentration is high, a number of phenomena manifest themselves in the behavior of phosphorus in silicon. These phenomena must be taken into account in any simulation.

When the majority-carrier concentration exceeds the intrinsic carrier concentration at a certain annealing temperature, the diffusion of phosphorus in silicon is enhanced. This enhancement manifests itself in the concentration dependence of the diffusion coefficient and the formation of a plateau and a tail on the concentration distribution of the impurity [1, 2]. At high phosphorus concentrations, a polytropic effect occurs, which leads to an incomplete electrical activation of part of the phosphorus impurity in heavily doped Si layers [1, 3]. When silicon is thermally oxidized in the so-called passive mode (at relatively high oxidant partial pressures), the generation of Si interstitials at the SiO_2/Si interface, which leads to the oxidation-enhanced diffusion of phosphorus [4], has to be taken into account. Ion-implanted Si layers exhibit the phenomenon of so-called transient enhanced diffusion, in which the diffusion of an impurity into Si is caused by the high concentration of intrinsic point defects formed during the annealing of radiation defects introduced by the ion implantation [5].

The thermal oxidation of silicon is accompanied by a segregation of the impurities at the moving SiO_2/Si interface. When Si is doped with donors (P, As, or Sb), this effect leads to the enrichment of the surface Si layer with impurities, due to the rejection of impurity atoms by the growing dioxide. Along with the ordinary segregation-induced accumulation occurring during the thermal oxidation of heavily doped Si layers, the formation of a peak of the surface impurity concentration within a much smaller localization region (1–10 nm) is observed. This surface concentration peak cannot be described within the ordinary diffusion–segregation model based on the Fick equation [6]. The main systematic features of the manifestation of the surface concentration peak with respect to phosphorus, depending on the doping level, temperature, and annealing medium, have been established by Auger electron spectroscopy (AES) and secondary-ion mass spectrometry (SIMS) [7–10].

We previously proposed [11] a model of the formation of the surface concentration peak in silicon uniformly doped with phosphorus during annealing in a neutral medium. The existing models of the segregation-induced redistribution of phosphorus during annealing in an oxidizing medium disregard the formation of a peak of surface impurity concentration. For

example, in the model of nonequilibrium segregation [12], the segregation coefficient m_s was determined under the assumption that the impurity concentration at the SiO_2/Si interface originating from the Si side is equal to the impurity concentration at the foot of the surface concentration peak. In addition, the diffusion of the impurity in silicon was disregarded. It was proposed in [13, 14] that the surface concentration peak should be considered to be a result of the trapping of the impurity from both Si and SiO_2 in a thin intermediate layer (the third phase). In this case, the segregation coefficient was determined as the ratio of the concentrations in silicon and in the dioxide at different sides of the intermediate layer; i.e., the surface concentration peak was disregarded.

The purpose of this study is to develop a model of the diffusion–segregation redistribution of phosphorus in a SiO_2/Si system during thermal oxidation of heavily doped Si layers, taking into account the formation of a peak of surface impurity concentration at the interface.

2. MODEL OF THE DIFFUSION–SEGREGATION REDISTRIBUTION OF PHOSPHORUS

In a mathematical analysis of phosphorus redistribution in a SiO_2/Si system, along with the interface motion, a number of phenomena characteristic of high concentrations of phosphorus in silicon have to be taken into account; in particular, the formation of a surface concentration peak at the SiO_2/Si interface should be considered. Let us assume that the surface concentration peak is formed due to the difference between the free energy of the impurity atoms in the silicon surface region and that in the silicon bulk. In this case, the distribution of the surface impurity concentration corresponding to thermal equilibrium has the form

$$C^*(x) = C_0 \exp(-f(x) dE_b/kT), \quad (1)$$

where C_0 is the impurity concentration in the silicon bulk, dE_b is a decrease in the free energy of the impurity atoms at the silicon surface as compared with the bulk, k is the Boltzmann constant, and T is absolute temperature. The free energy distribution function $f(x)$ near the surface is assumed to be the same as the one we used in [11]:

$$f(x) = 1/(1 + x/x_m), \quad (2)$$

where x is a coordinate measured from the SiO_2/Si interface to the silicon bulk and x_m is the characteristic length of the surface free energy distribution.

At high impurity concentrations, the diffusion coefficient becomes concentration-dependent. Under equilibrium conditions with respect to intrinsic point defects, the diffusion coefficient of phosphorus in silicon is calculated using the concentration dependence [15]:

$$\begin{aligned} D^* &= D^0 + D^- + D^{2-} \\ &= D_i^0 + D_i^-(n/n_i) + D_i^{2-}(n/n_i)^2. \end{aligned} \quad (3)$$

Here, D^0, D^-, and D^{2-} are the partial coefficients of the diffusion via intrinsic point defects in, respectively, neutral, negative, and doubly negatively charged states; n is the electron concentration; and the subscript i corresponds to an intrinsic semiconductor. In the case of nonuniform doping, diffusion coefficient (3) is multiplied by the coefficient of diffusion enhancement by an internal electric field: $h_e = 2n/(2n - C_a)$. Here, C_a is the concentration of electrically active impurity.

Another specific feature of the behavior of phosphorus in heavily doped silicon is the polytropic effect, which leads to an incomplete electrical activation of the impurity. The reason for this phenomenon is the formation of Si_lP_m clusters of P and Si lattice atoms. The number of P atoms m in these clusters increases as the impurity concentration in the layer increases [3]. The clustering of some of the impurity atoms makes them not only electrically inactive but also unable to migrate. The impurity redistribution between the electrically active and inactive states is relatively fast; hence, the cluster formation is taken into account by the factor $\partial C/\partial C_a$ of diffusion coefficient (3) using the dependence $C(C_a)$ from [3].

When thermal oxidation is considered, it is necessary to take into account the perturbation in the ensemble of intrinsic point defects caused by the generation of intrinsic interstitials. Since P atoms diffuse through the silicon lattice via the dual vacancy–interstitial (indirect) mechanism, with the formation of *PV* and *PI* pairs, the corresponding diffusion coefficient is determined by the relation [16]

$$D = D_{PV}^* a_V + D_{PI}^* a_I, \quad (4)$$

where D_{PV}^* and D_{PI}^* are the coefficients of the diffusion of phosphorus via the *PV* and *PI* pairs ($D^* = D_{PV}^* + D_{PI}^*$); a_V and a_I are the degrees of supersaturation (or undersaturation) with respect to the vacancies and intrinsic interstitials, respectively; $a_V = C_V/C_{V_{eq}}$; $a_I = C_I/C_{I_{eq}}$; and $C_{V_{eq}}$, $C_{I_{eq}}$, C_V, and C_I are the equilibrium and nonequilibrium concentrations of vacancies and intrinsic interstitials, respectively. Under the conditions of local equilibrium, according to the law of mass action, $a_V = 1/a_I$. The contribution to the diffusion of different charge states of the intrinsic point defects is taken into account using the relative fractions of the vacancy and interstitial components of the diffusion, which, in the case of phosphorus in silicon, are [2] $f_I^0 = 0.95$ and $f_V^- = f_V^{2-} = 1$ ($f_V^z + f_I^z = 1$, where z is the charge of the intrinsic point defects). The degree of supersaturation with respect to the intrinsic interstitials in (4) is related to the oxidation rate by the empirical dependence [17]

$$a_I = 1 + aU_{ox}^b, \quad (5)$$

where U_{ox} is the thermal oxidation rate and a and b are parameters ($a = 1 \times 10^5$ (s/cm)$^{0.5}$ and $b = 0.5$).

The diffusion of dopants from the implanted Si layer is enhanced due to the excess of intrinsic point defects formed during the annealing of radiation defects introduced by the implantation. The effect of transient enhanced diffusion is taken into account by introducing a factor $A(t)$, the diffusion enhancement coefficient ($A(t) = a_V = a_I$), into diffusion coefficient (4) [18]:

$$A(t) = [1 + (A_0 - 1)\exp(-t/\tau)]. \quad (6)$$

Here, t is the annealing time; A_0 is the enhancement of the diffusion at the initial instant of time $t = 0$; and τ is the temperature-dependent characteristic time constant of the transient enhanced diffusion, $\tau = \tau_0\exp(E_\tau/kT)$.

The concentration distributions of phosphorus were analyzed within a model of the diffusion–segregation redistribution of the impurity in the bulk of the SiO_2/Si structure, which possesses moving gas/SiO_2 and SiO_2/Si interfaces with the coordinates $x_1(t)$ and $x_2(t)$, respectively. If the diffusive transport in SiO_2 is disregarded, the phosphorus redistribution can be described by the boundary-value problem for the equations

$$\partial C_{ox}/\partial t = \partial/\partial x((1-\alpha)U_{ox}C_{ox}), \quad x_1 < x < x_2, \quad (7)$$

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}\left[D\left(\frac{\partial C}{\partial x} - \frac{C}{C^*}\frac{\partial C^*}{\partial x}\right)\frac{\partial C}{\partial C_a}\right], \quad x_2 < x < l, \quad (8)$$

where x is a coordinate measured from the SiO_2/Si interface at the initial instant of time $x_2(0)$ to the silicon bulk; l is the width of the silicon solution region, which is chosen according to the condition for a constant impurity concentration during the annealing; α is the ratio of the thickness of the oxidized silicon layer to the thickness of the forming dioxide W_{ox} ($\alpha = 0.44$); and $U_{ox} = dW_{ox}/dt$ is the oxidation rate of silicon. An additional segregation term was introduced into the diffusion equation in [19] to take into account the gradient of the chemical potential of the diffusing particles. In the case under consideration, the second term in Eq. (8) describes the accumulation of phosphorus in the silicon surface region. Equation (8) takes into account a number of factors affecting the diffusion: (i) a high doping level (relation (3)), (ii) clusterization (the factor $\partial C/\partial C_a$); (iii) the thermal oxidation rate (relations (4) and (5)); and (iv) the effect of transient enhanced diffusion (the factor (6)).

System of equations (7) and (8) was supplemented with the condition of matching at the SiO_2/Si interface (a consequence of the law of conservation of the total amount of impurity passing through the interface as a result of its motion and due to the concentration gradient). Under the assumption of equilibrium segregation (with respect to the kinetic trapping), this condition takes the form

$$U_{ox}(1/m_s - \alpha)C = D\partial C/\partial C_a[\partial C_a/\partial x - C_a\partial(f(x)dE_b/kT)/\partial x], \quad (9)$$

where m_s is the segregation coefficient, determined as the ratio of the current impurity concentrations in Si and SiO_2 at the interface, and $m_s = C_s/C_{ox}$. The oxidation kinetics is described by the linear–parabolic dependence

$$t = W_{ox}^2/k_p + W_{ox}/k_l, \quad (10)$$

where k_p and k_l are the parabolic and linear constants of the silicon oxidation rate (taken from [20]). For the initial phosphorus profile after the implantation, we used the Pearson IV distribution, with the moments taken from [21].

The problem for equations (7) and (8) with boundary condition (9) was solved numerically by reducing it, with replacement of the variables, to a problem with immobile boundaries. The latter was solved by factorization using homogeneous implicit difference schemes [22].

3. SIMULATION RESULTS

Within the above-described model, we analyzed the experimental distributions of phosphorus in the SiO_2/Si system for different conditions of the thermal oxidation of the heavily (homogeneously and inhomogeneously) doped Si layers. The phosphorus segregation coefficient m_s was determined from the best correspondence between the solution to problem (7)–(9) and the experimental profiles of phosphorus in SiO_2 and Si.

In terms of the above model, the peak of surface phosphorus concentration is formed due to the diffusion of the impurity from the bulk to the surface under the action of the free energy gradient in the surface region. In this case, both the height and the width of the surface concentration peak are determined by the parameters of the surface free energy distribution and the phosphorus diffusion coefficient, which depends on the impurity concentration in the case of heavy doping. In Fig. 1, the effect of doping on the phosphorus distribution near the silicon surface is shown (curves *3*, *4*) in comparison with the experimental profiles (curves *1*, *2*) obtained in [9] by AES with sputtering of the sample surface by 1-keV Ne^+ ions. As can be seen from Fig. 1, the calculation within the model using the parameters listed in the caption to Fig. 1 makes it possible to adequately describe the phosphorus profiles near the oxidized silicon surface at different initial doping levels: $C_b = 2 \times 10^{20}$ cm^{-3} (curves *1*, *3*) and 8×10^{19} cm^{-3} (curves *2*, *4*). This discription also includes features such as the surface concentration peak and the depletion region near its base. The formation of the depletion

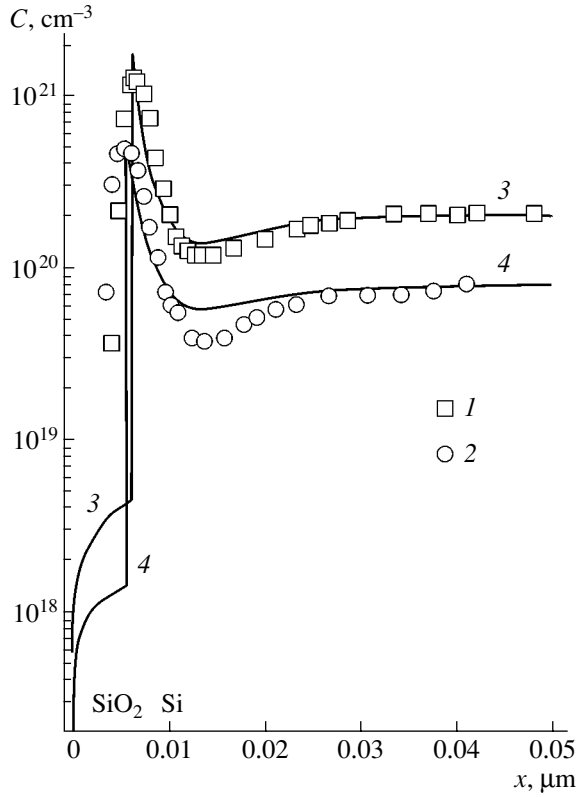

Fig. 1. Phosphorus distributions in the surface region of homogeneously doped Si(100) after thermal oxidation (T = 900°C, t = 10 min, dry O_2). The dopant concentration C_b = (*1*, *3*) 8×10^{19} and (*2*, *4*) 2×10^{20} cm^{-3}. (*1*, *2*) Experiment [9] and (*3*, *4*) calculation at dE_b = 0.215 eV, x_m = 3 nm, m_s = 380, and a_I is derived from (5).

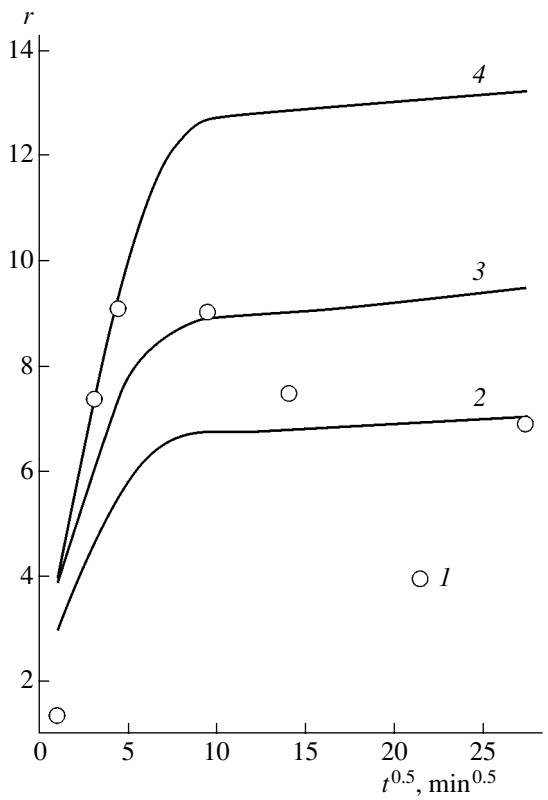

Fig. 2. Kinetics of the change in the relative height of the peak of surface P concentration during thermal oxidation (T = 900°C, dry O_2) of homogeneously doped Si(100) ($C_b = 2 \times 10^{20}$ cm^{-3}). (*1*) Experiment [9] and (*2*–*4*) calculation where x_m = 3 nm, m_s = 380, a_I is derived from (5), and dE_b = (*2*) 0.14, (*3*) 0.18, and (*4*) 0.22 eV.

region is due to an impurity outflow towards the region of the surface concentration peak.

The rate of the increase in the relative height of the surface concentration peak ($r = C_s/C_b$) within the above model is also determined by the parameters of the free energy distribution and the coefficient of the phosphorus diffusion near the silicon surface. The kinetics of changes in the relative height of the surface concentration peak in comparison with the experimental data of [9] is shown in Fig. 2 for different values of dE_b. As can be seen from Fig. 2, the calculated (for dE_b = const) height of the surface concentration peak attains maximum at 900°C for $t \approx 100$ min, after which it does not change, whereas the experimental value of the surface peak height drops with conservation of the total concentration in the peak. We believe that the decrease in the Auger signal related to the peak of surface phosphorus concentration at the interface and its broadening with an increase in the oxidation time and, accordingly, the oxide thickness is an instrumental effect. This effect is due (as was shown, for example, in [23]) to the increase in the microrelief caused by ion etching of the sample surface during the analysis.

The phosphorus distribution in the SiO_2/Si system after thermal oxidation of a homogeneously doped Si layer ($C_b = 6 \times 10^{20}$ cm^{-3}) is shown in Fig. 3, where it is compared to the experimental profiles obtained in [9] by AES (curve *1*) and SIMS using Cs^+ ions (curve *2*). As can be seen from Fig. 3 (curves *3*, *4*), taking into account the formation of the surface concentration peak makes it possible to adequately describe the experimental profile of phosphorus in the SiO_2/Si system at values of the parameters similar to those in Figs. 1 and 2.

The effect of taking into account the formation of the surface concentration peak on the phosphorus segregation in the case of the thermal oxidation of an inhomogeneously doped implanted layer is shown in Fig. 4. The SIMS profile from [13] is also shown for comparison. In this case, the phosphorus ions were implanted through a 25-nm dioxide layer and subsequently annealed at 900°C for 450 min in dry O_2 with 1% of HCl. As can be seen from Fig. 4, the calculation taking into account the formation of the surface concentration peak and its effect on the segregation ($m_s = C_s/C_{ox}$) makes it possible to describe the redistribution of phosphorus in the SiO_2/Si system more exactly than with the calculation disregarding the formation of the surface

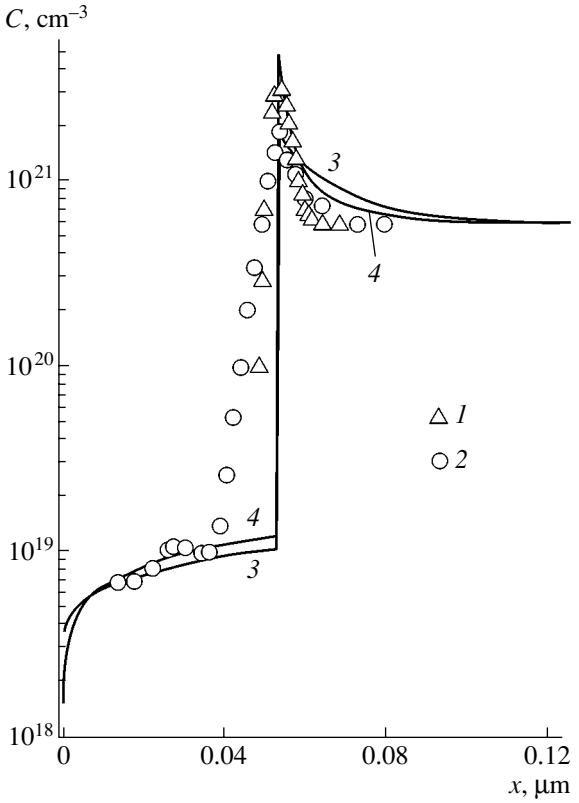

Fig. 3. Phosphorus distributions in the SiO_2/Si system after thermal oxidation (T = 900°C, t = 90 min, dry O_2) of homogeneously doped Si(100) ($C_b = 6 \times 10^{20}$ cm^{-3}). (*1*, *2*) Experiment [9] ((*1*) AES and (*2*) SIMS data), (*3*) calculation disregarding the formation of a surface concentration peak (m_s = 160), and (*4*) calculation taking into account the formation of a surface concentration peak (dE_b = 0.12 eV, x_m = 3 nm, and m_s = 380). a_I is derived from (5).

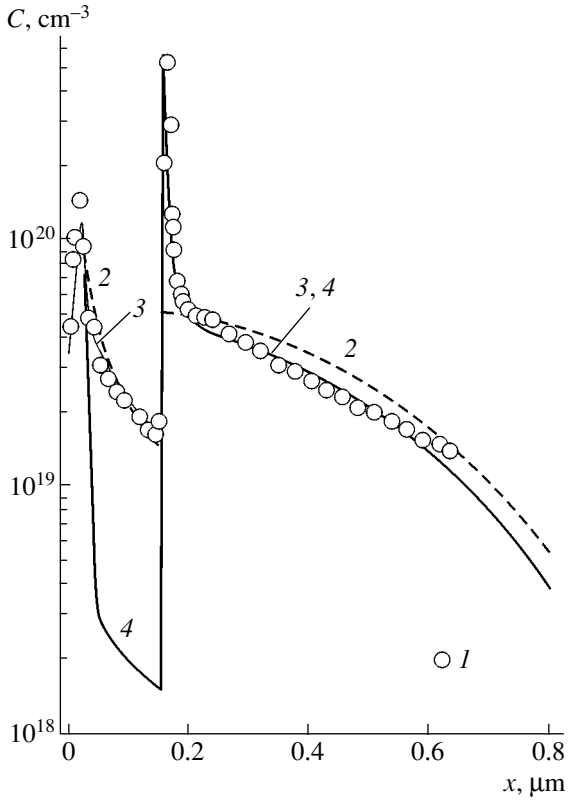

Fig. 4. Phosphorus distributions in the SiO_2/Si system after thermal oxidation (T = 900°C, t = 450 min, dry O_2 (1% HCl)) of an implanted Si(100) layer (E = 80 keV and $Q = 2.5 \times 10^{15}$ cm^{-2}). (*1*) Experiment [13], (*2*) calculation disregarding the formation of a surface concentration peak (m_s = 3.6), and (*3*, *4*) calculation taking into account the formation of a surface concentration peak (dE_b = 0.25 eV, x_m = 14 nm, m_s = 36). m_s = (*2*, *3*) C_s/C_{ox} and (*4*) C_{sb}/C_{ox}. a_I is derived from (5).

concentration peak (compare curves *3* and *2*). A large amount of the impurity passes from the silicon surface region to the region of the surface concentration peak during the annealing; therefore, disregard of the formation of the surface concentration peak in the simulation of the diffusion–segregation redistribution of phosphorus leads to an overestimation of the impurity concentration in silicon.

In contrast to this model, which uses the approximation of an ideally sharp interface, the model developed in [13] assumes that the surface concentration peak is formed due to the trapping of P atoms from Si and SiO_2 in some intermediate (transient) layer and that it does not affect the segregation coefficient. The latter is determined conventionally as the ratio $m_s = C_{sb}/C_{ox}$, where C_{sb} is the P concentration in bulk Si beyond the region of the surface concentration peak. These two approaches are shown in Fig. 4 (curves *3*, *4*). As can be seen from Fig. 4, the model proposed here describes the phosphorus profile in the SiO_2/Si system much better than in the case of the conventional determination of the segregation coefficient, which yields an underestimated impurity concentration in silicon dioxide.

When the oxidation of the implanted Si layers is of short duration, the effect of transient enhanced diffusion in the implanted layer should be taken into account. Figure 5 shows the SIMS profiles (obtained in this experiment) of phosphorus in the SiO_2/Si system after an implantation of P^+ ions (E = 100 keV and $Q = 7.5 \times 10^{15}$ cm^{-2}) and oxidation at 950°C for 20 min in wet O_2 (curve *1*) in comparison with the simulation both taking into account (using relation (6)) and disregarding the effect of transient enhanced diffusion (curves *2* and *3*, respectively). The magnitude of the effect of oxidation on the diffusion of phosphorus in silicon, estimated by the averaged degree of supersaturation with respect to the intrinsic interstitials $\langle a_I \rangle = \langle I \rangle / I_{eq}$, was taken from [24]. Analysis of the concentration profiles allowed us to determine the parameters of the distribution of the surface free energy of the P atoms, dE_b = 0.12 eV and x_m = 10 nm, and the parameters of the transient enhanced diffusion, A_0 = 31 and τ = 522 s.

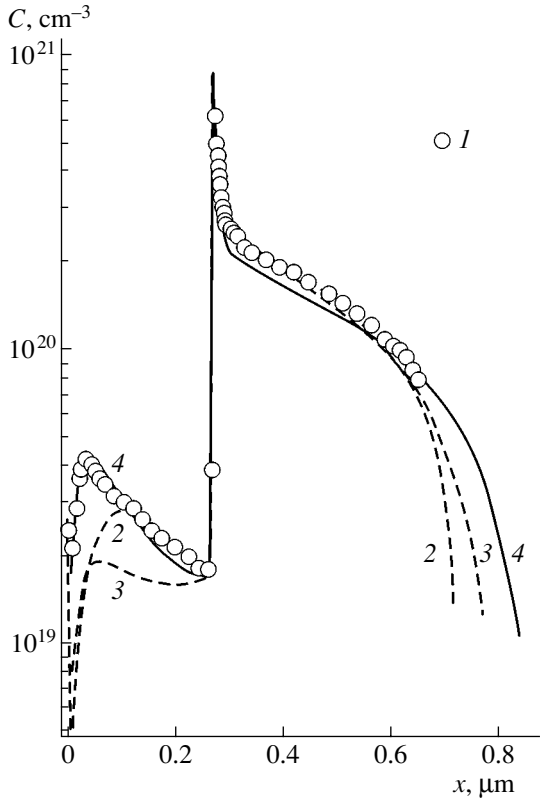

Fig. 5. Phosphorus distributions in the SiO_2/Si system after thermal oxidation (T = 950°C, t = 20 min, wet O_2) of an implanted Si(111) layer (E = 100 keV and $Q = 7.5 \times 10^{15}$ cm^{-2}). (*1*) Experiment (SIMS), (*2*) calculation disregarding the transient enhanced diffusion (m_{s0} = 50.5), (*3*) calculation taking into account the transient enhanced diffusion (m_{s0} = 50.5), and (*4*) calculation taking into account the transient enhanced diffusion and the dependence $m_s(t)$ according to (10) (dE_b = 0.12 eV, x_m = 10 nm, and m_{s0} = 50.5). a_I is taken from [24].

As can be seen from Fig. 5, the P profile cannot be adequately described using a constant segregation coefficient (m_s = 50.5 for curve *3*). As the analysis showed, a correspondence between the calculated concentration profiles in the SiO_2/Si system and the experimental distribution can be obtained using the time dependence of the segregation coefficient in the form (see curve *4*)

$$m_s(t) = m_{s0}/[1 + (B_0 - 1)\exp(-t/\tau)], \qquad (11)$$

where m_s0 is the segregation coefficient by the end of the process of transient enhanced diffusion, B_0 = 2.4, and the characteristic segregation time τ coincides with the corresponding value for the transient enhanced diffusion in expression (6).

Note that the time dependence of the segregation coefficient in form (11) makes it possible to describe the experimental P profiles in both Si and SiO_2 in the case of the thermal oxidation of ion-implanted Si layers in wet oxygen and at higher temperatures (in the range 950–1150°C). In this case, the temperature dependence of the characteristic time of transient enhanced diffusion and segregation has the form $\tau = 1.4 \times 10^{-7}\exp(2.33\ \text{eV}/kT)$ s.

4. RESULTS AND DISCUSSION

The redistribution of phosphorus during the thermal oxidation of homogeneously doped Si layers (Figs. 1–3) and inhomogeneously implanted Si layers subjected to long-term annealing ($t \gg \tau$) (Fig. 4) is adequately described by the proposed model provided that the formation of a surface concentration peak is taken into account and it is assumed that m_s = const. In contrast, when the formation of the surface concentration peak is disregarded, it is impossible to adequately describe the concentration profiles of P in the SiO_2/Si system (compare curves *3*, *4* and curves *2*, *3* in Fig. 4) in the case of the thermal oxidation of heavily doped Si layers. In a simulation of the phosphorus redistribution occurring during thermal oxidation that was carried out in [12–14], the segregation coefficient was determined conventionally as $m_s = C_{sb}/C_{ox}$, i.e., disregarding the formation of the surface concentration peak. This approach leads to an underestimation of the segregation flux in the dioxide, because the outflow of the impurity from the surface Si layer to the surface (into the region of forming surface concentration peak) is disregarded. At the same time, when the segregation coefficient is determined as $m_s = C_{sb}/C_{ox}$, i.e., the formation of the surface concentration peak is taken into account, we have, in contrast, an increase in the segregation flux into the dioxide due to the growth of the surface concentration peak in Si at the interface with SiO_2. Analysis of the concentration profiles of phosphorus in the SiO_2/Si system (Fig. 4, curves *1*, *3*, *4*) shows that it is the second approach that leads to good agreement between the experimental and calculated profiles not only in Si but also in SiO_2. This result indicates that the peak of surface phosphorus concentration is indeed in the surface Si layer (as was suggested in [8–10, 12]) rather than in the transient layer or in the dioxide layer near the interface (as was believed in [13, 14] and [7, 25, 26], respectively). It is the presence of this surface concentration peak in the surface Si layer that affects the chemical potential of P atoms and makes it possible to determine the impurity segregation coefficient as $m_s = C_s/C_{ox}$ rather than C_{sb}/C_{ox}.

In the case of the thermal oxidation of ion-implanted layers for a time no longer than the characteristic time of the transient enhanced diffusion ($t \le \tau$), it is necessary to take into account the time dependence of not only the enhancement coefficient of the transient enhanced diffusion $A(t)$ (compare curves *1* and *2* in Fig. 5) but also the segregation coefficient $m_s(t)$ (compare curves *3* and *4* in Fig. 5). The time dependence $m_s(t)$ in form (11) is inversely proportional to the enhancement of the transient diffusion in form (6) and has the same characteristic time constant τ. This result

indicates that the excess intrinsic point defects formed in the implanted layer during the annealing of radiation defects affects not only the transient enhanced diffusion in silicon but also the segregation transport of the impurity through the SiO_2/Si interface.

According to current concepts, thermal oxidation occurs in the reaction zone at the moving SiO_2/Si interface [27]. It is reasonable to suggest that, along with the Si atoms, dopant atoms passing into the reaction zone from silicon also become oxidized in this zone. The main component of the driving force behind this passage is dopant oxidation; i.e., the transport of impurity particles through the SiO_2/Si interface has a reaction character. Considering the mechanism involved in this reaction segregation, the impurity flux through the interface can be represented as an exchange of impurity atoms between the silicon phase and the reaction zone. In the same way as the Si atoms, the impurity atoms are oxidized in the reaction zone, with the resulting formation of structural fragments of their own oxides, which are then incorporated into the structure of silicon dioxide. Any process facilitating the oxidation of the impurity atoms leads to the enhancement of the segregation transport of these atoms through the interface from silicon to dioxide, which manifests itself in a decrease in the segregation coefficient m_s.

In the ion-implanted Si layers, one such process is the generation of intrinsic point defects during the annealing of the implantation-induced defects. Since the transient enhanced diffusion in implanted Si layers is observed both for the impurities diffusing predominantly via the indirect interstitial mechanism (B, P) and for the impurities diffusing via the vacancy mechanism (Sb) [18], it can be concluded that supersaturation with respect to both the intrinsic interstitials and the vacancies occurs. The vacancies are carriers of a free volume V_F that is necessary for solid-state reactions of the oxidation of both silicon and the impurity atoms, since the molecular volumes of the reaction products, silicon dioxide ($\Omega(SiO_2) \cong 0.044$ nm^3) and impurity oxide ($\Omega(P_2O_5) \cong 0.099$ nm^3), exceed the atomic volume of the starting reagent, namely, silicon ($\Omega(Si) \cong 0.02$ nm^3). This consideration is especially true for the oxidation of the impurity atoms, since $\Omega(P_2O_5) > \Omega(SiO_2)$. Therefore, it is reasonable to suggest that the coefficient of the distribution between the oxidized and unoxidized impurity components in the reaction zone, $K = C_{oxn}/C_{oxa}$, is proportional to the relative concentration of vacancies at the interface originating from the side of SiO_2, $a_{V_{ox}}$:

$$K = K_0 a_{V_{ox}}. \quad (12)$$

Here, K_0 is the distribution coefficient in the state of equilibrium with respect to the intrinsic point defects (at $a_{V_{ox}} = 1$). The segregation coefficient, determined experimentally as a discontinuity in the concentration distribution at the interface, is the ratio of the total impurity concentrations from different sides of the interface: $m_s = C_s/C_{ox}$. Taking into account expression (12) for the distribution coefficient K, we can write m_s in the form

$$m_s = \frac{C_s}{C_{ox}} = \frac{C_s}{C_{oxa} + C_{oxn}} = \frac{C_s}{C_{oxa}(1 + K)}. \quad (13)$$

Hence, at $K \gg 1$, we obtain the dependence

$$m_s \cong m_s^*/a_{V_{ox}}, \quad (14)$$

where m_s^* is the equilibrium segregation coefficient (with respect to the intrinsic point defects) for an unperturbed state of the ensemble of intrinsic point defects (at $a_{V_{ox}} \approx 1$) and m_s is the nonequilibrium segregation coefficient (at $a_{V_{ox}} \neq 1$).

The diffusion coefficients of the vacancies and Si interstitials in both silicon and silicon dioxide are rather large, which gives grounds to assume the segregation of the intrinsic point defects at the SiO_2/Si interface to be equilibrium. Therefore, the concentrations of intrinsic interstitials and vacancies normalized to the corresponding equilibrium concentrations (degrees of supersaturation with respect to the intrinsic point defects) from different sides of the interface are equal to each other; i.e., $a_{I_{ox}} = a_{I_s}$ and $a_{V_{ox}} = a_{V_s}$. This circumstance may be the reason for the correlation between the time dependence of the transient enhanced diffusion in silicon and the segregation coefficient of phosphorus at the SiO_2/Si interface. The difference between the amplitudes of the time dependences for the transient enhanced diffusion and the segregation ($A_0 = 31$ in (6) and $B_0 = 2.4$ in (11)) can be explained by the mutual recombination of excess vacancies and intrinsic interstitials (diffusing from the region of implantation defects in silicon) at the interface.

5. CONCLUSIONS

A model of the diffusion–segregation redistribution of phosphorus in an SiO_2/Si system is developed for the case of thermal oxidation of heavily doped Si layers. This model takes into account the formation of a peak of surface impurity concentration at the SiO_2/Si interface along with the effects of high impurity concentrations, clusterization of impurities, thermal oxidation of silicon, and implantation-induced defects on diffusion. The formation of the surface concentration peak is attributed to a change in the free energy of the P atoms near the Si surface and is modeled by a diffusion–segregation equation. It is shown by an analysis of the experimental data within the model proposed that, taking into account the formation of the surface concentration peak, it is possible to describe the specific features of phosphorus redistribution during the thermal oxidation of homogeneously and inhomogeneously heavily doped Si layers. In the case of the thermal oxidation of

ion-implanted Si layers, it is necessary to take into account the nonequilibrium state of the ensemble of intrinsic point defects caused by the annealing of implantation-induced defects. This state leads not only to the transient enhanced diffusion of phosphorus in the silicon bulk but also affects the accumulation of impurity atoms in the silicon surface region. It was ascertained that, in this case, the coefficient for the segregation of phosphorus at the SiO_2/Si interface is not constant but, in a similar way to the transient enhanced diffusion, depends on time. This phenomenon is explained by the reactive character of the impurity segregation during the thermal oxidation of silicon, at which the nonequilibrium intrinsic point defects in the implanted Si layer affect the processes of oxidation and trapping of the impurity atoms by the growing silicon dioxide. As a result, the impurity segregation coefficient for the SiO_2/Si system is found to be a function of the relative concentration of vacancies at the interface.

REFERENCES

1. R. B. Fair, in *Impurity Doping Processes in Silicon*, Ed. by F. F. Y. Wang (North-Holland, New York, 1981).
2. O. V. Aleksandrov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 1289 (2001) [Semiconductors **35**, 1231 (2001)].
3. O. V. Aleksandrov, N. V. Ashkinadze, and R. Z. Tumarov, Fiz. Tverd. Tela (Leningrad) **26**, 632 (1984) [Sov. Phys. Solid State **26**, 384 (1984)].
4. S. M. Hu, J. Appl. Phys. **45**, 1567 (1974).
5. R. B. Fair, J. J. Wortman, and J. Lin, J. Electrochem. Soc. **131**, 2387 (1984).
6. O. V. Aleksandrov and N. N. Afonin, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 12, 97 (1990).
7. N. J. Chou, Y. J. Van Der Meulen, R. Hammer, and J. Cahill, Appl. Phys. Lett. **24**, 200 (1974).
8. J. S. Johannessen, W. E. Spicer, J. F. Gibbons, and J. D. Plummer, J. Appl. Phys. **48**, 4453 (1978).
9. S. A. Schwarz, R. W. Barton, C. P. Ho, and C. R. Helms, J. Electrochem. Soc. **128**, 1101 (1981).
10. O. V. Aleksandrov, N. N. Afonin, and A. P. Kovarskiĭ, Élektron. Tekh., Ser. 6: Mater., No. 4 (241), 73 (1989).
11. O. V. Aleksandrov and N. N. Afonin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 1570 (1996) [Semiconductors **30**, 823 (1996)].
12. K. Sakamoto, K. Nishi, F. Ichikawa, and S. Ushio, J. Appl. Phys. **61**, 1553 (1987).
13. F. Lau, L. Mader, C. Mazure, *et al.*, Appl. Phys. A **49**, 671 (1989).
14. M. Orlowski, Appl. Phys. Lett. **55**, 1762 (1989).
15. C. P. Ho, J. D. Plummer, S. E. Hansen, and R. W. Dutton, IEEE Trans. Electron Devices **30**, 1438 (1983).
16. P. M. Fahey, P. B. Griffin, and J. D. Plummer, Rev. Mod. Phys. **61**, 289 (1989).
17. N. Jeng and S. T. Dunham, J. Appl. Phys. **72**, 2049 (1992).
18. O. V. Aleksandrov and D. S. Fedorov, Izv. S.-Peterb. Gos. Élektrotekh. Univ., No. 1, 16 (2002).
19. H. M. You, U. Gosele, and T. Y. Tan, J. Appl. Phys. **74**, 2461 (1993).
20. E. A. Taft, J. Electrochem. Soc. **136**, 3476 (1989).
21. A. F. Burenkov, F. F. Komarov, M. A. Kumakhov, and M. M. Temkin, *Spatial Distributions of the Energy Released in Cascades of Atomic Collisions in Solids* (Énergoizdat, Moscow, 1985) [in Russian].
22. A. A. Samarskiĭ, *The Theory of Difference Schemes*, 2nd ed. (Nauka, Moscow, 1983; Marcel Dekker, New York, 2001), Chap. 3, p. 141.
23. J. Y. Wang, A. Zalar, and E. J. Mittemeijer, Appl. Surf. Sci. **222**, 171 (2004).
24. O. V. Aleksandrov and N. N. Afonin, Zh. Tekh. Fiz. **73** (5), 57 (2003) [Tech. Phys. **48**, 580 (2003)].
25. Y. Sato, K. Imai, and N. Yabumoto, J. Electrochem. Soc. **144**, 2548 (1997).
26. P. B. Griffin, S. W. Crowder, and J. M. Knight, Appl. Phys. Lett. **67**, 482 (1995).
27. S. M. Repinskiĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 1050 (2001) [Semiconductors **35**, 1006 (2001)].

Translated by Yu. Sin'kov

Semiconductors, Vol. 39, No. 6, 2005, pp. 623–628. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 655–659.

ATOMIC STRUCTURE AND NONELECTRONIC PROPERTIES OF SEMICONDUCTORS

Thermodynamic Stability and Redistribution of Charges in Ternary AlGaN, InGaN, and InAlN Alloys

V. G. Deĭbuk^ and A. V. Voznyĭ^^

Fed'kovich National University, ul. Kotsyubinskogo 2, Chernovtsy, 58012 Ukraine

^e-mail: vdei@chnu.cv.ua

^^e-mail: vozny@elinity.com

Submitted October 14, 2004; accepted for publication November 10, 2004

Abstract—A model of the delta lattice parameter is used to study the thermodynamics of AlGaN, InGaN, and InAlN alloys. The phase diagrams obtained indicate that $Al_xGa_{1-x}N$ is stable in the entire range of x, whereas the miscibility gap corresponds to $0.2 < x < 0.69$ for $In_xGa_{1-x}N$ and to $0.16 < x < 0.7$ for $In_xAl_{1-x}N$ at 1000 K. Biaxial stresses lower the critical temperature and narrow the miscibility gap. The charge-density distribution is analyzed using the pseudopotential method to obtain an approximation of 32-atom supercells. The results of the analysis show that the stability of these alloys is controlled by the competition between the destabilizing contribution of strains related to the mismatch between the lattice constants and a stabilizing charge exchange between various chemical bonds. Biaxial stress reduces the charge redistribution caused by strains and thus increases the stability of an alloy. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

In recent years, considerable progress has been made in the fabrication of optical and electronic devices based on nitrides of Group III elements (AlN, GaN, and InN). These devices primarily include active optoelectronic devices that operate in the green, blue, and ultraviolet regions of the spectrum, and also high-frequency and high-temperature electronic devices, such as field-effect transistors [1]. A feature common to all these structures is the use of ternary $In_xGa_{1-x}N$ or $Al_xGa_{1-x}N$ alloys. The formation of alloys from nitrides of Group III elements basically makes it possible to vary the band gap from 0.7 eV in InN to 6.28 eV in AlN, with an intermediate value of 3.44 eV in GaN (at 300 K) [2].

However, the presence of the miscibility gap makes it impossible to obtain these alloys with an arbitrary composition of components. In particular, the critical temperature of decomposition in an $Al_xGa_{1-x}N$ system possessing a small mismatch of its lattice constants is fairly low, which makes it possible to obtain a more or less regular solid solution in the entire range of compositions under normal conditions. At the same time, fairly wide miscibility gaps have recently been experimentally observed in the $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$ alloys [3].

Under normal conditions, AlN, GaN and InN crystals exhibit a wurtzite hexagonal structure. Ternary alloys can also be grown into wurtzite-type crystals, irrespective of the method of deposition: molecular-beam epitaxy (MBE), vapor-phase epitaxy from metal–organic compounds (MOVPE), hydride vapor-phase epitaxy (HVPE), and so on. It has recently been shown that thin AlN, GaN, and InN films with a zinc blende cubic structure can be obtained [4] on GaAs (001) and 3*C*-Si (001) substrates using a plasma MBE method. In addition, layers of $Al_xGa_{1-x}N$, $In_xGa_{1-x}N$, and $In_xAl_{1-x}N$ ternary compounds possessing this cubic structure have been synthesized [3]. A characteristic feature of these nitrogen-containing alloys is the presence of internal stresses related to the lattice mismatch and to a difference in the thermal-expansion coefficients. For example, the mismatch of the GaN and AlN lattice constants amounts to 2.5 and 3.9% for the *a* and *c* hexagonal directions, respectively. The lattice mismatch increases to 10.7 and 9% for InN and GaN, while this mismatch is as large as 13.5 and 12.6% for AlN and InN. The thermal-expansion coefficients vary from 5.6×10^{-6} to 4.2×10^{-6} K^{-1} in the direction of the *a* axis and from 3.2×10^{-6} to 5.3×10^{-6} K^{-1} for hexagonal GaN and InN [5]. A large difference between the above-listed parameters gives rise to appreciable internal stresses and, as a consequence, to the miscibility gap for ternary alloys that involve nitrogen. These stresses and composition fluctuations affect the majority of physical properties, including solubility.

In this paper, we report the results of studying the thermodynamics of stressed $Al_xGa_{1-x}N$, $In_xGa_{1-x}N$, and $In_xAl_{1-x}N$ alloys and the effect of biaxial stresses originating in the substrate on diagrams of the spinodal decomposition in thin films of the compounds under consideration. When performing the calculations, we assumed that the pseudobinary nitride alloys form crystals that have a tetrahedrally coordinated cubic lattice with a zinc blende structure. The results of the thermodynamic consideration of these systems can be also generalized to a tetragonal system. We also assumed

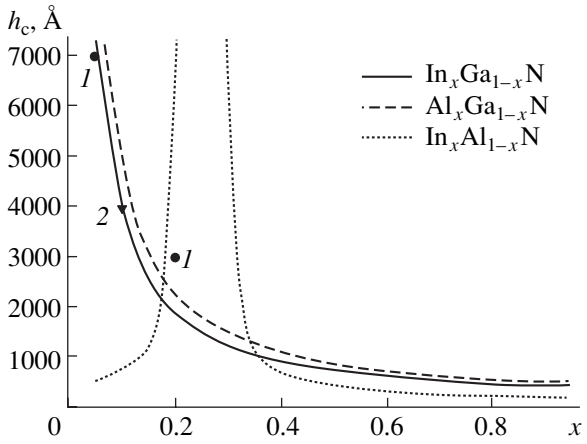

Fig. 1. Dependences of the critical thickness h_c on the composition of the epitaxial films: the lines represent the results of calculations, circles *1* correspond to experimental data for $Al_xGa_{1-x}N$/GaN [11], and triangle *2* corresponds to the experiment for $In_xGa_{1-x}N$/GaN [11].

that the Ga, Al, and In atoms are distributed randomly over the sites of the cationic face-centered sublattice but that the N atoms reside in the other sublattice. Although the calculations reported below are carried out for a cubic phase, similar results can also be expected for a wurtzite phase [6]. We also calculated the spatial distribution of the charge, taking into account the alloying effects in the approximation of 32-atom supercells. This approach allowed us to analyze the charge-transport processes that occur during the formation of the alloys and also to gain insight into the phenomenon of thermodynamic instability of ternary nitrogen-containing alloys from the standpoint of a microscopic approach.

2. AN ANALYSIS OF THERMODYNAMIC STABILITY

2.1. $Al_xGa_{1-x}N$

Substitutional $Al_xGa_{1-x}N$ solid solutions are experimentally obtained regardless of the composition (for all values of x) in the form of both bulk crystals and epitaxial layers; therefore, the thermodynamic analysis that we perform can be considered as a check of sorts.

The thermodynamics of $Al_xGa_{1-x}N$ stability has been studied theoretically using various methods (see [6–9]). In particular, the critical temperature T_c of spinodal decomposition has been found to equal 368 K when the regular-solution model is applied [9], whereas estimation using a generalized quasi-chemical approximation (GQCA) yields a value that is obviously too small (T_c = 87 K) [6]. In our calculations, we used the delta lattice-parameter (DLP) model, as it yields good results for alloys based on III–V compounds, and generalized it to epitaxial films. We provided a detailed description of this model in [10]. The critical thickness (h_c), calculated within the modified Matthews–Blakeslee approximation, of $Al_xGa_{1-x}N$ films in relation to their composition is shown in Fig. 1 and is in good agreement with experimental data: 300 nm < h_c < 700 nm at $0.05 < x < 0.2$ [3, 11].

We calculated the critical temperature as T_c = 107 K at $x = x_c = 0.53$, which explains why there is no miscibility gap in this alloy at the characteristic growth temperatures. As the film thickness decreases, the significant contribution of biaxial stresses to the strain energy brings about a decrease in the critical decomposition temperature and a narrowing of the miscibility gap.

2.2. $In_xGa_{1-x}N$

In contrast to an $Al_xGa_{1-x}N$ system, the $In_xGa_{1-x}N$ alloys exhibit a fairly wide miscibility gap. These alloys have been studied rather extensively in recent years, both experimentally and theoretically [6–8, 12–14]. The phase separation observed at the growth temperatures is caused by internal strains that arise as a result of mixing the two components (GaN and InN) possessing mismatched lattice constants. As a consequence, the limiting solubility of InN (GaN) in GaN (InN) is lower than 5% (10%) at the typical temperature of alloy growth T = 1000 K [15]. However, the phase diagram also shows that a mismatched alloy can exist in the form of a metastable solid solution in a wide interval between the binodal and spinodal curves. Figure 1 shows the results of our calculations of the critical thickness h_c of a thin $In_xGa_{1-x}N$/GaN film as a function of the composition. The calculated value of h_c at $x = 0.1$ is in good agreement with the corresponding experimental value (h_c = 400 nm) [3]. In the context of the DLP model, we calculated the curves of the binodal and spinodal decomposition for $In_xGa_{1-x}N$/GaN films with different thicknesses (Fig. 2). As can be seen, a film with the thickness h = 1 μm is completely relaxed. It is worth noting that, in this case, T_c = 1322 K at x_c = 0.44, which is in satisfactory agreement with the results of theoretical calculations performed in the context of the GQCA [6] (T_c = 1295 K) and valence-force field [15] (T_c = 1473 K) models. According to our calculations of the phase diagram for a relaxed film at T = 1000 K, the range of the binodal decomposition equals $0.09 < x_b < 0.86$ (as compared to $0.05 < x_b < 0.90$ in [6]) and that of the spinodal decomposition is $0.20 < x_s < 0.69$ ($0.16 < x_s < 0.75$ in [6]).

Experimental data [3, 16] show that the spinodal decomposition sets in at an In content higher than 30%. As the film thickness decreases, the effect of the biaxial stresses originating in the substrate should progressively manifest itself in the shape of the decomposition curves (Fig. 2). In particular, the critical decomposition temperature decreases, while the spinodal decomposition region shifts to higher indium concentrations. For example, the decomposition range is given by 0.76 <

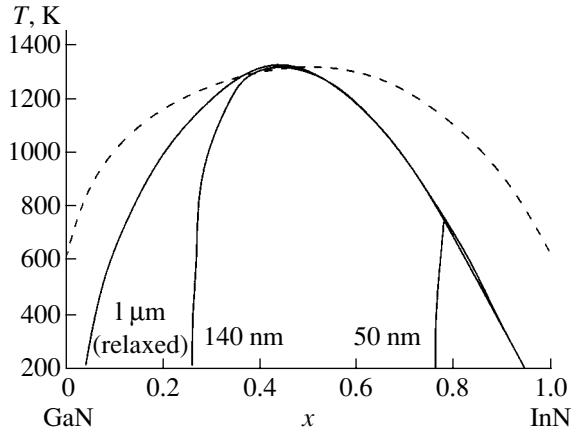

Fig. 2. Phase diagrams of the spinodal decomposition in epitaxial $In_xGa_{1-x}N$/GaN films of different thicknesses. The 1-μm-thick film is not stressed. The dashed line corresponds to the binodal curve, and the solid lines correspond to spinodal curves.

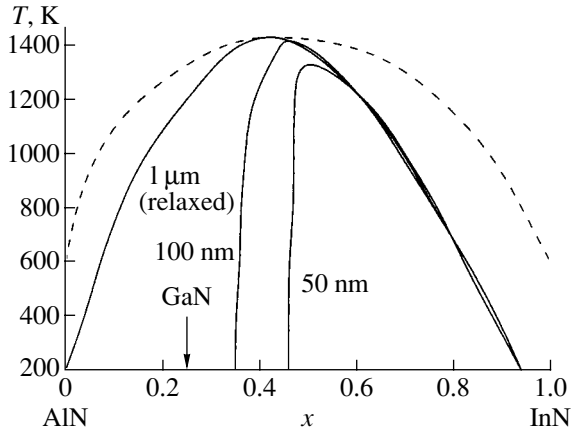

Fig. 3. Phase diagrams of the spinodal decomposition in $In_xGa_{1-x}N$/GaN films of different thicknesses. The 1-μm-thick film is not stressed. The dashed line corresponds to the binodal curve, and the solid lines correspond to the spinodal curves.

$x_s < 0.83$ according to our calculations for a 50-nm-thick film; at the same time, experimental data show that In at a content of up to 80% can be found in the solution for such films [17].

2.3. $In_xAl_{1-x}N$

In contrast to the $Al_xGa_{1-x}N$ and $In_xGa_{1-x}N$ compounds, the $In_xAl_{1-x}N$ alloy has been studied much less extensively due to the difficulties encountered in growing this system [18, 19]. The wide range of variations in its band gap makes it possible to hope that this alloy will eventually find a wide number of applications, in particular, as an insulating barrier for devices based on GaN. The appreciable mismatch of the InN and AlN lattice parameters gives rise to considerable internal stresses and it is these that represent the main cause of thermodynamic instability in an $In_xAl_{1-x}N$ alloy.

We calculated the Helmholtz free energy $\Delta F(x, T)$ of the alloy. The results of this calculation made it possible to draw the T–x phase diagram and determine the critical immiscibility temperature $T_c = 1450$ K at $x_c = 0.42$. Theoretical estimations using the GQCA method yield $T_c = 1485$ K [20]. The results of our calculations are shown in Fig. 3 and indicate that, in the alloy under consideration, the phase separation should be observed in a wide range of concentrations for the typical growth temperatures. For example, at $T = 1000$ K, the spinodal decomposition of relaxed films should occur if the In content varies from 16 to 70% (in comparison, this content ranges from 15 to 70% in the GQCA method [20]), whereas, in relaxed $In_xAl_{1-x}N$/GaN films, the phase separation has been experimentally observed to begin at $x > 0.17$ [19].

If the substrate-induced biaxial stresses in thin $In_xAl_{1-x}N$/GaN films are taken into account, the critical temperature decreases; however, this decrease is not as large as in the case of $In_xGa_{1-x}N$/GaN films. In addition, the miscibility gap is narrowed. This effect is particularly rapid on the substrate side of the phase diagram, which is in satisfactory agreement with the data of experimental studies of corresponding films [18].

3. REDISTRIBUTION OF THE VALENCE-ELECTRON CHARGE DENSITY

In order to gain more reliable information about the stability of the alloys, we can use self-consistent ab initio calculations of the total energy. However, these calculations are rather time-consuming, especially for disordered alloys, which are simulated using supercells with a large number of atoms. The causes of the instability of the alloys on a microscopic level can be studied by analyzing the charge-density distribution of the valence electrons, since the total energy is closely related to this distribution.

The calculations were carried out using a model empirical pseudopotential and took into account internal local strains, compositional disorder, and structural relaxation; supercells containing 32 atoms were used in the simulation. The procedure for selecting the pseudopotentials and the method of calculation for the approximation of the supercells were described in detail in [21, 22].

Let us consider the process involved in the formation of an alloy, which can be divided into several stages, using the example of InGaN. In Fig. 4, we show the distribution of the charge density ρ in binary GaN and InN compounds. The maximum of the charge density in InN is larger than in GaN and is located closer to the anion, which corresponds to the higher degree of ionicity present in InN.

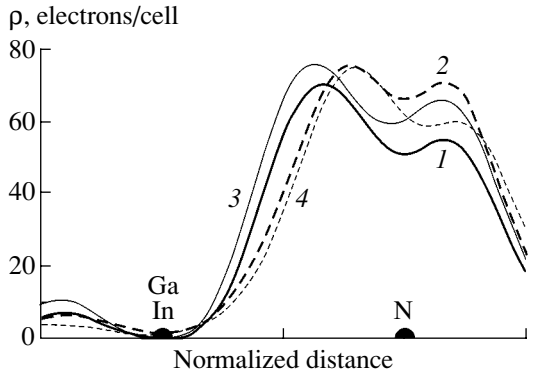

Fig. 4. The charge-density distribution between different bonds during formation of the InGaN alloy: (*1*, *2*) in the GaN and InN binary alloys, (*3*, *4*) in the $In_{0.5}Ga_{0.5}N$ ternary alloy, (*1*, *3*) along the Ga–N bonds, and (*2*, *4*) along the In–N bonds.

The first stage of the alloy's formation under simulation consists in bringing the lattice constants of the binary compounds into correspondence with that of the resulting ternary alloy, which means that the InN lattice should be compressed whereas the GaN lattice should be stretched. As a result, the charge density along the Ga–N bonds increases appreciably, which corresponds to an increase in the degree of ionicity, and becomes higher than that of InN, where the charge density along the bonds and the degree of ionicity decrease under the effect of strain.

In the second stage, we join the GaN and InN cells (so that the ternary alloy is formed) without relaxation of the bond lengths and angles. Since the electronegativity of the Ga–N bonds differs from that of the In–N bonds, a charge exchange should occur between these bonds. The charge is transferred from the Ga–N bonds, with a lower degree of ionicity, to the In–N bonds, with a higher degree of ionicity, which leads to a decrease in the difference between the charge distribution along these bonds. In the final stage, we take into account the relaxation of the bond lengths, which leads to a shift of atoms from the ideal positions specified by Vegard's law [22–24]. In addition, the strains arising from this shift result in the charge redistribution.

Thus, the net effect of all the factors under consideration leads to the equalization of the charge distributions along different bonds and to the fact that the properties of an actual alloy approach those of the virtual crystal (Fig. 4, curves *3*, *4*). We recently performed a detailed study of the influence of the alloying effects under consideration on the electronic spectrum of the alloys (in particular, on the dependence of the band gap on composition) [22].

The total variation in the charge density ($\Delta\rho$) along the bonds found as a result of considering the ternary alloy instead of a binary alloy is illustrated in Fig. 5. For example,

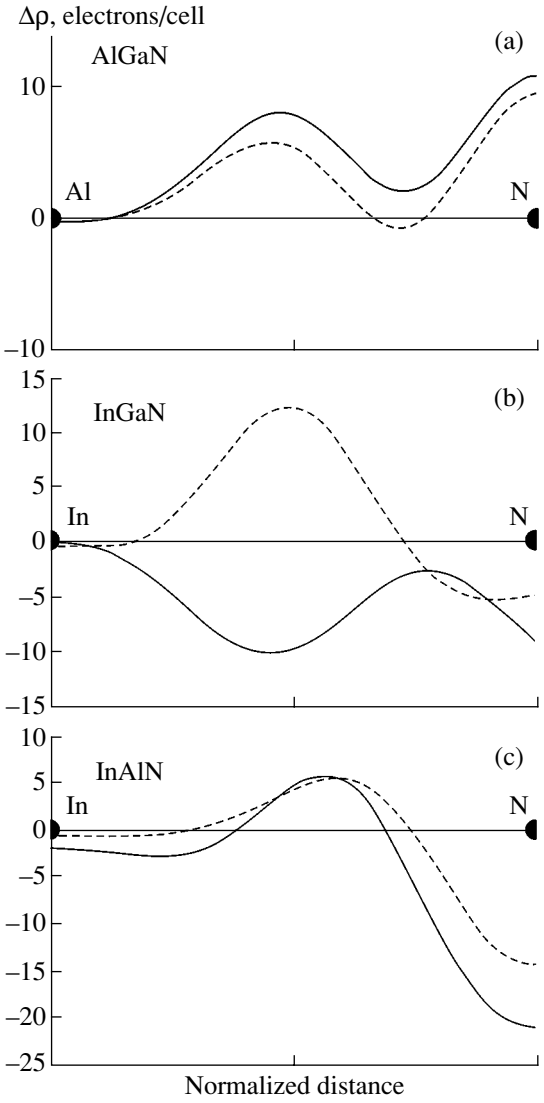

Fig. 5. The charge-density redistribution along the bonds with a higher degree of ionicity if the binary compounds are replaced by ternary alloys: (a) $Al_{0.5}Ga_{0.5}N$, (b) $In_{0.5}Ga_{0.5}N$, and (c) $In_{0.5}Al_{0.5}N$. The solid lines correspond to the bulk samples, and the dashed lines correspond to pseudomorphic films on a GaN substrate.

$$\Delta\rho(\text{In–N}) = \rho(\text{along the In–N bonds in InGaN}) - \rho(\text{InN}).$$

As ab initio calculations have shown [24], a transfer of charge from a bond with a higher ionicity to a bond with a lower ionicity leads to the instability of the system. In the nitrides of Group III elements under discussion, the strain caused by bringing the lattice constants of the alloys' components into correspondence with the average lattice constant is accompanied by a charge transfer from the bonds with a higher degree of ionicity to the bonds with a lower degree of ionicity. In contrast,

the difference between the electronegativities and structural relaxation bring about the transfer of charge from the bonds with a lower degree of ionicity to the bonds with a higher degree of ionicity; as a result, the system is stabilized. The difference between the electronegativities and structural relaxation in AlGaN prevails over the contribution of the strain; as a result, the net charge transfer occurs from the Ga–N bonds, with a lower degree of ionicity, to the Al–N bonds, with a higher degree of ionicity (Fig. 5a). In this study, to model the ionicity characteristic, we use the Garcia–Cohen coefficient of charge asymmetry [25]. In contrast to other models of ionicity, this coefficient indicates that the ionicity is higher in AlN than in GaN [22, 25, 26]: g(GaN) = 0.54, g(AlN) = 0.61, and g(InN) = 0.64. In InGaN and InAlN (Figs. 5b, 5c), the charge transfers from the bonds with a higher degree of ionicity to the bonds with a lower degree of ionicity; i.e., these compounds are unstable. This instability is caused by the much stronger effect of the strains resulting from the lattice-constant mismatch (as large as 10–12%).

In Fig. 5, the dashed lines show similar results obtained for thin pseudomorphic films of ternary alloys consisting of nitrides of Group III elements and grown on GaN substrates. Our calculations show that the biaxial stresses arising in the films appreciably reduce the charge redistribution caused by strains but that the chemical transport is nearly unaffected by these stresses. The distribution of $\Delta\rho$ along the bonds with a higher degree of ionicity (i.e., the Al–N bonds) varies only slightly if the bulk $Al_{0.5}Ga_{0.5}N$ samples are replaced by pseudomorphic films of the same material (Fig. 5a), and the values of $\Delta\rho$ remain positive, which indicates that the alloy under consideration is stable. The pattern is changed most significantly in the case of $In_{0.5}Ga_{0.5}N$ (Fig. 5b): as the alloy is formed in the pseudomorphic film, the charge is generally redistributed from the Ga–N bonds, with a lower degree of ionicity, to the In–N bonds, with a higher degree of ionicity; it is worth noting that $\Delta\rho > 0$ for almost the entire length of the In–N bonds, which indicates that the alloy is stable for the composition under consideration. In the $In_{0.5}Al_{0.5}N$ films (Fig. 5c) (as in the bulk material), the charge is generally redistributed from the In–N bonds, with a higher degree of ionicity, to the Al–N bonds, with a lower degree of ionicity. As a result, the alloy remains unstable, even in the form of a film, for the composition under consideration. However, the absolute value of $\Delta\rho$ in the film is lower, which gives grounds to expect that the concentration range where the alloy can become stable widens.

Thus, the concentration range where the chemical charge transport prevails over the strain contribution extends; as a result, we can state that the thin films of the alloys under consideration are more stable.

4. CONCLUSION

In this paper, we reported the results obtained from studying the thermodynamics of the stability of alloys that consisted of nitrides of Group III elements. The calculations that we carried out within a model of the delta lattice parameter show that AlGaN alloys are stable in the entire range of concentrations at typical growth temperatures; in contrast, InGaN and InAlN alloys exhibit a miscibility gap, which is in good agreement with the available experimental data. It is shown that the biaxial stresses arising in thin films lead to a narrowing of the miscibility gaps and to a decrease in the critical temperatures.

An analysis of the charge-density redistribution along the different bonds in the alloys made it possible to study, on the microscopic level, the processes that occur during the alloys' formation. It is established that the stability of the alloys depends not only on the difference between the lattice constants of the binary compounds but also on the charge exchange between bonds with different degrees of ionicity.

REFERENCES

1. S. Nakamura, *Introduction to Nitride Semiconductor Blue Lasers and Light Emitting Diodes* (Taylor and Francis, London, 2000).
2. J. Wu, W. Walukiewicz, and K. M. Yu, Solid State Commun. **127**, 411 (2003).
3. S. C. Jain, M. Willander, J. Narayan, and R. Van Overstraeten, J. Appl. Phys. **87**, 965 (2000).
4. M. J. Paisley, Z. Sitar, J. B. Posthil, and R. F. Davis, J. Vac. Sci. Technol. A **7**, 701 (1989).
5. A. Chen and A. Sher, *Semiconductor Alloys: Physics and Material Engineering* (Plenum, New York, 1995).
6. L. K. Teles, J. Furthmuller, L. M. Scolfaro, *et al.*, Phys. Rev. B **62**, 2475 (2000).
7. T. Ito, Phys. Status Solidi B **217**, R7 (2000).
8. D. Doppalapudi, S. N. Basu, K. F. Ludwig, and T. D. Moustakas, J. Appl. Phys. **84**, 1389 (1998).
9. V. A. Elyukhin and S. A. Nikishin, Semicond. Sci. Technol. **11**, 917 (1996).
10. V. G. Deĭbuk, Fiz. Tekh. Poluprovodn. (St. Petersburg) **37**, 1179 (2003) [Semiconductors **37**, 1151 (2003)].
11. I. Akasaki and H. Amano, in *GaN*, Ed. by J. I. Pankove and T. D. Moustakas (Academic, New York, 1998), Vol. 1, p. 459.
12. A. N. Westmeyer and S. Mahajan, Phys. Status Solidi B **228**, 161 (2001).
13. A. Kaschner, A. Hoffmann, C. Thomsen, *et al.*, Phys. Status Solidi A **179**, R4 (2000).
14. Y.-T. Moon, D. J. Kim, and K.-M. Song, Phys. Status Solidi B **210**, 167 (1999).
15. I. Ho and G. B. Stringfellow, Appl. Phys. Lett. **69**, 2701 (1996).
16. T. Matsuoka, J. Cryst. Growth **189**, 19 (1998).

17. R. Singh and T. D. Moustakas, in *Proceedings of First International Symposium on Gallium Nitride and Related Materials* (Pittsburgh, PA, 1996), Mater. Res. Soc. Symp. Proc., p. 163.
18. K. Murano, T. Inushima, Y. Ono, *et al.*, Phys. Status Solidi B **228**, 31 (2001).
19. S. Yamaguchi, M. Kariya, and S. Nitta, J. Cryst. Growth **195**, 309 (1998).
20. L. K. Teles, L. M. R. Scolfaro, J. Furthmuller, *et al.*, Phys. Status Solidi B **234**, 956 (2002).
21. V. G. Deĭbuk, A. V. Voznyĭ, M. M. Sletov, and A. M. Sletov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 420 (2002) [Semiconductors **36**, 398 (2002)].
22. A. V. Voznyĭ and V. G. Deĭbuk, Fiz. Tekh. Poluprovodn. (St. Petersburg) **38**, 316 (2004) [Semiconductors **38**, 304 (2004)].
23. J. L. Martins and A. Zunger, Phys. Rev. B **30**, 6217 (1984).
24. G. P. Srivastava, J. L. Martins, and A. Zunger, Phys. Rev. B **31**, 2561 (1985).
25. A. Garcia and M. L. Cohen, Phys. Rev. B **47**, 4215 (1993).
26. K. Karch, J.-M. Wagner, and F. Bechstedt, Phys. Rev. B **57**, 7043 (1998).

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 629–635. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 660–666.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Determination of the Concentration of Deep Levels in Semi-insulating CdS Single Crystals by Photoinduced-Current Transient Spectroscopy

A. P. Odrinskiĭ

Institute of Technical Acoustics, National Academy of Sciences of Belarus, pr. Lyudnikova 13, Vitebsk, 210717 Belarus
e-mail: odra@mail333.com

Submitted July 13, 2004; accepted for publication September 9, 2004

Abstract—Deep levels in semi-insulating CdS single crystals grown with a variation in the stoichiometric composition are investigated by photoinduced-current transient spectroscopy (PICTS). A series of deep levels with a thermal activation energy ranging from 0.066–0.54 eV is revealed. It is found that the signal ratio in the set of spectra disagrees with the basic PICTS model. A procedure for evaluation of the concentration of deep levels is developed for these conditions. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Currently, in order to study deep levels in high-resistivity semiconductors, researchers make efficient use of photoinduced-current deep-level transient spectroscopy (PICDLTS [3] or PICTS), which was first suggested in [1, 2]. This method is based on an analysis of the temperature dependence of the photocurrent relaxation kinetics. It makes it possible to obtain the same information as with the method of thermally stimulated conductivity (TSC) (the energy positions of deep levels E_t and the capture cross section of nonequilibrium carriers S_{nt}) from a single temperature scan, but it has a higher sensitivity [4] and accuracy when determining the parameters of deep levels [5].

However, opinions differ as to whether this method can be used to estimate the concentration of deep levels, which casts doubt on its role as a valuable method for studying this phenomenon. For example, in [6–9], the PICTS method was used to estimate the concentration of deep levels, and estimates [10] made using this method and the method ofoptical absorption have been found to coincide with an accuracy of 25%. At the same time, the concentration of deep levels is often determined using less effective procedures. For example, in [11], the PICTS method was used to estimate the relative concentration of deep levels, and the estimate was calibrated by measuring the space-charge limited current. It was noted in [11–15] that it is difficult to estimate many parameters of the base model in practice and that the method cannot be used to estimate the concentration of deep levels. This opinion is often related to the disagreement between the parameters of deep-level signals (peak heights) in a set of PICTS spectra[1] and the phenomenological model [5]. In this study, we demonstrate the possibility of making such an estimate using the example of a series of semi-insulating CdS single crystals.

2. EXPERIMENTAL RESULTS

Using the PICTS method, we studied several semi-insulating CdS crystals (with a resistivity of 10^8–10^{12} Ω cm) grown using the same method as described in [16] and possessing a varying stoichiometric composition. Except for the ratio of the partial vapor pressures of Cd and S p_S/p_{Cd} at the solidification front, the crystals were grown with constant growth parameters [17]. In [18], results from studying the effects of excitation using light with a photon energy $h\nu$ = 2.55 eV were reported. Based on the requirement for a more uniform excitation of the sample bulk and a decrease the effect of the surface, the PICTS measurements were carried out using excitation by quasi-monochromatic light with $h\nu \le E_g$ ($h\nu$ = 2.3–2.5 eV), where E_g is the band gap. The radiation of a DKSSh500 xenon lamp, which was passed through an MSD1 monochromator, was used. The width of the entrance and exit slits was 5 mm, which corresponds to a spectral range of ~300 Å. A partial set of results for excitation with $h\nu \le E_g$ was presented in [4]. It was concluded that the mechanism of trap filling due to the thermal capture of nonequilibrium carriers from the band (thermal mechanism), considered in the PICTS model, is insufficient to explain the detection of a dominant deep level in the spectra, and an optical mechanism of filling due to direct optical transitions between the level and the bands was proposed.

Under excitation at $h\nu$ = 2.35 eV, several deep levels were found (see the spectra in Fig. 1) that were not detected under excitation with $h\nu$ = 2.55 eV, and some

[1] The derivation of the peak-height ratio for the set of spectra is given in the Appendix.

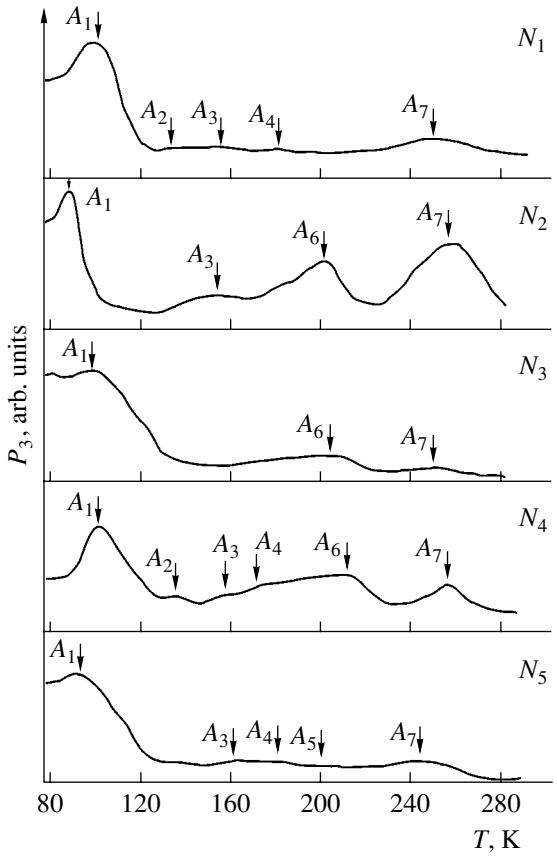

Fig. 1. PICTS spectra of the crystals when excited at $h\nu$ = 2.35 eV, which corresponds to a thermal emission of ~230 s^{-1}. The curves are normalized to the height of the dominant peak. The crystals were grown with the following ratios of partial Cd and S vapor pressures p_S/p_{Cd}: (N_1) 0.8, (N_2) 1.7, (N_3) 4.2, (N_4) 6.3, and (N_5) 7.6.

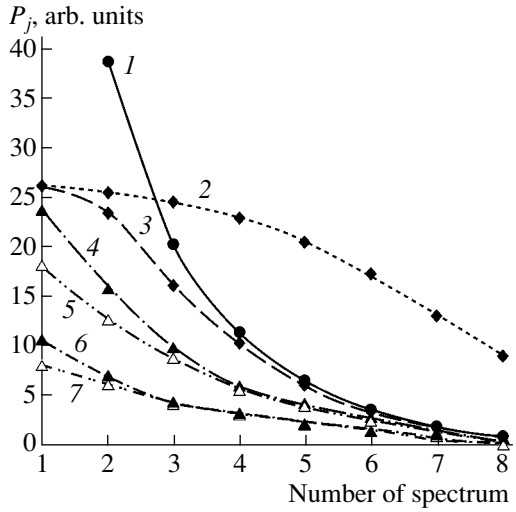

Fig. 2. Height of dominant peak A_1 in the set of spectra recorded for sample N_4. Curves *1* and *3* were obtained for the planar arrangement of contacts in the structure and excitation at $h\nu$ = 2.35 and 2.55 eV, respectively. Pairs of curves *4*, *5* and *6*, *7* were obtained for the structure with sandwich contacts (for various voltage polarities on the illuminated electrode) and excited by light at $h\nu$ = 2.35 and 2.55 eV, respectively. Curve *2* shows the peak heights calculated from Eq. (1) taking into account peak height P_1 in curve *3*.

special features were revealed in the change in the peak height in the PICTS spectra.

(i) As the photon energy of the excitation radiation decreases, peak A_1 (see Fig. 1) remains dominant in the spectra of all the samples in the range 80–110 K. Then, the spectral structure becomes complicated: new peaks (denoted as A_2–A_7) are observed. These peaks, although somewhat lower, are of the same order of magnitude as the height of peak A_1. However, when excitation is produced by radiation at $h\nu$ = 2.55 eV, peak A_1 is at least by two orders of magnitude higher than the other peaks. Measurements with sandwich contacts showed that peaks A_1–A_7 correspond to deep donor levels.

(ii) Figure 2 shows a comparison of the height of peak A_1 in the set of spectra recorded for sample N_4 using different excitation energies and locations of contacts. Sandwich (a semitransparent electrode on the illuminated face and the second electrode on the rear face) and planar (strip electrodes on the illuminated face) contacts were used. It can be seen that the signal excited by radiation at $h\nu$ = 2.35 eV is higher. Such an increase in the height of peak A_1 was observed for all the samples.

(iii) It follows from the phenomenological PICTS model (see Appendix) that the peak height ratio in the *i*th and *j*th spectra is

$$(P_i)_{\max}/(P_j)_{\max} = W_i/W_j, \quad (1)$$

where W is the characteristic numerical coefficient of the spectrum. In Fig. 2, curve *2* is calculated from the condition for the validity of Eq. (1) for curve *3*. It can be seen that the change in the height of peak A_1 in the set of spectra does not follow Eq. (1). Figure 3 shows changes in the peak heights in the set of spectra for other deep levels, which were detected for sample N_4 under excitation at $h\nu$ = 2.35 eV. It can be seen that, for almost all the deep levels, the following inequality is valid:

$$(P_i)_{\max}/(P_j)_{\max} < W_i/W_j. \quad (2)$$

Here, the *i*th spectrum corresponds to a lower rate of thermal emission.

3. ESTIMATION OF THE PHOTOGENERATION RATE OF NONEQUILIBRIUM CARRIERS

For a correct comparison of the data obtained under different excitation conditions, we selected the photogeneration rate of nonequilibrium carriers g_l as a criterion for characterizing the degree of sample excitation. The applicability of g_l as compared to the luminous flux Φ is due to the following reasons: (i) The variation in

the light penetration depth at constant Φ (measured in photon/(cm^2 s)) changes the concentration of nonequilibrium free carriers, which characterizes the degree of sample excitation. (ii) The photogeneration rate g_l enters the expression for the detected deep-level signal (see Eq. A5) and serves as a calibration parameter in the estimation of the concentration of deep levels.

In [11, 19], it was assumed that the light entering a sample is completely absorbed. Hence, the photogeneration rate of nonequilibrium carriers was determined via the incident luminous flux:

$$g_l = Q\Phi\alpha(1 - R). \quad (3)$$

Here, Q is the quantum yield of photoconductivity, α is the optical absorption coefficient, and R is the reflectance. To carry out such an estimation, it is necessary to measure the incident and reflected luminous fluxes. Samples are often small and, therefore, this estimation is a complex problem. An additional error results from the existence of surface recombination and the absence of data on Q.

However, the photogeneration rate can be estimated from measurements of the steady-state photocurrent. In this case, the surface recombination, the effects related to the reflection, and the quantum yield of photoconductivity are automatically taken into account. For steady-state optical excitation,

$$g_l = n_0/\tau_0. \quad (4)$$

Here, n_0 is the concentration of nonequilibrium charge carriers under steady-state optical excitation and τ_0 is their lifetime. The quantity g_l is temperature-independent, while n_0 and τ_0 depend strongly on temperature. Then, by measuring the lifetime of nonequilibrium carriers from the frequency dependence of the photocurrent and knowing the steady-state photocurrent at the same temperature, we can estimate the photogeneration rate. Such a procedure provides a more accurate estimate than optical measurements.

4. ANALYSIS OF THE CHANGE IN PICTS SIGNALS FROM DEEP LEVELS

As the energy of the excitation photons decreases, the light penetration depth in the samples increases and an increase in the height of the deep-level peak indicates that the concentration of the corresponding traps increases. The photogeneration rate of nonequilibrium carriers for sample N_4 was estimated to be $\sim 2 \times 10^{16}$ and 2×10^{14} cm^{-3} s^{-1} for excitation at $h\nu$ = 2.55 and 2.35 eV, respectively. In the latter case, the photogeneration rate is two orders of magnitude lower. Therefore, it can be suggested that the concentration changes by more than two orders of magnitude for the dominant deep level as the light penetration depth increases; moreover, the concentration is at least three orders of magnitude higher for deep levels, which are detected only under excitation at $h\nu$ = 2.35 eV. However, after their preparation (mechanical treatment and etching in a polishing etchant), the samples were not subjected to any thermal treatment that could introduce a gradient of deep-level concentration near the surface. Such inhomogeneities appeared during the growth of only one of the crystals. However, an increase in the peak height is characteristic of the entire series of samples. Therefore, an explanation based on the formation of concentration inhomogeneities during the crystal growth is not convincing.

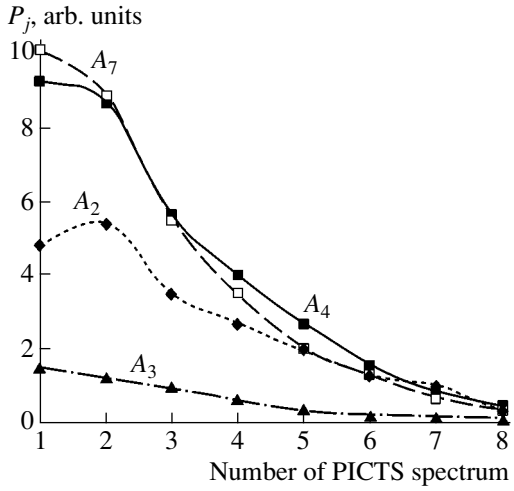

Fig. 3. Heights of peaks in the set of PICTS spectra of sample N_4 when excited at $h\nu$ = 2.35 eV.

In the phenomenological model [5], the normalization of the photocurrent relaxation signal makes it unnecessary, when estimating the energy positions and capture cross sections of deep levels, to consider the size of the region of the crystal in which photocurrent excitation and relaxation occurs. However, the geometrical factor can play a significant role in determining the deep-level concentration. When processing the data, the steady-state photocurrent is compared to the relaxation photocurrent at a certain instant after switching off the illumination. The latter quantity is proportional to the concentration of deep levels filled under nonequilibrium conditions, while the steady-state photocurrent is determined by the concentration of free nonequilibrium carriers. Consideration of the concentrations is justified if the regions of a sample in which the processes under study occur coincide. Depending on the experimental conditions, i.e., on arrangement of the contacts and the excitation radiation used (either in the intrinsic or impurity absorption range), three regions, which may not coincide with each other, should be distinguished in a sample. These are (i) the region of photogeneration of nonequilibrium carriers, (ii) the region in which excess charge carriers are generated, and (iii) the region where the processes involved in the nonequilibrium occupation of deep levels occur. The photogeneration region is determined by the light penetration depth in the crystal, which is equal to $1/\alpha$. During optical excitation, the generated nonequilibrium carri-

ers diffuse to a certain depth and can drift in the field. In this case, the region with excess carriers will differ from the region of carrier photogeneration.

If the deep levels are occupied due to the capture of nonequilibrium carriers from the band, the region with excess carrier concentration geometrically coincides with the region in which the nonequilibrium occupation of deep levels occurs. If, according to [4], the deep levels are occupied in accordance with the optical mechanism, the photogeneration region coincides with the region in which the deep levels are occupied under optical excitation. In this case, when estimating the concentration of deep levels, it is necessary to introduce a coefficient taking into account the ratio between the sizes of the region with excess carriers and the region of deep-level occupation.

When sandwich contacts are used the analysis is more complicated [20]. Therefore, below, we consider the planar arrangement of contacts. When free carriers of both signs are generated, ambipolar diffusion is controlled by the diffusion length of the minority carriers:

$$L_p = (D_p\tau_p)^{1/2}. \quad (5)$$

Here, τ_p is the lifetime and D_p is the diffusivity of the minority carriers. It follows from the Einstein relation that

$$D_p = \mu_p kT/q, \quad (6)$$

where μ_p is the mobility, k is the Boltzmann constant, and T is temperature. According to [21], $D_p = 0.25$ cm^2 s^{-1} at $\mu_p = 29$ cm^2 V^{-1} s^{-1}. According to [22], for typical values of τ_p in the range 10^{-6}–10^{-10} s, the diffusion length of the minority carriers is 5×10^{-4}–5×10^{-6} cm at T = 100 K, which is comparable with the size of the photogeneration region $1/\alpha = 10^{-5}$ cm under excitation at $h\nu = 2.55$ eV. In this case, the change in the signal that occurs with a change in excitation (Fig. 2, curves *1, 3*) can be explained by the diffusion effect. This explanation can be used, with some reservations, to interpret the change in the height of peak A_1, but the diffusion effect cannot account for the absence of signals from other deep levels under excitation with $h\nu = 2.55$ eV.

However, another explanation can be offered. If the optical mechanism, which is independent of the concentration of free nonequilibrium carriers, makes a significant contribution to the occupation of the deep levels, a larger number of traps are occupied during excitation at $h\nu = 2.35$ eV, as the penetration depth increases due to direct optical transitions. This circumstance should lead to an increase in deep-level signals. Let us assume that the light penetration depth increases by a factor of K at a constant luminous flux, which is absorbed into the samples. In this case, the photogeneration rate decreases by a factor of K. The concentration of deep levels occupied under nonequilibrium conditions and steady-state optical excitation is

$$n_{t0} = N_t/(1 + \beta), \quad (7)$$

where N_t is the concentration of deep levels and β is the ratio of the rates of level depletion (occupation):

$$\beta = (e_{nt}^{th} + e_{nt}^{o})/(g_l^*\tau_n^* C_{nt} + e_{pt}^{o}). \quad (8)$$

Here, e_{nt}^{th} is the rate of thermal emission from the deep levels, e_{nt}^{o} is the rate of optical depletion of the deep levels (transition of electrons from the deep levels to the conduction band with accompanying photon absorption), e_{pt}^{o} is the rate at which the optical occupation of the deep levels occurs (transition of electrons from the valence band to the deep levels with accompanying photon absorption), and C_{nt} is the coefficient of electron capture from the conduction band. If only the thermal mechanism operates, according to Eqs. (7) and (A5), for partially occupied deep levels[2] ($\beta \gg 1$),

$$(P_j)'_{\max} = N_t W_j K/[g_l(1 + \beta K)] \approx (P_j)_{\max}. \quad (9)$$

Thus, the signal does not change. For the optical mechanism of occupation, since the rates of optical occupation (depletion) remain constant as the light penetration depth increases, the peak height increases by a factor of K. Thus, with two competitive mechanisms for the occupation of the deep levels, the increase in the light penetration depth by a factor of K should lead to an increase in the signal by a factor of $1 - K$. Then, the increase in the signal that occurs with an increase in the penetration depth of the excitation radiation can be explained by the contribution of the optical mechanism to the occupation of the deep levels by nonequilibrium carriers, rather than by the change in the concentration of the corresponding deep level. Another consequence of this consideration is that the most reliable estimate of the concentration of deep levels can be obtained from the PICTS data using excitation by light when the penetration depth is comparable to the sample thickness. The fact that "many deep levels can be more easily detected" under excitation in the impurity absorption region was noted in [23]. The use of light-induced excitation for which the penetration depth is comparable to the sample thickness was also recommended in [19], although without any reliable justification.

Let us analyze the inconsistency between the peak height ratio in the set of spectra and the phenomenological model. It was found from the dependence of the photocurrent on the intensity of excitation light that linear recombination occurs. Traps for majority carriers were also detected. In this case, to explain the transformation of Eq. (1) into an inequality, we should assume that the steady-state optical-induced occupancy of the deep levels n_{t0} depends on temperature. This dependence distorts the peak height ratio and is associated with the partial occupation of a level under optical excitation. In [9, 11, 24], the estimation of the concentration of deep levels from PICTS data was based on the suggestion that their occupancy attains a constant value.

[2] Occupancy of deep level A_1 was estimated in [4].

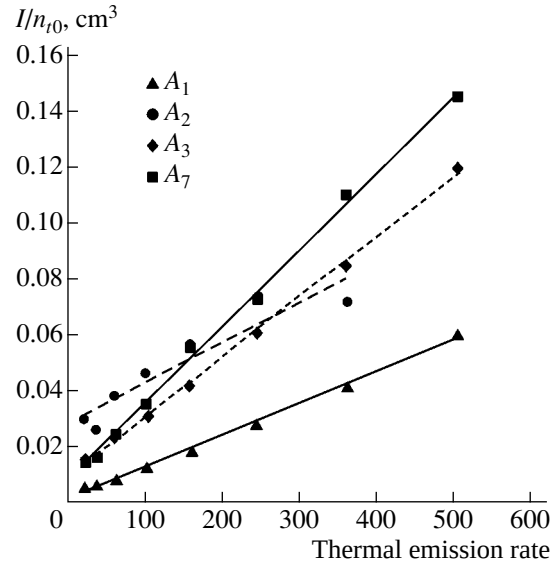

Fig. 4. Determination of the deep-level concentration by changing the optical occupation of the level. The intersection of the straight line $1/n_{t0} = f(e_{nt}^{th})$ with the ordinate yields the inverse concentration.

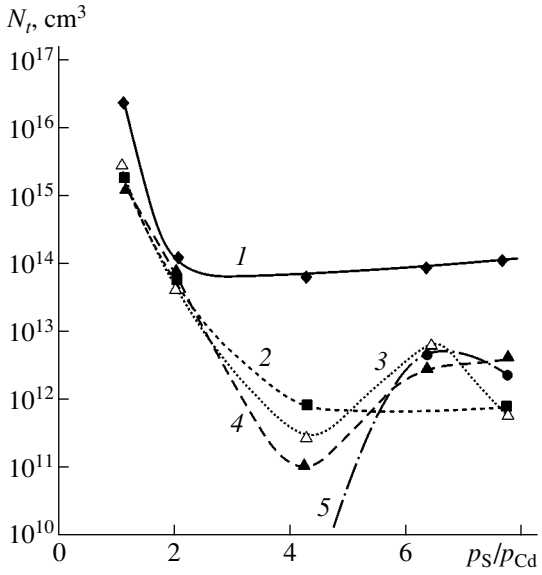

Fig. 5. Concentrations of detected deep levels versus the ratio of the partial pressures of the intrinsic components Cd and S during the crystal growth: (*1*) A_1, (*2*) A_6, (*3*) A_7, (*4*) A_2, and (*5*) A_5.

However, the occupation of traps depends on their parameters, and we can always find a situation in which it is, in principal, impossible to attain a sufficient level of optical excitation.

We can also estimate the concentration in the case of the partial occupancy of the deep levels. Based on Eqs. (7) and (8), and taking into account that the rate of thermal emission (depleting traps) varied by a factor of ~50 in the case under consideration in the temperature range T_{PICTS} in which deep level signals are detected, it is logical to suggest that the change in n_{t0} is associated specifically with the temperature dependence of the thermal emission rate. In this case, using Eq. (A5), we can estimate the steady-state optical-induced occupancy of a level for the corresponding rate of thermal emission for the *i*th spectrum from the height of the peak $(P_i)_{\max}$. As a result, we obtain a set of the pairs of values e_{nt}^{th} and n_{t0} for the temperature points of the peak position in the set of spectra. Making the classic assumption[3] that the rate at which the optical depletion of the deep levels occurs is negligible,

$$e_{nt}^{o}/(g_l^*\tau_n^* C_{nt} + e_{pt}^{o}) \ll 1, \qquad (10)$$

and linearizing formula (6), we can estimate the concentration of deep levels by the least-squares method using the set of values e_{nt}^{th} and n_{t0}. Figure 4 shows the plots used to determine the concentration of deep levels in sample N_4. For most of the deep levels, the experimental points are fitted by straight lines in the coordinates e_{nt}^{th} and $1/n_{t0}$. The significant spread of points for deep level A_2 is apparently due to the fact that peak A_2 is observed on the high-temperature shoulder of more intense peak A_1.

5. DISCUSSION

The table contains the parameters of the deep levels detected under excitation at $h\nu = 2.35$ eV and the estimates of their concentration for the crystals grown with a varying stoichiometric composition. Figure 5 shows the dependences of the concentrations of several detected deep levels on the ratio of S and Cd vapor pressures during the crystal growth. For the deep level responsible for dominant peak A_1, the data obtained agree with the equilibrium concentrations of intrinsic defects calculated in [25] as functions of the ratio of vapor pressures of the initial components. We can isolate a portion in the vicinity of the stoichiometric composition $p_S/p_{Cd} = 4$–5, where the lowest concentration of this deep level is observed. The increase in the concentration with a decrease in the S vapor pressure agrees with the donor nature of deep levels. This particular level has been attributed to interstitial cadmium Cd_i [18], and it was also observed using the PICTS method [5, 26].

[3] The rate at which the traps are thermally depleted is disregarded in an explicit or implicit form in all the methods using the optical excitation of a sample to estimate the deep-level concentration (TSC, PICTS, etc.). We can justify the assumption in the following way. It was noted in [19] that the Frank–Condon shift can only be moderate, and when photoexcitation with $h\nu \approx E_g$ is used for a level located in the upper part of the band gap, it is reasonable to assume that $e_{nt}^{o} < e_{pt}^{o}$. The thermal mechanism can also make a significant contribution to the occupation of the deep levels.

Table

Peak	T_{PICTS}, K	Type of trapped carriers	E_t, eV	S_{nt}, cm^2	Concentration of deep levels N_t, cm^{-3}				
					N_1	N_2	N_3	N_4	N_5
A_1	80–110	Majority carrie	0.066	10^{-19}	$2.2\times10^{+16}$	$1.2\times10^{+14}$	$6.4\times10^{+13}$	$8.5\times10^{+13}$	$1.1\times10^{+14}$
A_2	120–130	Majority carrie	0.27	10^{-12}	$1.3\times10^{+15}$	$8.0\times10^{+13}$	–	$2.6\times10^{+12}$	$3.7\times10^{+12}$
A_3	130–150	Majority carrie	0.28	10^{-13}	$1.4\times10^{+15}$	–	–	$1.4\times10^{+11}$	$2.9\times10^{+12}$
A_4	150–170	Majority carrie	0.39	1.1×10^{-12}	$1.9\times10^{+15}$	$6.4\times10^{+13}$	–	–	$2.1\times10^{+12}$
A_5	170–185	Majority carrie	0.32	10^{-13}	–	–	–	$4.6\times10^{+12}$	$2.1\times10^{+12}$
A_6	190–200	Majority carrie	0.27	10^{-12}	$1.9\times10^{+15}$	$5.6\times10^{+13}$	$7.8\times10^{+11}$	$6.0\times10^{+11}$	$6.8\times10^{+11}$
A_7	220–260	Majority carrie	0.54	10^{-13}	$2.9\times10^{+15}$	$4.8\times10^{+13}$	$2.6\times10^{+11}$	$5.5\times10^{+12}$	$5.3\times10^{+11}$

In [27], the donor level responsible for peak A_2 was attributed to either an interstitial lithium impurity or interstitial copper [28]. Weak peak A_3, also observed by the PICTS method in [26], can be interpreted in a similar manner.

The donor deep level responsible for peak A_5 was observed only for crystals N_4 and N_5, which were grown under higher S vapor pressure with respect to the stoichiometric composition. According to an equilibrium concentration diagram for intrinsic defects [25], this level may be related to the doubly ionized S vacancies.

6. CONCLUSIONS

Thus, the peak height ratio in the set of spectra indicates that the traps are only partially filled when a deep-level signal is detected. A procedure is proposed to estimate the concentration of traps for majority carriers in the case of partial trap filling. The increase observed in the deep-level signal excitation when shifting from excitation at $h\nu \geq E_g$ to excitation at $h\nu < E_g$ is due to the contribution of the optical mechanism (i.e., level-to-band direct optical transitions) to the occupation of the deep levels.

APPENDIX

Numerous types of signal processing [29], developed within the DLTS method [30], are used to analyze the temperature dependence of relaxation kinetics. The essence of this analysis is in the calculation of a set of data describing the temperature dependence of the thermal emission rate and allowing the calculation of the parameters of deep levels. The phenomenological model of the PICTS method [5, 19] was described using double-gate processing. The experimental data under consideration were obtained using another type of processing,[4] which is similar to double-gate processing but makes it possible to improve the signal-to-noise ratio. When recombination is linear and recapture is disregarded, the nonequilibrium filling of traps for majority carriers, according to the model [5], gives a component in the photocurrent relaxation kinetics, which can be written as

$$i_n'(t) = (n_{t0}/g_l)e_{nt}^{th}\exp(-e_{nt}^{th}t). \quad \text{(A1)}$$

Here, the signal is normalized to the steady-state photocurrent.

In the PICTS experiment, during the slow heating of the samples under periodic optical excitation, a set of spectra was constructed by processing the photocurrent relaxation kinetics:

$$P_j(T) = \int_0^\infty i_n(t)F_j(t)dt. \quad \text{(A2)}$$

Here, j = 1–8 is the spectrum number, $i_n(t)$ is the normalized photocurrent relaxation signal, and $F_j(t)$ is the weight function. The following values of the weight function were used:

$$F_j(t) = \begin{cases} 0 & \text{at } \Delta t + \tau_j \geq t > \Delta t + 3\tau_j, \\ +1 & \text{at } \Delta t + \tau_j < t \leq \Delta t + 2\tau_j, \\ -1 & \text{at } \Delta t + 2\tau_j < t \leq \Delta t + 3\tau_j. \end{cases} \quad \text{(A3)}$$

Here, τ_j is the characteristic parameter of the jth spectrum and Δt is the time delay. It can be shown that, for such processing, with the presence of the component described by Eq. (A1) in the relaxation kinetics, the jth spectrum will contain a peak at the temperature at which the rate of thermal emission from a deep level attains the value

$$r_{nt}^{th} = 1/\tau_j \ln[1 + 2\tau_j/(\tau_j + \Delta t)]. \quad \text{(A4)}$$

We find that the peak height in the jth spectrum is

$$(P_j)_{\max} = n_{t0}W_j/g_l, \quad \text{(A5)}$$

[4] This is associated with the development of computer engineering, which made it possible to change the algorithm for processing large data sets. Such sets are characteristic of experiments using methods involving kinetic spectroscopy. Previously, double-gate processing of relaxation kinetics was implemented in an analog form based on specialized experimental equipment (a double-box-car integrator). Now, relaxation kinetics is processed in a digital form based on a personal computer.

where W_j is the characteristic numerical coefficient of the spectrum. This coefficient is determined by the processing parameters:

$$W_j = 4\alpha^{\alpha}/(\alpha + 2)^{(\alpha + 2)}, \quad \text{(A6)}$$

where $\alpha \equiv (\Delta t/\tau_j + 1)$.

Assuming, as in [11, 19], that constant occupancy of the deep levels ($n_{t0} = N_t$) can be attained provided that the level of optical excitation is sufficient, we obtain Eq. (1) from Eq. (A5). Equation (1) describes the peak height ratio in the set of spectra. Thus, the criterion of reliability of the concentration estimation using a procedure similar to that carried out in [9, 11, 24] using Eq. (A5) shows the validity of Eq. (1).

REFERENCES

1. Ch. Hurter, M. Boilou, A. Mitonneau, and D. Bois, Appl. Phys. Lett. **32**, 821 (1978).
2. G. M. Martin and D. Bois, in *Semiconductor Characterization Techniques*, Ed. by P. A. Barnes and G. A. Rozgonyi (The Electrochemical Society, Princeton, N.J., 1978), PV 78-3, p. 32.
3. V. P. Kuznetsov, É. M. Omel'yanovskiĭ, A. Ya. Polyakov, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **19**, 735 (1985) [Sov. Phys. Semicond. **19**, 452 (1985)].
4. A. P. Odrinskiĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **38**, 310 (2004) [Semiconductors **38**, 298 (2004)].
5. O. Yoshie and M. Kamihara, Jpn. J. Appl. Phys. **22**, 621 (1983).
6. N. Benjelloun, M. Tapiero, J. P. Zielinger, *et al.*, J. Appl. Phys. **64**, 4013 (1988).
7. A. Blondeel and P. Clauws, J. Appl. Phys. **86**, 940 (1999).
8. S. Gariazzo and A. Serpi, Phys. Rev. B **41**, 7718 (1990).
9. R. Kozłowski, P. Kamiński, and E. Nossarzewska-Orłowska, Nucl. Instrum. Methods Phys. Res. A **476**, 639 (2002).
10. É. M. Omel'yanovskiĭ, A. Ya. Polyakov, N. S. Rytov, and V. I. Raĭkhshteĭn, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 1428 (1986) [Sov. Phys. Semicond. **20**, 897 (1986)].
11. M. Ayoub, M. Hage-Ali, J. M. Koebel, *et al.*, Mater. Sci. Eng. B **83**, 173 (2001).
12. S. R. Blight and H. Thomas, J. Appl. Phys. **65**, 215 (1989).
13. A. Zerrai, G. Marrakohi, and G. Bremond, J. Appl. Phys. **87**, 4294 (2000).
14. A. Castaldini, A. Cavallini, B. Fraboni, *et al.*, Phys. Rev. B **56**, 14897 (1997).
15. D. Seghier, J. Phys. D: Appl. Phys. **29**, 1071 (1996).
16. E. V. Markov and A. A. Davydov, Izv. Akad. Nauk SSSR, Neorg. Mater. **11**, 1755 (1975).
17. O. V. Bogdankevich, N. N. Kostin, E. M. Krasavina, *et al.*, Izv. Akad. Nauk SSSR, Neorg. Mater. **23**, 1618 (1987).
18. O. F. Vyvenko, I. A. Davydov, A. P. Odrinskiĭ, and V. A. Teplitskiĭ, Fiz. Tekh. Poluprovodn. (Leningrad) **28**, 721 (1990) [Sov. Phys. Semicond. **28**, 425 (1990)].
19. J. P. Zielinger and M. Tapiero, J. Phys. III **3**, 1327 (1993).
20. É. M. Omel'yanovskiĭ, A. Ya. Polyakov, N. S. Rytov, and V. I. Raĭkhshteĭn, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 1428 (1986) [Sov. Phys. Semicond. **20**, 897 (1986)].
21. R. Baubinas, B. P. Kietris, R. Reksnys, and A. Sakalas, Phys. Status Solidi A **50**, K63 (1978).
22. V. E. Lashkarev, A. V. Lyubchenko, and M. K. Sheĭnkman, *Nonequilibrium Processes in Photoconductors* (Naukova Dumka, Kiev, 1981) [in Russian].
23. Kousuke Ikeda and Yoshikazu Ishii, Jpn. J. Appl. Phys. **26**, 377 (1987).
24. J. C. Balland, J. P. Zielinger, M. Tapiro, *et al.*, J. Phys. D: Appl. Phys. **19**, 71 (1986).
25. *Physics of II–VI Compounds*, Ed. by A. N. Georgobiani and M. K. Sheĭnkman (Nauka, Moscow, 1986) [in Russian].
26. H. Ashour and F. El Akkad, Phys. Status Solidi A **184**, 175 (2001).
27. M. K. Sheĭnkman, N. E. Korsunskaya, I. V. Markevich, and T. G. Torchinskaya, Fiz. Tekh. Poluprovodn. (Leningrad) **14**, 438 (1980) [Sov. Phys. Semicond. **14**, 259 (1980)].
28. Yu. I. Émirov, S. S. Ostapenko, M. A. Rizakhanov, and M. K. Sheĭnkman, Fiz. Tekh. Poluprovodn. (Leningrad) **16**, 1371 (1982) [Sov. Phys. Semicond. **16**, 879 (1982)].
29. A. A. Istratov, O. F. Vyvenko, H. Hleslmair, and E. R. Weber, Meas. Sci. Technol. **9**, 477 (1998).
30. D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).

Translated by N. Korovin

Semiconductors, Vol. 39, No. 6, 2005, pp. 636–637. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 667–668.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Determination of the Charge Carrier Concentration in Lead Selenide Polycrystalline Layers Using Reflectance Spectra

A. E. Gamarts^, Yu. M. Kanageeva, and V. A. Moshnikov

St. Petersburg State Electrotechnical University, St. Petersburg, 197376 Russia

^e-mail: gamarts@mail.ru

Submitted August 31, 2004; accepted for publication September 17, 2004

Abstract—The possibility of using reflectance spectra in a nondestructive and contactless measurement of charge-carrier concentrations in polycrystalline lead selenide layers is shown. Changes in the carrier concentrations of such samples as a result of short-term sensitizing heat treatments are studied. © *2005 Pleiades Publishing, Inc.*

Polycrystalline layers based on lead chalcogenides have been used for many years in the production of photodetecting and emitting devices with an operating range of 2–5 μm [1, 2]. Interest in these materials has recently been renewed due to the development of a new generation of small-sized infrared (IR) absorption gas analyzers [3–5].

The structure and composition of the polycrystalline lead selenide (PbSe) films used in these devices are subjected to special modification (sensitizing) treatments, in which oxygen diffusion and active oxidation take place [6]. Such structural modification results in changes at various property levels [7]. These processes significantly depend on the characteristics of the initial layers, in particular, on their charge-carrier concentration.

Room-temperature observation of the photosensitivity effect in a lead layer requires the existence of *n*-type grains that change their conductivity type and have a barrier at the interface [8–10]. A network-type structure (Fig. 1), in which a percolation cluster is formed, is the most viable for this purpose.

The existence of a barrier structure with a complex current flow complicates the determination of carrier concentrations using methods based on kinetic effects. The IR reflection method, due to the absence of current, makes it possible to refine the Hall measurement data and results of the quantitative thermal probe method [11]. Their combined analysis should enable the development of an adequate model of the processes occurring in this type of nanostructured heterophase system.

The objective of this paper was to study the potential of the IR reflection method for a contactless nondestructive estimation of the carrier concentration and its subsequent use in the study of the physicochemical processes occurring during the sensitization of polycrystalline PbSe layers.

Anomalous dispersion of reflectance R is observed in the IR spectral region. The reflectance R tends to unity as the incident radiation frequency approaches the plasma frequency, which (in solids) is given by [12]

$$\omega_p = \sqrt{\frac{Ne^2}{m^*\varepsilon_\infty\varepsilon_S}}, \quad (1)$$

where N is the carrier concentration, m^* is the effective carrier mass, e is the elementary charge, and ε_S and ε_∞ are the static and dynamic permittivities. The reflectance R reaches its lowest level at the frequency

$$\omega_{\min} \approx \omega_p\sqrt{\frac{\varepsilon_\infty}{\varepsilon_\infty - 1}}. \quad (2)$$

Thus, the plasma frequency can be determined and the carrier concentration can be calculated using the plasma minimum position in a sample. The structures under study were polycrystalline PbSe films grown by vacuum thermal evaporation onto substrates that were transparent in the IR range (Si, BaF_2, and glass). These

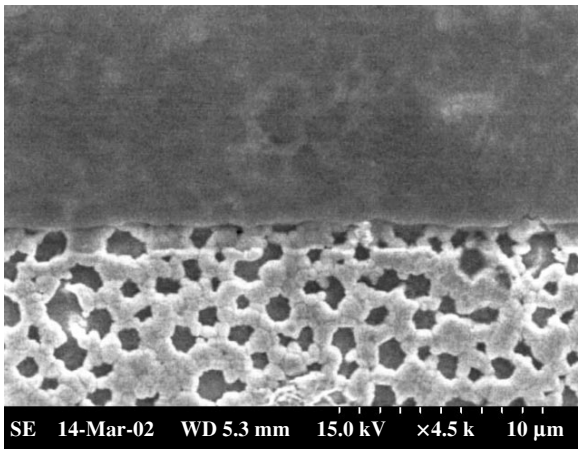

Fig. 1. A photosensitive lead selenide layer with a partially removed passivating coating (the image was obtained using a Hitachi S-3500N electron microscope).

1063-7826/05/3906-0636$26.00

Carrier concentration in the polycrystalline PbSe films, as determined from the reflectance spectra, and the calculation parameters

Nos.	Sample	Treatment	λ_{min}, μm	λ_p, μm	ε_∞	m^*	N, 10^{20} cm^{-3}
1	Si/PbSe	–	6.26	6.74	24	0.22 m_0	1.3
2	Si/PbSe	T = 693 K	18.74	19.12	24	0.11 m_0	0.081

films were then subjected to multistage sensitizing annealing. A more detailed description of the fabrication technology of the photosensitive elements can be found in [13].

In this study, we used an IKS-29 infrared spectrophotometer. The carrier concentration was calculated using a LabVIEW-6.0 code based on the Kukharsky and Subashiev method [14]. Its advantage is that not only the plasma minimum position but also the value of this minimum and (most importantly) its spectral shape in the region $\lambda > \lambda_{min}$ can be analyzed. This factor allows the determination of a larger number of the parameters that characterize the physical properties of crystals.

The results of the study are presented in Fig. 2 and the table. Similar experiments were performed with structures possessing grain sizes ranging from 0.1 to 0.3 μm. For practical purposes, short-term heat treatments were used, during which a plasma minimum shift was observed. Such a shift is an analytical signal for determining the diffusion parameters of a sensitizing impurity. The results confirm the previously obtained data on oxygen diffusion in polycrystalline PbSe layers [15].

Thus, the possibility of nondestructive contactless measurements of the carrier concentration in polycrystalline layers with submicrometer grain sizes and a network structure, used as photodetectors and emitters in the mid-IR range, is shown.

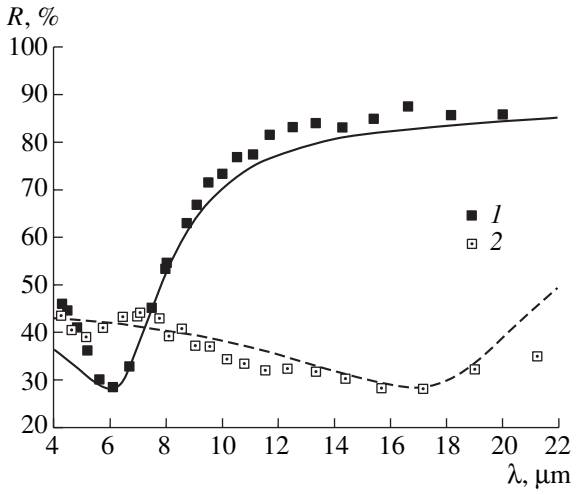

Fig. 2. Reflectance spectra of the (*1*) initial and (*2*) annealed polycrystalline lead selenide layers. The experimental data is indicated by dots, and the curves show the results calculated using the Kukharsky and Subashiev method.

ACKNOWLEDGMENTS

The authors are grateful to Prof. A.N. Veis for participating in a detailed discussion of the results and for his helpful suggestions.

This study was supported by a grant for the support of post-graduate students provided by the Ministry of Education of the Russian Federation, no. A04-3.15-410.

REFERENCES

1. A. F. Gibson, W. D. Lawson, and T. S. Moss, Proc. Phys. Soc. London, Sect. A **64**, 1054 (1957).
2. J. N. Hamphry and W. W. Scanlon, Phys. Rev. **5** (3), 256 (1957).
3. E. M. Gamarts and V. A. Krylov, Peterb. Zh. Élektron., No. 1, 54 (2004).
4. I. G. Lukitsa, E. M. Gamarts, V. A. Krylov, and S. I. Frantsuzova, Eur. Patent No. 98,403,035.3-2204 (16 February 1999).
5. T. Boudet, J. Fantini, E. Gamarts, and V. Krilov, Eur. Patent No. 01,401,408.8-2204 (17 January 2001).
6. S. M. Repinskiĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 1050 (2001) [Semiconductors **35**, 1006 (2001)].
7. A. I. Popov, V. A. Vorontsov, and I. A. Popov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 665 (2001) [Semiconductors **35**, 637 (2001)].
8. L. N. Neustroev and V. V. Osipov, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 59 (1986) [Sov. Phys. Semicond. **20**, 34 (1986)].
9. L. N. Neustroev and V. V. Osipov, Mikroélektronika **17**, 399 (1988).
10. V. T. Trofimov, Yu. G. Selivanov, and E. G. Chizhevskiĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 755 (1996) [Semiconductors **30**, 405 (1996)].
11. V. A. Moshnikov and S. L. Miloslavov, in *Production and Properties of II–VI and IV–VI Semiconductor Compounds and Solid Solutions Based on These Compounds* (Mosk. Inst. Stali Splavov, Moscow, 1977), p. 302 [in Russian].
12. Yu. A. Ukhanov, *Optical Properties of Semiconductors* (Nauka, Moscow, 1977) [in Russian].
13. E. M. Gamarts, N. V. Golubchenko, V. A. Moshnikov, and D. B. Chesnokova, Peterb. Zh. Élektron., No. 4, 11 (2003).
14. A. A. Kukharskiĭ and V. K. Subashiev, Fiz. Tverd. Tela (Leningrad) **8**, 753 (1966) [Sov. Phys. Solid State **8**, 603 (1966)].
15. A. E. Gamarts, V. M. Lebedev, V. A. Moshnikov, and D. B. Chesnokova, Fiz. Tekh. Poluprovodn. (St. Petersburg) **38**, 1195 (2004) [Semiconductors **38**, 1160 (2004)].

Translated by A. Kazantsev

Semiconductors, Vol. 39, No. 6, 2005, pp. 638–641. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 669–672.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Energy Parameters of Two-Electron Tin Centers in PbSe

S. A. Nemov*, F. S. Nasredinov*^, P. P. Seregin*, N. P. Seregin, and É. S. Khuzhakulov*****

**St. Petersburg State Polytechnical University, ul. Politekhnicheskaya 25, St. Petersburg, 195251 Russia*

^e-mail: nasredinov@tuexph.stu.neva.ru

***Institute of Analytical Instrumentation, Russian Academy of Sciences, St. Petersburg, 198103 Russia*

****Tashkent Oblast State Pedagogical University, Angren, 702500 Uzbekistan*

Submitted September 9, 2004; accepted for publication September 27, 2004

Abstract—Relations that make it possible to use an experimentally measured temperature dependence of carrier concentration to determine the Hubbard energy U and temperature dependence of the Fermi level F for two-electron tin centers in lead selenide are derived. A study of $Pb_{1-x-y}Sn_xNa_ySe$ solid solutions shows that their Fermi level in the temperature region 100–600 K lies below the valence band top E_V and that their $F(T)$ dependences are linear, with extrapolation to $T = 0$ yielding $E_V - F = 210 \pm 10$ meV. The Hubbard energy of the two-electron tin centers in PbSe is found to be $U = -80 \pm 20$ meV. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

According to transport studies [1] and ^{119}Sn Mössbauer spectroscopy data [2], the tin impurity in PbSe is a donor; indeed, Mössbauer spectra of degenerate *n*-type samples of $Pb_{1-x}Sn_xSe$ containing an excess of lead have a shape characteristic of divalent tin Sn^{2+}; the spectra of overcompensated degenerate *p*-type samples of $Pb_{1-x-y}Sn_xNa_ySe$ ($y \gg 2x$) are characteristic of tetravalent tin Sn^{4+}; and those of partially compensated degenerate *p*-type samples of $Pb_{1-x-y}Sn_xNa_ySe$ ($y \le 2x$), of a superposition of divalent and tetravalent tin. It is believed that tin atoms replace divalent lead in the PbSe cubic lattice and form donor states, so that the lines characteristic of Sn^{2+} and Sn^{4+} in the Mössbauer spectra correspond to the neutral ($[Sn]^0$) and doubly ionized states ($[Sn]^{2+}$) of the tin donor center, respectively. The fact that only the Sn^{4+} line (without the Sn^{2+} line) is observed in the Mössbauer spectra, and only in strongly overcompensated $Pb_{1-x-y}Sn_xNa_ySe$ samples, suggests that the tin levels are located against the background of the valence band. The absence of Sn^{3+} (the singly ionized tin donor center) in the Mössbauer spectra of partially compensated samples implies that, in PbSe, tin forms two-electron donor centers with negative Hubbard energy.

Thus, tin forms, against the valence band background of $Pb_{1-x}Sn_xSe$, two bands of localized tin states separated by the Hubbard energy

$$U = E_1 - E_2, \tag{1}$$

where E_1 is the energy of the level to which an electron is excited and transfers Sn^{3+} to the Sn^{2+} center, and E_2 is the energy of the level to which an electron is excited and transforms the Sn^{4+} center into the Sn^{3+} center. The density of states as a function of the energy in the $Pb_{1-x-y}Sn_xNa_ySe$ band gap was suggested by Nasredinov *et al.* [2]. A similar situation arises with thallium centers in lead chalcogenides [3]. The aim of this study was to determine the temperature dependences of the Fermi level and estimate the Hubbard energy for $Pb_{1-x-y}Sn_xNa_ySe$ solid solutions.

2. TEMPERATURE DEPENDENCES OF THE FERMI LEVEL AND CARRIER CONCENTRATION

The electroneutrality equation for $Pb_{1-x-y}Sn_xNa_ySe$ solid solutions can be written as

$$2N_{Sn^{4+}} + N_{Sn^{3+}} + p = N_{Na}, \tag{2}$$

where $N_{Sn^{3+}}$ and $N_{Sn^{4+}}$ are the concentrations of Sn^{3+} and Sn^{4+} centers, respectively; p is the hole concentration in the valence band; and N_{Na} is the concentration of ionized one-electron acceptors (sodium).

According to the Gibbs distribution, the concentration of impurity centers with different numbers of electrons is given by the relation

$$\frac{N_s}{N_{s-1}} = \frac{g_s}{g_{s-1}} \exp\left[\frac{F - E_s}{kT}\right], \tag{3}$$

where N_s and N_{s-1} are the concentrations of centers with s and $s-1$ electrons, g_s and g_{s-1} are the spin degeneracy factors for the corresponding centers, F is the Fermi level, E_s is the energy of the level to which the sth electron becomes attached, and k is the Boltzmann constant. Thus, we have

$$\frac{N_{Sn^{2+}}}{N_{Sn^{3+}}} = \frac{g_{Sn^{2+}}}{g_{Sn^{3+}}} \exp\left[\frac{F - E_1}{kT}\right] \tag{4}$$

and

$$\frac{N_{Sn^{3+}}}{N_{Sn^{4+}}} = \frac{g_{Sn^{3+}}}{g_{Sn^{4+}}}\exp\left[\frac{F-E_2}{kT}\right], \quad (5)$$

where $N_{Sn^{2+}}$ is the concentration of Sn^{2+} centers, and $g_{Sn^{2+}}$, $g_{Sn^{3+}}$, $g_{Sn^{4+}}$ are degeneracy factors for the Sn^{2+}, Sn^{3+}, and Sn^{4+} centers, respectively (if 5*s* electrons are responsible for the donor properties of tin, then $g_{Sn^{2+}} = 1$, $g_{Sn^{3+}} = 2$, and $g_{Sn^{4+}} = 1$).

Since

$$N_{Sn^{2+}} + N_{Sn^{3+}} + N_{Sn^{4+}} = N_{Sn}, \quad (6)$$

where N_{Sn} is the total tin concentration, we can introduce the following expressions for $N_{Sn^{3+}}$ and $N_{Sn^{4+}}$ into Eq. (2):

$$N_{Sn^{3+}} = N_{Sn}\left\{1 + \frac{g_{Sn^{2+}}}{g_{Sn^{3+}}}\exp\left[\frac{F-E_1}{kT}\right] + \frac{g_{Sn^{4+}}}{g_{Sn^{3+}}}\exp\left[\frac{E_2-F}{kT}\right]\right\}^{-1},$$
$$N_{Sn^{4+}} = N_{Sn}\left\{1 + \frac{g_{Sn^{3+}}}{g_{Sn^{4+}}}\exp\left[\frac{F-E_1}{kT}\right] + \frac{g_{Sn^{2+}}}{g_{Sn^{4+}}}\exp\left[\frac{2F-E_1-E_2}{kT}\right]\right\}^{-1}. \quad (7)$$

For the hole concentration in the valence band, we can write

$$p = \int_{-\infty}^{E_v} g(E)[1-f(E)]dE, \quad (8)$$

where $g(E)$ is the density of states in the valence band, $f(E) = \dfrac{1}{\exp\left[\dfrac{E-F}{kT}\right]+1}$ is the Fermi function, and E_v is the energy of the valence-band top.

If $E = E_v - kT\varepsilon$ and $F = E_v - kT\mu$, then $1 - f(E) = \dfrac{1}{\exp(\varepsilon-\mu)+1}$, which yields the following expression

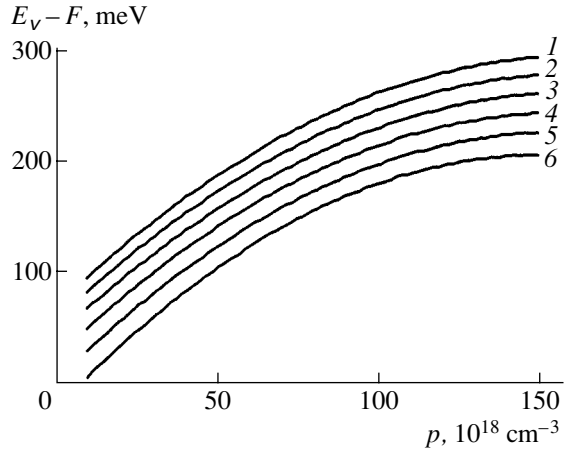

Fig. 1. Values of the integral $p = \frac{4\pi}{h^3}(2m_dkT)^{3/2}\int_0^\infty \varepsilon^{1/2}\left(1+2\varepsilon\frac{kT}{E_g}\right)\left(1+\varepsilon\frac{kT}{E_g}\right)^{1/2}\frac{1}{\exp(\varepsilon-\mu)+1}d\varepsilon$ tabulated for the temperatures 100, 200, 300, 400, 500, and 600 K (curves *1–6*, respectively).

for Cane's model that takes into account electron interaction in the valence and conduction bands:

$$g(E) = \frac{4\pi}{h^3}(2m_d)^{3/2}(E_v-E)^{1/2} \times\left(1+2\frac{E_v-E}{E_g}\right)\left(1+\frac{E_v-E}{E_g}\right)^{1/2} = \frac{4\pi}{h^3}(2m_d)^{3/2}\varepsilon^{1/2}\left(1+2\varepsilon\frac{kT}{E_g}\right)\left(1+\varepsilon\frac{kT}{E_g}\right)^{1/2}, \quad (9)$$

where E_g is the band gap, $m_d = (4)^{2/3}\sqrt[3]{m_\perp^2 m_\parallel}$ is the effective mass of the density of states near the valence-band top; and $m_\perp$ and $m_\parallel$ are the transverse and longitudinal effective masses, respectively.

The temperature dependence of the effective mass for PbSe is given by the relation [3]

$$m_d(T) = \frac{m_d(0)}{E_g(0)}E_g(T), \quad (10)$$

where $m_d = 0.11m_e$ and $E_g = 0.165$ eV are the effective mass and the band gap for $T = 0$, and the temperature dependence of E_g can be written as [3]

$$E_g(T) = E_g(0)(1 + 4\times10^{-4}T). \quad (11)$$

Thus, the hole concentration in the valence band is given by

$$p = \frac{4\pi}{h^3}(2m_dkT)^{3/2} \times\int_0^\infty \varepsilon^{1/2}\left(1+2\varepsilon\frac{kT}{E_g}\right)\left(1+\varepsilon\frac{kT}{E_g}\right)^{1/2}\frac{1}{\exp(\varepsilon-\mu)+1}d\varepsilon. \quad (12)$$

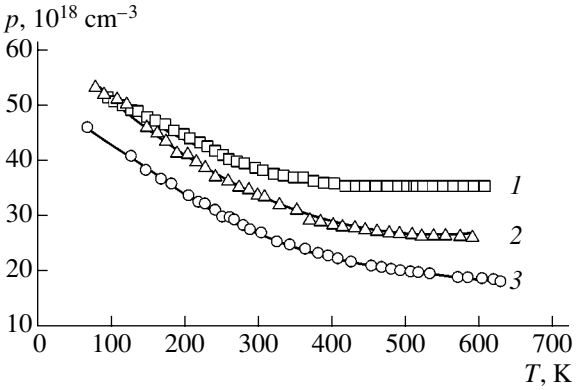

Fig. 2. Temperature dependences of the hole concentration for the $Pb_{1-x-y}Sn_xNa_ySe$ solid solutions: curve *1* corresponds to $N_{Sn} = 8.7 \times 10^{19}$ cm^{-3} and $N_{Na} = 5.22 \times 10^{19}$ cm^{-3}; curve *2*, $N_{Sn} = 1.74 \times 10^{20}$ cm^{-3} and $N_{Na} = 5.22 \times 10^{19}$ cm^{-3}; and curve *3*, $N_{Sn} = 3.48 \times 10^{20}$ cm^{-3} and $N_{Na} = 5.22 \times 10^{19}$ cm^{-3}.

Equation (12) permits us to determine the temperature dependence $F(T)$ by comparing the experimentally measured $p(T)$ relation with tabulated values of the integral on the right-hand side of this equation (see Fig. 1).

Electroneutrality relation (1) can be conveniently recast in the form

$$N_{Na} - \rho = p, \tag{13}$$

where ρ is the density of positive charge at the tin centers (in units of electron charge), or in the expanded form

$$\begin{aligned} N_{Na} - N_{Sn} &= \frac{2 + \dfrac{g_{Sn^{3+}}}{g_{Sn^{4+}}}\exp\left[\dfrac{F - E_2}{kT}\right]}{1 + \dfrac{g_{Sn^{3+}}}{g_{Sn^{4+}}}\exp\left[\dfrac{F - E_2}{kT}\right] + \dfrac{g_{Sn^{2+}}}{g_{Sn^{4+}}}\exp\left[\dfrac{2F - E_1 - E_2}{kT}\right]} \\ &= \frac{4\pi}{h^3}(2m_d kT)^{3/2}\int_0^\infty \varepsilon^{1/2}\left(1 + 2\varepsilon\frac{kT}{E_g}\right) \\ &\quad \times \left(1 + \varepsilon\frac{kT}{E_g}\right)\frac{1}{\exp(\varepsilon - \mu) + 1}d\varepsilon, \end{aligned} \tag{14}$$

with the charge at the localized centers represented on the left-hand side of the equation, and that of the free holes in the valence band, on the right-hand side.

The mean energy $E_0 = \dfrac{E_1 + E_2}{2}$ and the Hubbard energy U can be determined by calculating the occupancy of the tin impurity centers by electrons:

$$\xi = \frac{\rho}{eN_{Sn}} = \frac{N_{Na} - p}{N_{Sn}}. \tag{15}$$

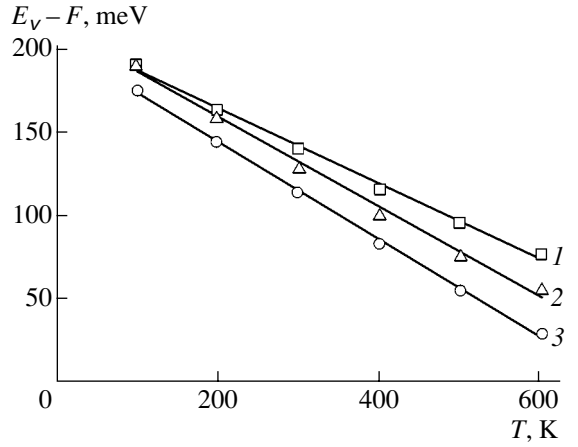

Fig. 3. Temperature dependences of the Fermi level for the $Pb_{1-x-y}Sn_xNa_ySe$ solid solutions: curve *1* corresponds to $N_{Sn} = 8.7 \times 10^{19}$ cm^{-3} and $N_{Na} = 5.22 \times 10^{19}$ cm^{-3}; curve *2*, $N_{Sn} = 1.74 \times 10^{20}$ cm^{-3} and $N_{Na} = 5.22 \times 10^{19}$ cm^{-3}; and curve *3*, $N_{Sn} = 3.48 \times 10^{20}$ cm^{-3} and $N_{Na} = 5.22 \times 10^{19}$ cm^{-3}.

This equation can be written as

$$\begin{aligned} \xi &= 2\frac{1 + \exp\left(\dfrac{F - E_2}{kT}\right)}{1 + 2\exp\left(\dfrac{F - E_2}{kT}\right) + \exp\left(2\dfrac{F - E_a}{kT}\right)} \\ &= 2\frac{1 + \exp\left(\dfrac{U}{2kT}\right)\exp\left(\dfrac{F - E_0}{kT}\right)}{1 + 2\exp\left(\dfrac{U}{2kT}\right)\exp\left(\dfrac{F - E_0}{kT}\right) + \exp\left(2\dfrac{F - E_0}{kT}\right)} \end{aligned} \tag{16}$$

or, introducing the notation $w = \exp\left(\dfrac{U}{2kT}\right)$ and $z = \exp\left(\dfrac{F - E_0}{kT}\right)$,

$$\xi = 2\frac{1 + wz}{1 + 2wz + z^2}, \tag{17}$$

with $w \ll 1$ for a degenerate semiconductor.

3. EXPERIMENTAL RESULTS

Figure 2 shows experimental temperature dependences of the hole concentration for $Pb_{1-x-y}Sn_xNa_ySe$ samples differing in relation to the degree of compensation of the two-electron tin donor impurity by the one-electron acceptor. The integral on the right-hand side of Eq. (12) was tabulated for the temperatures 100, 200, 300, 400, 500, and 600 K, and the resulting values

were used to determine the $F(T)$ dependences. Figure 3 shows the results of the calculation. It can be seen that, for all the samples studied, the Fermi level in the temperature interval 100–600 K lies in the valence band and that the $F(T)$ dependences are linear:

$$F = F_0 + \alpha kT. \quad (18)$$

Here, $\alpha = \frac{1}{k}\frac{dF}{dT}$. Extrapolation to $T = 0$ yields $E_v - F_0 = 210 \pm 10$ meV for all the samples.

To find the mean energy $E_0 = \frac{E_1 + E_2}{2}$, we should rewrite expression (16) in the form

$$\xi = 2\frac{1 + \exp\left(\frac{U}{2kT}\right)\exp\left(\frac{F_0 - E_0}{kT}\right)\exp(\alpha)}{1 + 2\exp\left(\frac{U}{2kT}\right)\exp\left(\frac{F_0 - E_0}{kT}\right)\exp(\alpha) + \exp\left(2\frac{F_0 - E_0}{kT}\right)\exp(2\alpha)}. \quad (19)$$

It follows that, for $T \longrightarrow 0$, we have $F_0 - E_0 > 0$ (the Fermi level is controlled by valence-band holes) if $p_0 = N_{Na}$, where p_0 is the hole concentration at $T = 0$ (for $E_V - F_0 < 210$ meV, this is possible for $N_{Na} < 3 \times 10^{19}$ cm^{-3}, i.e., for $x < 0.018$) and $F_0 - E_0 < 0$ (Fermi level is controlled by the tin level occupation) if $N_{Na} > 2N_{Sn} + 3 \times 10^{19}$ cm^{-3}. For intermediate sodium concentrations, $F_0 = E_0$.

The composition of the samples under study suggests that in all cases the relation $F_0 = E_0$ holds; in this situation,

$$\xi = 2\frac{1 + \exp\left(\frac{U}{2kT}\right)\exp(\alpha)}{1 + 2\exp\left(\frac{U}{2kT}\right)\exp(\alpha) + \exp(2\alpha)}. \quad (20)$$

The Hubbard energy U can be derived from the latter relation using the experimental values of α and ξ:

$$U = 2kT\ln\left[\frac{\exp(\alpha)}{2}\frac{\xi}{1-\xi} - \frac{\exp(\alpha)}{2}\left(2 + \frac{\xi}{1-\xi}\right)\right]. \quad (21)$$

The energy U was determined for a temperature of 100 K (since, at higher temperatures, we have to take into account the temperature dependence of E_0) and found to be $U = -70 \pm 10$ meV.

Note that a Mössbauer study of the electron exchange between neutral and ionized tin centers in PbSe in the temperature range 100–400 K yielded $U = 60 \pm 20$ meV [2], which is in good agreement with our value of U.

4. CONCLUSION

Relations permitting the use of an experimental temperature dependence of the carrier concentration in lead selenide with two-electron tin centers to determine the Hubbard energy and temperature dependence of the Fermi level were derived from the Gibbs distribution. For all the $Pb_{1-x-y}Sn_xNa_ySe$ solid solutions studied, the Fermi level in the 100–600-K temperature region was located in the valence band, the $F(T)$ dependences were linear, and their extrapolation to $T = 0$ yielded $E_v - F_0 = 210 \pm 10$ meV. The Hubbard energy for the tin centers was found to be $U = -70 \pm 20$ meV, which is in a good agreement with Mössbauer spectroscopy data.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, project no. 02-02-17306.

REFERENCES

1. G. T. Alekseeva, E. A. Gurieva, P. P. Konstantinov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1388 (1995) [Semiconductors **29**, 719 (1995)].
2. F. S. Nasredinov, S. A. Nemov, V. F. Masterov, and P. P. Seregin, Fiz. Tverd. Tela (St. Petersburg) **41**, 1897 (1999) [Phys. Solid State **41**, 1741 (1999)].
3. S. A. Nemov and Yu. I. Ravich, Usp. Fiz. Nauk **168**, 817 (1998) [Phys. Usp. **41**, 735 (1998)].

Translated by G. Skrebtsov

Semiconductors, Vol. 39, No. 6, 2005, pp. 642–649. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 673–680.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Dynamic Chaos in a Partially Illuminated Compensated Semiconductor Under the Conditions of Impurity-Related Breakdown

K. M. Jandieri^, Z. S. Kachlishvili, and A. B. Stroganov

Tbilisi State University, Tbilisi, 380028 Georgia

^e-mail: kjandieri@yahoo.com

Submitted June 21, 2004; accepted for publication September 30, 2004

Abstract—The nonlinear dynamics of a compensated semiconductor in the case of an impurity-related electrical breakdown in a classically strong magnetic field under the effect of shortened Hall contacts and resonance radiation, the frequency of which corresponds to the ionization energy of hydrogen-like donor impurities, is evaluated. As a result, both regular and chaotic self-sustained oscillations are obtained. It is found that all three scenarios for the origination of chaos are realized for corresponding values of the bifurcation parameters. The results obtained have the potential become the theoretical foundation for the operation of an easily controlled high-frequency oscillator whose modes of operation (the switched-on mode and the transition from the regular mode to a chaotic mode and vice versa) could be changed by varying the illumination intensity. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The impurity-related electrical breakdown of a compensated semiconductor is one of the mechanisms conducive to the origination of oscillatory instability. There have been many publications concerned with theoretical and experimental studies in this field (see, for example, [1, 2]). The conditions for the origination of nonlinear oscillations and the characteristics of these oscillations depend on a number of external parameters. Among these, a magnetic field is of major importance. In this context, it is reasonable to consider the following two cases separately: (i) the mode of a specified direction of current (the Hall contacts are open; i.e., we have the typical Hall mode) and (ii) the mode of a specified direction of field (the Hall contacts are shorted).

In case (i), it is possible to obtain both regular and chaotic self-sustained oscillations if not only the dielectric relaxation of an applied field but also the relaxation of the Hall electric field is taken into account. The chaos originates according to the Feigenbaum scenario. Such a case was earlier considered in [3, 4].

Nonlinear oscillations in mode (ii) were studied in [5, 6] using a mathematical model based on differential equations that described generation–recombination processes at a hydrogen-like impurity center, the dielectric relaxation of an applied electric field in a semiconductor, and a lag of the electron temperature with respect to variations in the electric field. As a result, the phase diagram produced included both elementary and binary limiting cycles. This circumstance serves as a theoretical basis for a high-frequency oscillator that can operate in two different amplitude modes under the same conditions; switching from one mode to the other can be accomplished easily using a small-amplitude external current pulse.

The aforementioned mathematical model provided no way of describing chaotic oscillations. Apparently, this shortcoming was caused by the fact that the equation describing a lag in the electron temperature was written for a much smaller time scale than could be applied to the other equations. Consequently, the equation for the lag could not make a sufficiently large contribution to the total dynamics of a semiconductor.

This paper is a continuation of publications [5, 6]; i.e., we study the nonlinear oscillatory dynamics of a compensated semiconductor in a situation where the Hall contacts are shorted out (a magnetic field is applied perpendicularly to an electric field) under the conditions of an impurity-related electrical breakdown. However, in contrast to [5, 6], we consider a semiconductor that is partially exposed to light: a part of the sample is subjected to resonance optical radiation while the other part is not illuminated (the energy of the photons in the incident radiation corresponds to the ionization energy of the hydrogen-like impurity). As a result, the kinetic processes and concentrations of free charge carriers in different parts of the semiconductor differ. The latter circumstance should give rise to a diffusion current in the vicinity of the interface between the illuminated and unilluminated parts of the sample. However, we consider a situation in which the diffusion current can be disregarded (the sample length in the direction of the electric field is much larger than both the diffusion and drift lengths). Under these conditions, a highly accurate representation of a semiconductor can be attained with two series-connected illuminated and unilluminated samples. As will become clear from the

1063-7826/05/3906-0642$26.00

reasoning and results below, this system, including its chaos aspect, is of a certain academic interest, since all three universal scenarios of transitions from regular to chaotic oscillations are realized in the system under consideration. A corresponding four-dimensional mathematical model can be derived using the equations for kinetic processes and the electric-field dielectric relaxation for each sample taken separately (as was mentioned above, the equations describing electron-temperature lag are for use on a much smaller time scale; thus, we can state with confidence that the electron temperatures instantaneously take their quasi-stationary values). It is worth noting that a similar problem has been considered by Schoell [7], who studied the dynamics of two series-connected semiconductors with differing parameters. However, first, a magnetic field was not applied to the sample, and, second, the mathematical model was based on an equation for energy relaxation rather than on an equation for electric-field relaxation. In addition, the difference between the semiconductors was described using their internal parameters. In the case under consideration in this study, we use an external parameter, i.e., illumination of the sample, which, in our opinion, is more convenient from the standpoint of the practical use of the corresponding theoretical results.

2. MATHEMATICAL MODEL

Thus, the mathematical model of the system under consideration is represented by the following set differential equations (the problem is considered using the electron-temperature method):

$$\frac{dn_1}{dt} = \gamma_{opt}[N_D - N_A - n_1] + A_I^{(1)}(Z_1)[N_D - N_A - n_1]n_1 - B_T^{(1)}(Z_1)[N_A + n_1]n_1, \quad (1^0)$$

$$\frac{dn_2}{dt} = \gamma_T[N_D - N_A - n_2] + A_I^{(2)}(Z_2)[N_D - N_A - n_2]n_2 - B_T^{(2)}(Z_2)[N_A + n_2]n_2, \quad (2^0)$$

$$\frac{dE_1}{dt} = \frac{4\pi}{\varepsilon SR}[\mathscr{E} - E_1L - E_2L - eSRn_1\mu_1(Z_1)E_1], \quad (3^0)$$

$$\frac{dE_2}{dt} = \frac{4\pi}{\varepsilon SR}[\mathscr{E} - E_2L - E_1L - eSRn_2\mu_2(Z_2)E_2]. \quad (4^0)$$

Here, n_1 and n_2 are the free-electron concentrations in the illuminated sample (sample 1) and the unilluminated sample (sample 2), respectively; E_1 and E_2 are the strengths of the corresponding electric fields; $Z_1 \equiv [T_{e1}/T]$ and $Z_2 \equiv [T_{e2}/T]$ are dimensionless electron temperatures; T is the lattice temperature; N_D is the concentration of the hydrogen-like donor impurity; N_A is the concentration of compensating acceptors; $A_I^{(1)}$, $A_I^{(2)}$, $B_T^{(1)}$, and $B_T^{(2)}$ are the coefficients of impact ionization and thermal recombination; γ_T and γ_{opt} are the rates of thermal and optical ionization (thermal ionization is negligible compared with optical ionization in the illuminated sample as a result of a low lattice temperature, $T = 4.2$ K); μ_1 and μ_2 are the free-electron mobilities; R is the load resistance connected in series with the samples; $\mathscr{E}$ is the emf of the dc power supply; and S is the cross section, L is the length, and ε is the permittivity of the samples.

We now rewrite system of Eqs. (1^0)–(4^0) in a more convenient and compact form as

$$\frac{dn_1}{dt} = -a_1n_1^2 + b_1n_1 + d_1, \quad (1)$$

$$\frac{dn_2}{dt} = -a_2n_2^2 + b_2n_2 + d_2, \quad (2)$$

$$\frac{dE_1}{dt} = \frac{4\pi}{\varepsilon K}[\zeta - E_1 - E_2 - Ken_1\mu_1E_1], \quad (3)$$

$$\frac{dE_2}{dt} = \frac{4\pi}{\varepsilon K}[\zeta - E_2 - E_1 - Ken_2\mu_2E_2], \quad (4)$$

where the following notation is used:

$$a_1 = A_I^{(1)} + B_T^{(1)}, \quad a_2 = A_I^{(2)} + B_T^{(2)},$$
$$b_1 = -\gamma_{opt} - B_T^{(1)}N_dC + A_I^{(2)}N_d(1 - C), \quad (5)$$
$$b_2 = -\gamma_T - B_T^{(2)}N_dC + A_I^{(2)}N_d(1 - C).$$

Here, $d_1 = \gamma_{opt}N_d(1 - C)$, $d_2 = \gamma_TN_d(1 - C)$, $\zeta = \mathscr{E}/L$, C is the compensation factor, and $K = SR/L$ is a quantity with the dimensions of resistivity.

Since our studies are qualitative, we tried to use comparatively exact and as simple as possible expressions for the kinetic coefficients. In order to obtain the dependence $A_1(Z)$, we used a formula derived for the conditions of a constant ionization cross section [8], and, to obtain the dependence $B_T(Z)$, we use a formula calculated for a corrected Lax cascade capture [9]; however, we restrict the analysis to an accuracy of no higher than that provided by the $\sim Z^{-3/2}$ term. In a magnetic field, the following quantity plays the role of mobility (see, for example, [10]):

$$\mu = \frac{e}{m}\left\langle \frac{\tau}{1 + \omega_c^2\tau^2} \right\rangle. \quad (6)$$

Here, τ is the momentum-relaxation time, $\omega_c = eH/mc$ is the cyclotron frequency, H is the magnetic-field strength, and $\langle \ldots \rangle$ signifies averaging over the energy. In the case of a classically strong magnetic field ($\omega_c^2\tau^2 \gg 1$), formula (6) can be simplified and written as

$$\mu = \frac{e}{m}\frac{1}{\omega_c^2}\left\langle \frac{1}{\tau} \right\rangle. \quad (7)$$

It follows from (7) that $\mu(Z) \propto \tau^{-1}(Z)$. When momentum scattering by charged impurity atoms is prevalent, we can use the classical Brooks–Herring formula to find that $\mu(Z) \propto Z^{-3/2}$. This dependence is extremely favorable for the origination of nonlinear self-sustained oscillations [5, 6]. Taking this circumstance into account, we choose the sample parameters and external conditions in the computer simulation so that the prevalence of the above momentum-scattering mechanism is ensured. Specifically, we consider *n*-Ge with $N_D = 10^{16}$ cm^{-3} and $C = 0.9$ at $T = 4.2$ K and $H = 10^4$ Oe. In these conditions, the dependences of the kinetic coefficients on the dimensionless electron temperature are given by [8–10]

$$A_I(Z) \approx 1.25 \times 10^{-6}\sqrt{Z}\left(1 + \frac{27.6}{Z}\right)\exp\left(-\frac{27.6}{Z}\right),$$

$$B_T(Z) \approx \frac{3.36 \times 10^{-6}}{Z^{3/2}}, \qquad (8)$$

$$\mu(Z) \approx \frac{4.6 \times 10^7}{Z^{3/2}}.$$

The interrelation between the electron temperature and an applied electric field is determined from an energy-balance equation that, in the general case, can be written as [11]

$$e\mu(Z)E^2 = \left\langle \frac{dW(Z)}{dt} \right\rangle_{\text{ac, ph}} + \left\langle \frac{dW(Z)}{dt} \right\rangle_{\text{exc}} + \left\langle \frac{dW(Z)}{dt} \right\rangle_{\text{ion}}, \qquad (9)$$

where $\langle dW/dt\rangle_{\text{ac, ph}}$, $\langle dW/dt\rangle_{\text{exc}}$, and $\langle dW/dt\rangle_{\text{ion}}$ are the mean rates of the energy losses of electron gas (per electron) due to interaction with acoustic phonons and from the excitation and ionization of donor impurities, respectively. Since the impurity-related mechanisms are fairly complex, we use, in the case under consideration, the following approximate formulas derived on the basis of Eq. (9) (here, an electric field is measured in the cgse units):

$$Z \approx -0.93 + 13.04E + 4.68E^2 - 0.59E^3,$$
$$E \approx -2.34 \times 10^{-3} + 6.1 \times 10^{-2}Z - 3.8 \times 10^{-3}Z^2 + 2.2 \times 10^{-6}Z^3. \qquad (10)$$

3. THE EQUILIBRIUM POINTS

Setting the right-hand sides of Eqs. (1)–(4) equal to zero, we obtain a system of nonlinear algebraic equations that determines the stationary values of the variables for a specified set of samples under fixed external parameters of the problem. Taking into account the interrelation (formula (10)) between the electron temperature and electric field, we can write the above system as

$$\begin{cases} -a_1(E_1)n_1^2 + b_1(E_1)n_1 + d_1 = 0, & (11a) \\ -a_2(E_2)n_2^2 + b_2(E_2)n_2 + d_2 = 0, & (11b) \\ \zeta - E_1 - E_2 - Ken_1\mu_1(E_1)E_1 = 0, & (11c) \\ \zeta - E_2 - E_1 - Ken_2\mu_2(E_2)E_2 = 0. & (11d) \end{cases}$$

Equations (11a) and (11b) define the dependences $n_1(E_1)$ and $n_2(E_2)$:

$$n_1(E_1) = \frac{b_1(E_1) + \sqrt{b_1^2(E_1) + 4a_1(E_1)d_1}}{2a_1(E_1)}, \qquad (12)$$

$$n_2(E_2) = \frac{b_2(E_2) + \sqrt{b_2^2(E_2) + 4a_2(E_2)d_2}}{2a_2(E_2)}. \qquad (13)$$

Substituting (12) into Eq. (11c) and (13) into Eq. (11d), we obtain the following system of equations in E_1 and E_2:

$$\begin{cases} n_1(E_1)\mu_1(E_1)E_1 = n_2(E_2)\mu_2(E_2)E_2, & (14a) \\ \zeta - E_1 - E_2 - Ken_1(E_1)\mu_1(E_1)E_1 = 0. & (14b) \end{cases}$$

The solutions to this system depend on the parameters of the samples and on the external circuit. It is worth remembering that, according to the notation given by (5), K and ζ are defined as $K = SR/L$ and $\zeta = \mathscr{E}/L$, where L and S are the length and cross section of two samples identical in size. Thus, choosing the values of K and ζ, we thereby select the load resistance R and the emf $\mathscr{E}$ of the power supply for a specified set of the samples. Undoubtedly, system of Eqs. (14a)–(14b) can be solved directly using numerical methods; however, the numerical procedure has to be repeated again for each set of values of K and ζ. It is much more convenient if we use numerical methods for a specified set of samples to derive final universal dependences that would allow us to find the equilibrium points easily and, more importantly, to predict their number.

To this end, we use the following procedure: For a chosen value of ζ, we change E_1 with some step within the interval $(0, \zeta)$. For each value of E_1, we then solve Eq. (14a) and find all the possible values of E_2; finally, we substitute E_1 and E_2 into Eq. (14b) to find all the possible values of K. This procedure eventually makes it possible to determine the dependence $K(E_1)$ (see Figs. 1–3). Using the specified value of K and the curves in Figs. 1–3, we can determine the equilibrium values E_1^*; the corresponding equilibrium values E_2^*, n_1^*, and n_2^* can be determined using Eqs. (14b), (12), and (13), respectively.

The number of equilibrium points differs for different values of the system's parameters. For example, we

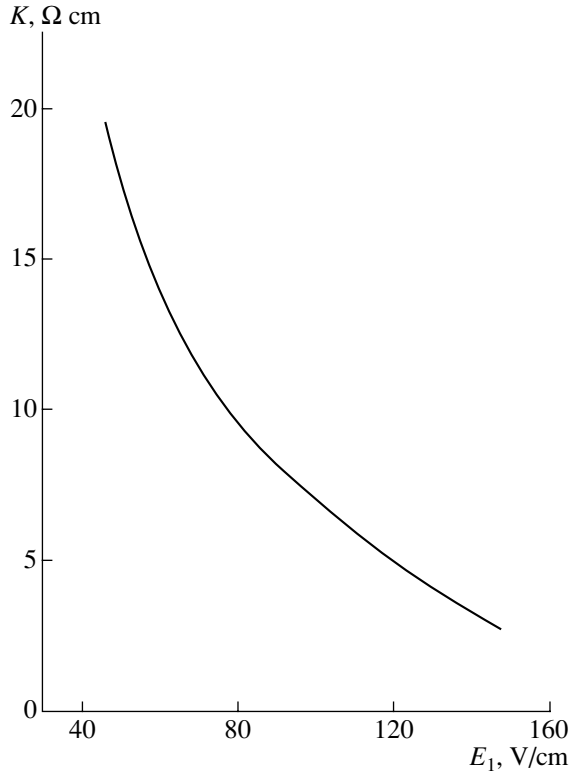

Fig. 1. The dependence $K(E_1)$ at $\gamma_{opt} = 5 \times 10^9$ s^{-1} and ζ = 720 V/cm.

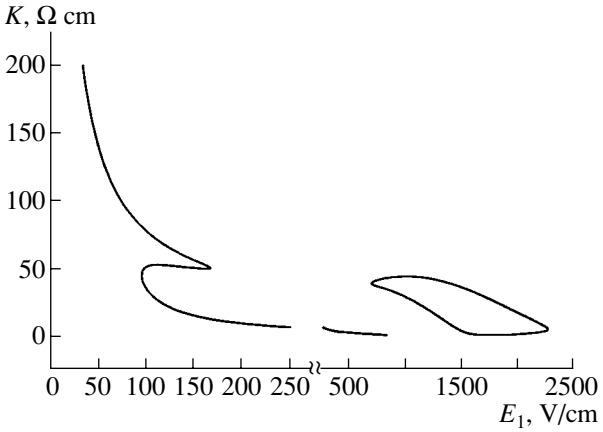

Fig. 2. The dependence $K(E_1)$ at $\gamma_{opt} = 5 \times 10^9$ s^{-1} and ζ = 3000 V/cm.

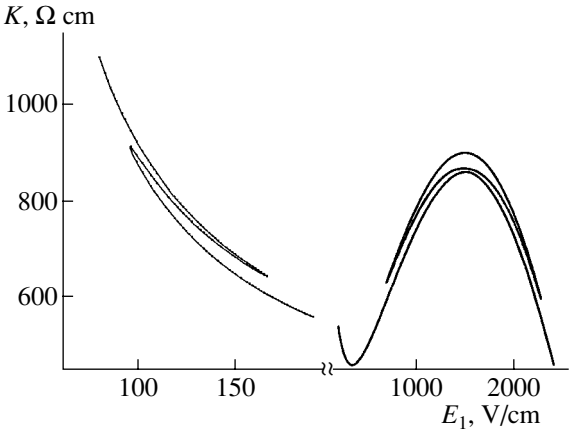

Fig. 3. The dependence $K(E_1)$ at $\gamma_{opt} = 5 \times 10^9$ s^{-1} and $\zeta = 3 \times 10^4$ V/cm.

have only a single equilibrium point in the case of $j_{opt} = 5 \times 10^9$ s^{-1} and ζ = 720 V/cm, irrespective of the value of K (Fig. 1). We can have one, two, or three equilibrium points at ζ = 3000 V/cm, depending on the value of K (Fig. 2). The number of equilibrium points can exceed three for larger values of K (Fig. 3).

4. RESULTS OF THE COMPUTER SIMULATION

For convenience, we rewrite Eqs. (1)–(4) in dimensionless variables:

$$\frac{dX_1}{dt} = -a_1 n_1^* (1 + X_1)^2 + b_1(1 + X_1) + \frac{d_1}{n_1^*}, \quad (15)$$

$$\frac{dX_2}{dt} = -a_2 n_2^* (1 + X_2)^2 + b_2(1 + X_2) + \frac{d_2}{n_2^*}, \quad (16)$$

$$\frac{dY_1}{dt} = \frac{4\pi}{\varepsilon K}\left[\frac{\zeta}{E_1^*} - 1 - Y_1 - \frac{E_2^*}{E_1^*}(1 + Y_2) - Ke\mu_1 n_1^*(1 + X_1)(1 + Y_1)\right], \quad (17)$$

$$\frac{dY_2}{dt} = \frac{4\pi}{\varepsilon K}\left[\frac{\zeta}{E_2^*} - 1 - Y_2 - \frac{E_1^*}{E_2^*}(1 + Y_1) - Ke\mu_2 n_2^*(1 + X_2)(1 + Y_2)\right]. \quad (18)$$

Here, $X_1 \equiv (n_1 - n_1^*)/n_1^*$, $Y_1 \equiv (E_1 - E_1^*)/E_1^*$, $X_2 \equiv (n_2 - n_2^*)/n_2^*$, and $Y_2 \equiv (E_2 - E_2^*)/E_2^*$.

The simulation was carried out for various values of the illumination intensity. Below, we report the results obtained at $j_{opt} = 5 \times 10^9$ s^{-1}, for which we obtain the most interesting and diverse oscillation dynamics.

Chaotic self-sustained oscillations (in addition to regular self-sustained oscillations) were observed at the above illumination intensity. It is important that all three universal scenarios of chaos origination were realized, depending on the value of ζ (and, consequently, on the value of $\mathscr{E}$ for a given sample), in the form of a bifurcation parameter. At small values of ζ, the chaotic oscillations originated according to the Feigenbaum scenario (see, for example, [12]). Chaos was observed in both the illuminated and unilluminated parts of the samples. The system features a single equilibrium point for the value of ζ under consideration (Fig. 1). A transition of the system from regular oscillations to chaos was accomplished by varying the second

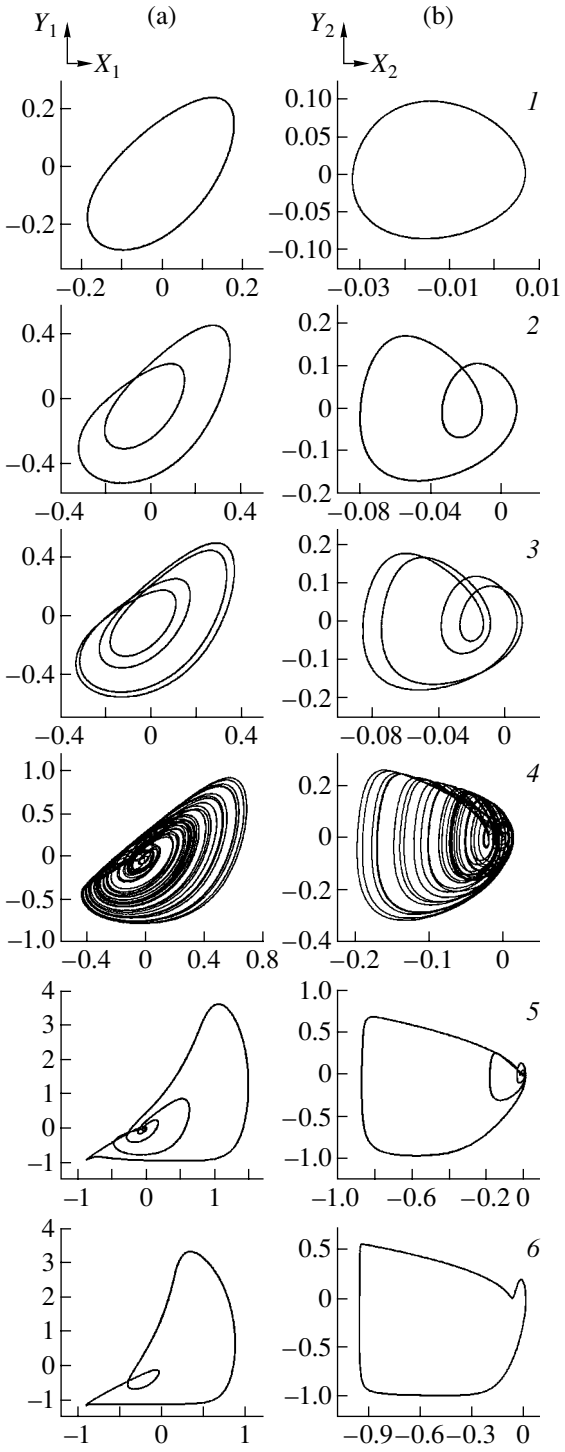

Fig. 4. Phase portraits in the (a) (x_1, y_1) and (b) (x_2, y_2) planes at ζ = 720 V/cm and K = (*1*) 7.57, (*2*) 7.18, (*3*) 7.15, (*4*) 6.92, (*5*) 6.33, and (*6*) 2.94 Ω cm. Chaotic oscillations are generated according to the Feigenbaum scenario.

bifurcation parameter K (i.e., R for a given sample). In particular, the following pattern is observed at ζ = 720 V/cm (Fig. 4): (i) the oscillations are regular at K = 7.57 Ω cm (E_1^* = 94.8 V/cm, E_2^* = 391.8 V/cm, n_1^* = 1.74×10^{14} cm^{-3}, and $n_2^* = 3.98 \times 10^{14}$ cm^{-3}) (Fig. 4, the top row); (ii) a doubling of the period is observed at K = 7.18 Ω cm (E_1^* = 98.1 V/cm, E_2^* = 394.8 V/cm, $n_1^* = 1.80 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.11 \times 10^{14}$ cm^{-3}) (Fig. 4, the second row from the top); (iii) a fourfold increase in the period is observed at K = 7.15 Ω cm (E_1^* = 98.7 V/cm, E_2^* = 395.4 V/cm, $n_1^* = 1.83 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.16 \times 10^{14}$ cm^{-3}) (Fig. 4, the third row); (iv) the oscillations are chaotic at K = 6.92 Ω cm (E_1^* = 100.2 V/cm, E_2^* = 396.9 V/cm, $n_1^* = 1.85 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.20 \times 10^{14}$ cm^{-3}) (Fig. 4, the fourth row); (v) a fourfold increase in the period is again observed at K = 6.33 Ω cm (E_1^* = 105.9 V/cm, E_2^* = 402.3 V/cm, $n_1^* = 1.97 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.42 \times 10^{14}$ cm^{-3}) (Fig. 4, the fifth row); and (vi) a doubling of the period is again observed at K = 2.94 Ω cm (E_1^* = 144.9 V/cm, $E_2^* = 452.7$ V/cm, $n_1^* = 2.82 \times 10^{14}$ cm^{-3}, and $n_2^* = 6.01 \times 10^{14}$ cm^{-3}) (Fig. 4, the sixth row). A further decrease in K gives rise to regular oscillations.

Chaos originates via alternation at comparatively large values of ζ [12], as is illustrated in Fig. 5 at ζ = 3000 V/cm. Under these conditions, the system can exhibit one or three equilibrium points (Fig. 2). As K increases, the oscillation dynamics of the system varies in the following way:

(1) at K = 2.5 Ω cm, we have three equilibrium points: (*a*) with E_1^* = 417 V/cm, E_2^* = 2436 V/cm, $n_1^* = 7.82 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.72 \times 10^{14}$ cm^{-3}; (*b*) with E_1^* = 1512 V/cm, E_2^* = 1410 V/cm, $n_1^* = 9.84 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.75 \times 10^{14}$ cm^{-3}; and (*c*) with E_1^* = 2244 V/cm, E_2^* = 645 V/cm, $n_1^* = 9.85 \times 10^{14}$ cm^{-3}, and $n_2^* = 8.48 \times 10^{14}$ cm^{-3}. It is noteworthy that point *a* is a stable focus, point *c* is a stable node, and point *b* is unstable. Consequently, the phase trajectory leaves point *b* and tends to points *c* or *a*, depending on the initial conditions (in the first row from the top in Fig. 5, the tendency towards point *a* is illustrated, and the second equilibrium point corresponds to the zero values of the variables). Thus, the oscillations are either not generated at all or are damped.

(2) K = 41.3 Ω cm (Fig. 5, the second row). All three equilibrium points, (*a*) with E_1^* = 96 V/cm, E_2^* = 1623 V/cm, $n_1^* = 1.76 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.79 \times 10^{14}$ cm^{-3}; (*b*) with E_1^* = 765 V/cm, E_2^* = 462 V/cm, $n_1^* = 9.4 \times 10^{14}$ cm^{-3}, and $n_2^* = 6.23 \times 10^{14}$ cm^{-3}; and (*c*) with E_1^* = 1257 V/cm, E_2^* = 400 V/cm, n_1^* =

9.78×10^{14} cm^{-3}, and $n_2^* = 4.32 \times 10^{14}$ cm^{-3}, are unstable. The limiting cycle is closed around the first point. The oscillations are regular.

(3) $K = 41.5\ \Omega$ cm (Fig. 5, the third row). All three equilibrium points, (*a*) with $E_1^* = 98$ V/cm, $E_2^* = 1626$ V/cm, $n_1^* = 1.77 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.80 \times 10^{14}$ cm^{-3}; (*b*) with $E_1^* = 774$ V/cm, $E_2^* = 462$ V/cm, $n_1^* = 9.42 \times 10^{14}$ cm^{-3}, and $n_2^* = 6.23 \times 10^{14}$ cm^{-3}; and (*c*) with $E_1^* = 1257$ V/cm, $E_2^* = 400$ V/cm, $n_1^* = 9.78 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.32 \times 10^{14}$ cm^{-3}, are again unstable; however, in contrast to the case 1, the second point (in addition to the first point) is also involved in the formation of the regular oscillatory behavior of the system. In particular, the first point is an unstable focus; the phase point moves along the unwinding spiral in the vicinity of this focus. This pattern is retained until the phase point is found in close proximity to the second equilibrium point. Then, the phase trajectory makes one turn about the second point, arrives in the vicinity of the first point, and the process occurs over and over; i.e., complex regular self-sustained oscillations are generated.

(4) $K = 51.3\ \Omega$ cm (Fig. 5, the fourth row from the top). All three unstable equilibrium points are involved in the formation of oscillatory behavior. These points correspond to (*a*) $E_1^* = 100.5$ V/cm, $E_2^* = 1248$ V/cm, $n_1^* = 1.85 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.68 \times 10^{14}$ cm^{-3}; (*b*) $E_1^* = 138$ V/cm, $E_2^* = 794.7$ V/cm, $n_1^* = 2.66 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.12 \times 10^{14}$ cm^{-3}; and (*c*) $E_1^* = 165.3$ V/cm, $E_2^* = 519.3$ V/cm, $n_1^* = 3.28 \times 10^{14}$ cm^{-3}, and $n_2^* = 7.29 \times 10^{14}$ cm^{-3}. The phase trajectory exhibits a chaotic alternation and makes a different number of turns sequentially about the first, second, and third equilibrium points. These points form an attractive conglomerate (a strange attractor). The oscillations are chaotic.

(5) $K = 119.1\ \Omega$ cm (Fig. 5, the bottom row). In this case, we have a single equilibrium point ($E_1^* = 61.86$ V/cm, $E_2^* = 366$ V/cm, $n_1^* = 1.08 \times 10^{14}$ cm^{-3}, and $n_2^* = 2.67 \times 10^{14}$ cm^{-3}). The oscillations are regular.

It is worth noting that the above refers to the illuminated part of the semiconductor. Deviations from regularity are almost unobserved in the unilluminated part.

For larger values of ζ, the chaotic oscillations originate according to the Ruelle–Takens–Newhouse scenario [12]. This pattern is illustrated in Fig. 6 for the case of $\zeta = 3 \times 10^4$ V/cm. The variations in the system's behavior as K increases are as follows.

(1) $K = 890\ \Omega$ cm (Fig. 6, panel *1*). We have the following three equilibrium points: (*a*) with $E_1^* = 100.2$ V/cm, $E_2^* = 1250$ V/cm, $n_1^* = 1.84 \times 10^{14}$ cm^{-3}, and $n_2^* = 9.68 \times 10^{14}$ cm^{-3}; (*b*) with $E_1^* = 1338$ V/cm, $E_2^* = 395.4$ V/cm, $n_1^* = 9.8 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.13 \times 10^{14}$ cm^{-3}; and (*c*) with $E_1^* = 1650$ V/cm, $E_2^* = 395$ V/cm, $n_1^* = 9.85 \times 10^{14}$ cm^{-3}, and $n_2^* = 4.12 \times 10^{14}$ cm^{-3}. All of these points are unstable. The limiting cycle is closed around the first point.

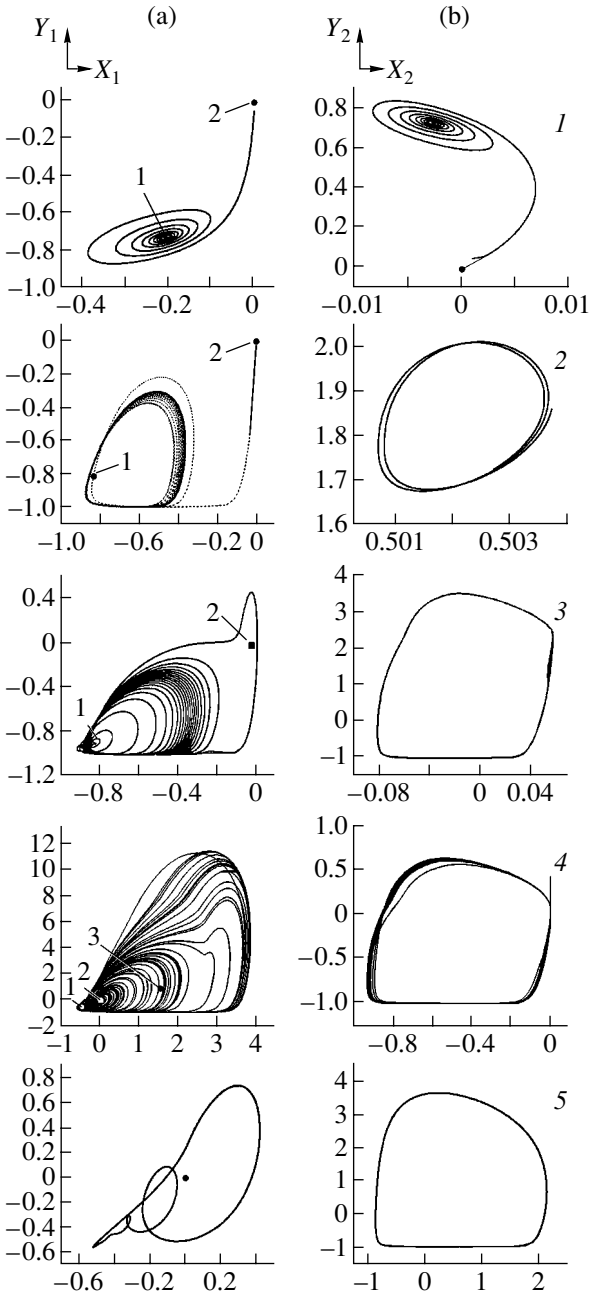

Fig. 5. Phase portraits in the (a) (x_1, y_1) and (b) (x_2, y_2) planes at $\zeta = 3000$ V/cm and $K =$ (*1*) 2.5, (*2*) 41.3, (*3*) 41.5, (*4*) 51.3, and (*5*) 119.1 Ω cm. Chaotic oscillations are generated owing to alternation.

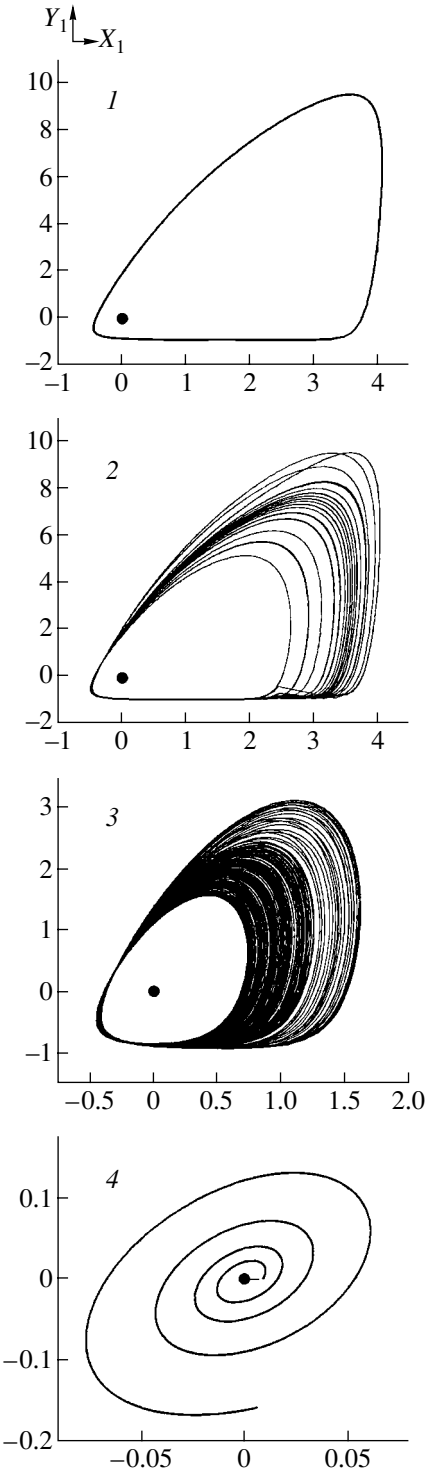

Fig. 6. Phase portraits in the (x_1, y_1) plane at $\zeta = 3 \times 10^4$ V/cm and K = (*1*) 890, (*2*) 917, (*3*) 1086, and (*4*) 1487 Ω cm. Chaotic oscillations are generated according to the Ruelle–Takens–Newhouse scenario.

At $K < 875$ Ω cm, additional stable points of equilibrium, resulting in the disappearance of undamped oscillations, appear in the system (see Fig. 3).

(2) $K = 917$ Ω cm (Fig. 6, panel *2*). There is a single unstable equilibrium point. The oscillations are quasi-periodic.

(3) $K = 1086$ Ω cm (Fig. 6, panel *3*). Quasi-periodicity transforms into chaos.

(4) $K = 1487$ Ω cm (Fig. 6, panel *4*). A single equilibrium point is represented by a stable focus, meaning that the oscillations are damped.

As the illumination intensity varies, the pattern of system's behavior changes radically. For example, at $j_{opt} = 5 \times 10^8$ s^{-1}, the oscillations are always regular at any values of the bifurcation parameters. As the illumination intensity is decreased further, the oscillations are not observed at all.

5. DISCUSSION

As can be seen, radically different oscillation dynamics is observed depending on the illumination intensity, emf of the power supply, and the load resistance. In what follows, we attempt to physically interpret some aspects of the results obtained.

Since the free-electron concentration is higher in the illuminated part of the semiconductor, the major fraction of the voltage is applied to its unilluminated part. As a result, this part experiences a breakdown and is the first to develop nonlinear oscillatory processes, which ultimately gives rise to oscillations in the illuminated part of the sample as well. In the general case, the instability of the system is controlled by the following three factors: (i) the extent to which the conduction-electron concentration is close to its limiting value ($N_D(1 - C)$); (ii) whether or not there is a breakdown in the unilluminated part of the semiconductor; and (iii) the extent of the effect of a decrease in the mobility (see formula (8)) on the current flowing in the semiconductor as the electric field (and, consequently, the electron temperature) increases.

At $j_{opt} = 5 \times 10^9$ s^{-1} and at small values of ζ ($\zeta = 720$ V/cm), the equilibrium free-electron concentrations in both parts of the semiconductor are far from saturation for all the values of K under consideration. Furthermore, the unilluminated part of the semiconductor is either subjected to breakdown or is close to the breakdown point; consequently, it is quite clear that the system is unstable and this instability is related to the concentration.

The system exhibits several equilibrium points at larger values of ζ (for example, at $\zeta = 3000$ V/cm). The points for which the free-electron concentrations are fairly far from saturation and the unilluminated part of the semiconductor is either in a state of breakdown or is close to it are again unstable; however, in addition, instability also arises when the concentrations are eventually saturated but a decrease in the mobility as the electric field increases gives rise to a negative differential conductivity (NDC). In this study, in the case where a Lorentz attractor is formed, the second point exhibits

the instability under consideration. For this point, the equilibrium electric fields applied to the illuminated and unilluminated parts of the semiconductor are approximately equal to each other. Apparently, this point is of a saddle type. Even the smallest deviation from equilibrium leads to a redistribution of the field between different parts of the semiconductor; under certain conditions, this redistribution is chaotic. Thus, mobility-related instability (in addition to concentration-related instability) is important in the case under consideration. It is noteworthy that, even if only one of the two parts of the semiconductor in the equilibrium state is outside the NDC region, this state is stable and the oscillations either do not develop at all or are damped.

6. PRACTICAL IMPLEMENTATION OF THE RESULTS OBTAINED

In our opinion, the results obtained are of interest from the standpoint of their practical implementation. As we have shown, the system's dynamics depends, to a great extent, on the values of a number of easily controllable parameters, such as the emf of the power supply, the load resistance, and the illumination intensity. By combining these parameters, it becomes possible to transfer the corresponding high-frequency oscillator from the regular signal mode to a mode of random-number generation and vice versa; alternatively, the oscillator can be switched off completely.

ACKNOWLEDGMENTS

This study was supported by the International Science and Technology Centre, grant no. G-394.

REFERENCES

1. N. Balkan, *Hot Electrons in Semiconductors* (Clarendon, Oxford, 1998), Part 2, p. 209.
2. R. P. Huebener, J. Peinke, and J. Parisi, Appl. Phys. A **48**, 107 (1989).
3. G. Hupper and E. Scholl, Phys. Rev. Lett. **66**, 2372 (1991).
4. Z. S. Kachlishvili and K. M. Jandieri, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 340 (1998) [JETP Lett. **67**, 358 (1998)].
5. Z. S. Kachlishvili and K. M. Jandieri, Pis'ma Zh. Tekh. Fiz. **23** (16), 62 (1997) [Tech. Phys. Lett. **23**, 643 (1997)].
6. K. M. Jandieri and Z. S. Kachlishvili, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 909 (2001) [Semiconductors **35**, 873 (2001)].
7. E. Schoell, Solid-State Electron. **31**, 539 (1988).
8. Z. S. Kachlishvili, Fiz. Tekh. Poluprovodn. (Leningrad) **2**, 580 (1968) [Sov. Phys. Semicond. **2**, 478 (1968)].
9. M. Lax, Phys. Rev. **119**, 1502 (1960).
10. V. L. Bonch-Bruevich and S. G. Kalashnikov, *Physics of Semiconductors* (Nauka, Moscow, 1977) [in Russian].
11. Z. S. Kachlishvili, Phys. Status Solidi B **48**, 65 (1971).
12. H. G. Schuster, *Deterministic Chaos: an Introduction* (Physik, Weinheim, 1984; Mir, Moscow, 1988).

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 650–657. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 681–688.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Modulation of the Characteristics of Intense Picosecond Stimulated Emission from GaAs

N. N. Ageeva*, I. L. Bronevoĭ*^, A. N. Krivonosov*, S. E. Kumekov, T. A. Nalet***, and S. V. Stegantsov***

**Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Moscow, 125009 Russia*
^e-mail: bil@mail.cplire.ru
***Kazakhstan National Technological University, Almaty, 480013 Kazakhstan*
****Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia*
Submitted July 6, 2004; accepted for publication November 2, 2004

Abstract—Intense picosecond stimulated emission from the face of a thin GaAs film is studied. The emission is observed when GaAs is pumped with high-power picosecond optical pulses. It is found that the dependences of the emission energy on the photon energy, the picosecond delay between two pump pulses, and the distance between the active region and the face are modulated. Modulation is taken to mean the appearance of protrusions or peaks in any of the mentioned dependences. The modulation parameters for the dependences under consideration are found to be related by expressions that make it possible to suggest the following. The modulation of characteristics is caused by a common (not yet conclusively identified) mechanism of self-modulation of the emission spectrum. This mechanism is related to an ultrafast nonlinear interaction between a highly photoexcited semiconductor and the pump radiation and stimulated emission. There is indirect evidence that this mechanism also gives rise to an amplitude modulation of the emission in a picosecond time interval. © *2005 Pleiades Publishing, Inc.*

The importance of studying the intense picosecond emission generated in a semiconductor is related to an inadequate knowledge of the kinetics of ultrafast interaction between a highly photoexcited semiconductor and high-intensity inherent stimulated emission. Here and below, "ultrafast" means no slower than in the picosecond time range. Interaction between stimulated emission and nonequilibrium charge carriers can manifest itself in the variations in a number of characteristics of both charge carriers and emission. For example, this interaction can give rise to an appreciable spatial inhomogeneity and picosecond pulsations in the temperature and charge-carrier concentration during the emission relaxation after an ultrashort GaAs pumping, as numerical simulation has shown [1]. In particular, the spatial nonuniformity of the charge-carrier concentration, optical gain, and the emission intensity have been the focus of previous analysis [2–4]; however, the charge-carrier temperature was disregarded in these studies. The interaction between emission and charge carriers can give rise to various phenomena related to the emission modulation [5, 6]. It is known that the modulation phenomena are highly sensitive to even a very weak feedback or, for example, to the presence of even insignificant passive (i.e., without a population inversion) regions [5]. Emission modulation resulting from interaction with a passive medium is particularly important in the case of the generation of ultrashort optical pulses in semiconductor lasers. Manifestation of the interaction between radiation and charge carriers via the modulation of the emission characteristics also makes study of this interaction appropriate in the context of practical implementations. Specifically, such study can lead to the observation of modulation processes that affect, adversely or beneficially, the generation of ultrashort pulses in semiconductor lasers, the potential for the optical transmission of information, and, generally, the operation of devices employing ultrafast semiconductor optoelectronics in a situation in which stimulated emission is used. Naturally, it is important to find methods for controlling the modulation.

The modulation of the emission characteristics observed in this study belongs to the aforementioned range of problems. The characteristics of the intense picosecond stimulated emission generated during the formation of a hot dense electron–hole plasma in GaAs are found to be modulated. The plasma was formed as a result of pumping the thin GaAs layer with picosecond optical pulses. The observed modulation exhibits a number of nontrivial features. We note once again that, in this study, modulation is understood as the appearance of local protrusions or peaks in the plot of a specific characteristic. The appearance of local protrusions in the emission spectrum was observed when this spectrum was measured with a sufficiently high resolution. It was found that a smooth spectrum was replaced alternately by a modulated spectrum. This transformation of the spectra occurred as the distance between the active

region in the GaAs layer and the cross-sectional layer face from which the measured emission emerged increased. Dependence of the emission energy (integrated over the spectral range under study) on the distance between the active region and the face was also found to be modulated. The emission-spectrum modulation was much deeper than might be expected if the cavity had been formed by the sample ends. The interval between local peaks in the dependence of the emission energy on the distance between the active region and the sample end was two orders of magnitude larger than the emission wavelength. Additional excitation of the semiconductor by a second pulse, which followed after the first pulse with a picosecond-scale delay τ and excited electrons to the energy level located in the conduction band but far below the level excitation resulting from the first pulse, also made it possible to observe the following behavior. It was found out that the dependence of the emission energy (with a fixed photon energy) on the delay time τ became modulated under certain conditions, which are described in detail below. We compared the intervals between local maxima in the dependences of the emission energy on (1) the delay time τ, (2) the emitted-photon energy, and (3) the distance between the active region and the sample end. We found that these intervals were related by expressions that suggested the following. The modulation of the emission characteristics is caused by a common (as yet unidentified) mechanism of emission-spectrum self-modulation under the conditions of an ultrafast interaction between a highly photoexcited semiconductor and pump radiation and stimulated emission. Apparently, this mechanism gives rise to an amplitude modulation of the emission on the picosecond time scale, which accompanies the spectrum modulation. Thus, it is possible that the observed modulation of the emission characteristics exhibits, in turn, a picosecond-scale amplitude modulation of the emission. The self-modulation mechanism appears and disappears (or is appreciably enhanced and attenuated) quasi-periodically as a result of variations in (i) the path traversed by the emission on its way from the active region to the sample end and back and (ii) the delay time τ in the situation where two pump pulses are used. In both cases, quasi-periodicity corresponded to an estimate, obtained on the basis of the spectrum modulation, of the amplitude-modulation period. In what follows, we provide a detailed description of the experimental results obtained in this study.

We studied the picosecond stimulated emission that emerged from the face of a thin (~1 μm) GaAs layer at room temperature. High-power 12-ps pulses of linearly polarized light were used for pumping; as a result, a dense ($n = p > 10^{18}$ cm^{-3}) electron–hole plasma was formed in GaAs. The degree of population inversion was high enough to cause stimulated emission during the pumping [7]. The emission intensity was estimated at $>10^8$ W/cm^2. The main decay of the emission intensity should be completed in 60 ps (under the conditions of this study) after the pump is switched on. The stage including a slow decay of emission with much lower intensity should follow the above main decay [8].

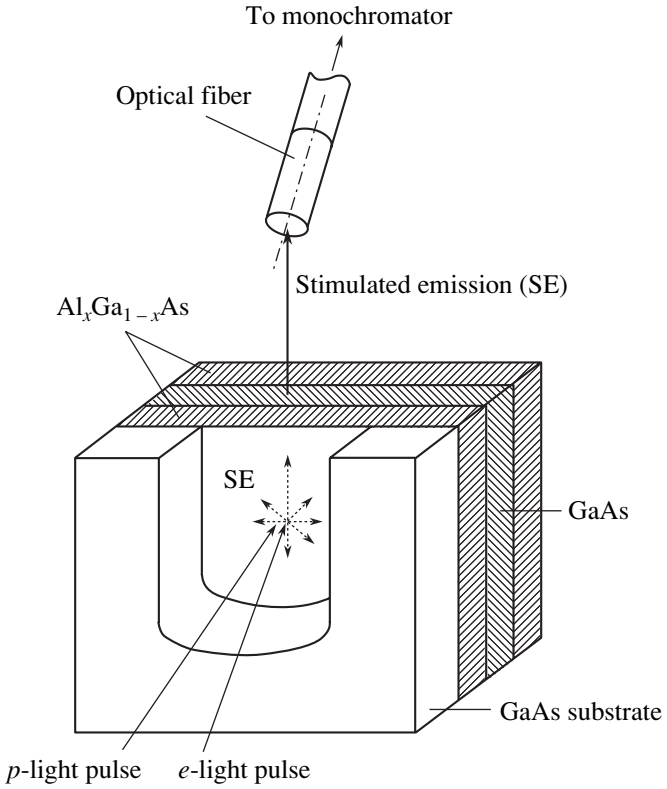

Fig. 1. Schematic representation of the experiment.

The sample under study was an $Al_{0.22}Ga_{0.78}As$–GaAs–$Al_{0.4}Ga_{0.6}As$ heterostructure with layer thicknesses of 1.2, 1.6, and 1.2 μm, respectively. The heterostructure was grown by molecular-beam epitaxy on a GaAs (100) substrate. The largest part of the heterostructure area (including one of its side faces) was then relieved of the substrate. The remaining part of the substrate fringed the heterostructure from three sides, in a framelike form (Fig. 1). The concentrations of donor and acceptor impurities in the heterostructure were no higher than 10^{15} cm^{-3}. The $Al_xGa_{1-x}As$ layers were intended for stabilization of the surface recombination and for mechanical strength; furthermore, they were transparent in the light used in our experiments. An antireflection coating was deposited onto the outer surfaces of the $Al_xGa_{1-x}As$ layers. As a result, the reflectance of light along the normal to the layers' surfaces was no larger than 2%. An antireflection coating was not deposited onto faces of the structure.

The heterostructure was irradiated with either one (*e*) or two (*e* and *p*) pulses focused to a single spot in the area from which the substrate had been removed. The pulses *e* were incident on the sample at an angle of 10° with respect to the normal to the sample surface, whereas pulses *p* were incident approximately at the

normal to the surface. The duration of each pulse, determined at the half-height of time distribution of the light intensity, was 12 ps, whereas the duration determined at the pulse pedestal was about 28 ps. The spatial distribution of the pulses' energy was approximately Gaussian, and the diameter of the focus spot (at the half-height of this distribution) was ~0.5 mm. The delay time τ of the *p* pulse relative to the *e* pulse was varied within the picosecond scale. This variation was attained by changing the optical path for the *p* pulse. According to estimations, the possible displacement of the focus spot for a *p* beam in the heterostructure plane could be no larger than 0.07 μm as the value of τ was changed by 1 ps. The photon energy in an *e* pulse was $\hbar\omega_e = 1.56$ eV and that in a *p* pulse was $\hbar\omega_p = 1.43$ eV (in our opinion, it was desirable that the *p* and *e* pulses excited charge carriers to different energy levels). Emission of LO phonons by the electrons generated by *p* pulses with the aforementioned $\hbar\omega_p$ should facilitate the transfer of these electrons to the level at which they experience stimulated recombination [9]. Excitation of the charge carriers with an *e* pulse was much more intense than that with a *p* pulse. For example, if the sample was only pumped by a *p* pulse, the time-integrated energy W_e of emission with $\hbar\omega_s = 1.386$ eV was an order of magnitude lower than in the case of pumping by an *e* pulse (here, $\hbar\omega_s$ is the photon energy in stimulated emission). The results of our earlier measurements provided some insight into the enhancement of emission as a result of the population inversion [10]. We found [10] that, in the case of pumping by an *e* pulse (of the same type as in this study), the optical gain in the active region (equidistant from the sample's ends) increased from 90 cm^{-1} at $\hbar\omega_s = 1.367$ eV to 440 cm^{-1} at $\hbar\omega_s = 1.41$ eV.

In the present study, stimulated emission propagated predominantly in the plane of the heterostructure exhibiting the properties of a waveguide. The fraction of the emission under study emerged from the face of the region relieved of the substrate (Fig. 1), entered a quartz optical waveguide with a length of 1.2 m, and was finally fed into a monochromator. The aperture angle of the emission under study was limited by the diameter of the light guide's strand (1 mm) and the distance between the light guide end and the heterostructure's face (10 mm). The axis of the aperture angle was located approximately in the heterostructure plane, passed through the focus spot of the pumping, was orthogonal to the face of the heterostructure, and made an angle of 3.5° with the normal to the entrance of the light guide. The latter circumstance ensured that a cavity could not form between the light-guide end and the face of the sample. The emission that left the exit of the light guide was focused onto the monochromator slit. The spatial positions of the light guide and the optical-pump focus remained unchanged during each of the experiments. A variation in the distance from the focus spot (i.e., the active region in the GaAs layer) to the face of the heterostructure was accomplished by displacing the sample in the direction of the emission under study. It was difficult to accurately determine the distance between the focus spot and the face owing to the spatial extent of the spot. Therefore, in what follows, we mention the displacement δY from a certain initial position of the sample. In each particular experiment, the samples were positioned so that the distance between the focus-spot center and the face of the sample was equal to 1.2 mm to within a satisfactory accuracy. At this distance, the emission reflected by the ends was returned to the active region immediately upon completion of the pump pulse. A displacement δY that decreased the distance between the active region and the face of the sample was considered to be positive. We restricted our studies to only a part of the emission spectrum, since we had to carry out measurements involving a very small step. For brevity, we refer to this part simply as the spectrum. It is worth noting that, in each particular experiment, we measured one or several of the dependences reported below. When comparing the plots of these dependences, it should be taken into account that the conditions of pumping in the experiments that were carried out on different days could only be reproduced to a limited accuracy, with reference to, in particular, the focus-spot diameter, the initial position of the spot with respect to the face of the sample, and, correspondingly, the displacement δY. However, this circumstance did not affect the main conclusions of this study.

We studied first the emission that was observed when the sample was pumped by an *e* pulse. As the displacement δY was increased (i.e., as the active region became closer to the face), the time-integrated emission energy W_e increased (see the inset in Fig. 2). This effect could be caused by a decrease in the emission absorption in the passive (unirradiated with an *e* pulse) GaAs region, since this region was narrowed as δY increased. At a fixed displacement δY, the higher the emission-photon energy $\hbar\omega_s$ was, the higher the emission energy W_e became. Correspondingly, the spectrum of emission emerging from the face of the sample (i.e., the emission that had passed the passive region) differed from the spectrum of emission from the pumped (active) region of the GaAs layer. For example, the spectrum of emission from the active region, measured for the same sample and the same pumping pulse (see Fig. 2 in [10]), featured a peak at the photon energy $\hbar\omega_s \approx 1.39$ eV. In contrast, in this study, the portion of the spectrum involving the photon energy $\hbar\omega_s = 1.39$ eV was found at the falloff in the spectrum of emission that emerged from the face of the sample. It is noteworthy that the emission absorption in the passive region should give rise to a certain increase in the transmittance of this region.

As can be seen from the inset in Fig. 2, the dependence $W_e = f(\delta Y)$ is found to be almost periodically modulated at a fixed value of $\hbar\omega_s$. Measurements of the emission energy *S* integrated over the spectrum and time showed that the dependence $S = f(\delta Y)$ was modulated in the same way as the dependence $W_e = f(\delta Y)$, which was obtained simultaneously (Fig. 2). However,

as will be explained below, the dependence $W_e = f(\delta Y)$ is modulated in a different way from the dependence $S = f(\delta Y)$ at certain specific energies $\hbar\omega_s$. It is noteworthy that the emission energy shown in the plot in the inset is higher than that in the plot $W_e = f(\delta Y)$ in Fig. 2 itself. This circumstance apparently the reason for the quasi-period in the plot in the inset (~40 μm) being smaller than the quasi-period (the interval between the neighboring peaks) in the plot in Fig. 2 itself (~60 μm). The modulation of the dependence of the emission-spectrum integrated energy on the distance between the active region and the face with values of the quasi-period that exceeded the emission wavelength in GaAs by two orders of magnitude seemed, to a certain extent, to be unusual. This modulation cannot be attributed to simple interference by the generated emission and the radiation reflected from the face (i.e., the modulation cannot be explained in the domain of applicability of the superposition principle). It only remains to assume that the emission under study is a component of a nonlinear system that also includes a photoexcited semiconductor and pumping pulses.

The modulation of the dependence $S = f(\delta Y)$ signified, first, that the spectrum-integrated energy of emission generated in a chosen direction with respect to the face was affected by a feedback via the emission reflected from the face. Second, the emission itself appears to be amplitude-modulated. Since the active medium apparently interacts consistently with both the emission generated in this medium and the emission reflected from the face, we can assume the following. The modulation of the dependence $S = f(\delta Y)$ is representative of a variation in the phase difference $\varphi_1(r) - \varphi_2(r)$ within the active region (here, φ_1 is the phase of a generated wave that travels towards the face, φ_2 is the phase of a wave that returns to the active region after reflection from the face, and r are the coordinates of a point within the active region). In this case, a variation in the distance between the active region and the face by a value ensuring the variation in the phase difference $\varphi_1(r) - \varphi_2(r)$ by 2π is equal, on the one hand, to the quasi-period of the modulation of the dependence $S = f(\delta Y)$ $\Delta Y \approx 60$ μm and, on the other hand, to $L/4$, where L is the modulation wavelength. The variation in the phase difference $\varphi_1(r) - \varphi_2(r)$ caused by the displacement δY is approximately independent of r. As a result, we obtain the estimate $L \approx 240$ μm. Taking into account the fact that the modulation wave propagates with the group velocity $c_g = c_0/n_g$, we obtain an expression,

$$Tc_g \approx 4\Delta Y, \quad (1)$$

that makes it possible to estimate the modulation period: $T \approx 4$ ps. In expression (1), c_0 is the speed of light in free space and $n_g = 5.1$ is the "group" index of refraction [11, 12]. The value of n_g is caused by a pronounced dispersion in the semiconductor in the spectral region that involves the stimulated emission. It is noteworthy that the obtained estimate for the modulation

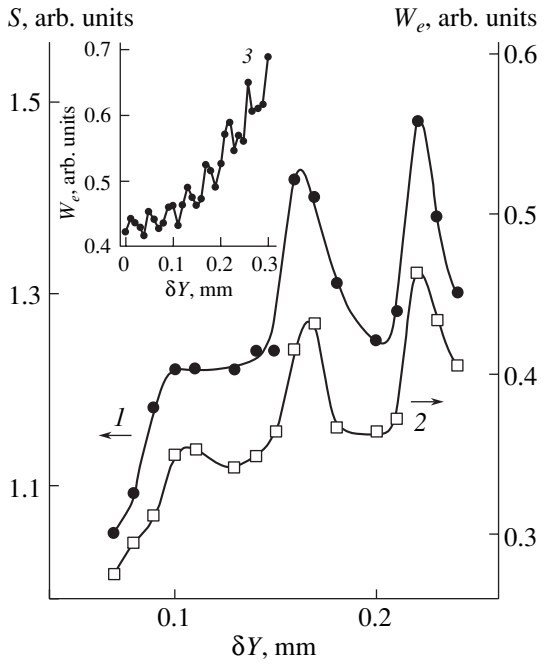

Fig. 2. Dependences of (*1*) the emission energy integrated over the spectrum (in the range $\hbar\omega_s = 1.38517$–1.38982 eV); (*2*) and (*3*) (in the inset) the emission energy W_e at the photon energy $\hbar\omega_s = 1.38578$ eV (curve *3* was measured in a separate experiment, where the emission energy was higher than in the measurements of curve *2*) on the displacement δY.

period T is consistent with the parameters of the observed modulation for two other emission characteristics (see below). First, the dependence of the emission energy on the delay τ between two pumping pulses was found to be modulated with approximately the same period (~4 ps). Second, the observed modulation of the emission spectrum corresponded to the time modulation of the emission with a period ~4 ps.

As was already mentioned, the modulation quasi-period for the dependence $S = f(\delta Y)$ differed from that for the dependence $W_e = f(\delta Y)$ (see the inset in Fig. 2); in the latter case, the quasi-period was equal to ~40 μm. For this quasi-period, a similar estimate for the period T was equal to 2.7 ps. This moderate difference from the estimate $T \approx 4$ ps obtained above apparently corresponds to the fact that, as was already mentioned, the energy of the emission at which the plot $W_e = f(\delta Y)$ was measured (see the inset) was higher than the energy of the same emission at which the plot $S = f(\delta Y)$ was measured.

As δY was varied (with a resolution of 0.01 mm), the observed shape of the emission spectrum $W_e = f(\hbar\omega_s)$ changed radically. Alternately, the smooth spectrum was replaced by a spectrum with local protrusions or steps, and then the spectrum became smooth again, and so on. This behavior is illustrated in Fig. 3 using the example of the spectra with the most clearly pronounced shape (i.e., the most sensitive to the value of

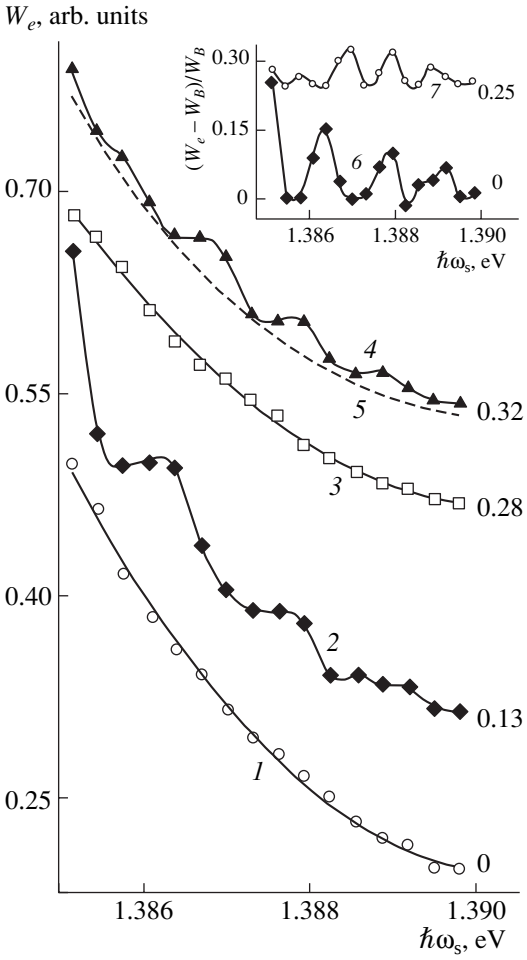

Fig. 3. Emission spectra at the displacements δY = (*1*) 0.16, (*2*) 0.18, (*3*) 0.20, and (*4*) 0.24 mm. Curve *5* represents the dependence $W_B = f(\hbar\omega_s)$ (see the text for details). In the inset, we show the relative-modulation spectra $(W_e - W_B)/W_B = f(\hbar\omega_s)$ for δY = (*6*) 0.18 and (*7*) 0.24 mm. For clarity, each curve is shifted along the vertical axis with respect to the true position; the value of the shift is written on the right of the curve.

the displacement δY). Each smooth spectrum can be adequately approximated with a portion of a parabolic curve. Therefore, the degree of modulation for the modulated spectrum $W_e = f(\hbar\omega_s)$ can be determined from the plot $(W_e - W_B)/W_B = f(\hbar\omega_s)$ (see the inset in Fig. 3). Here, $W_B = f(\hbar\omega_s)$ is the portion of the parabolic curve that is tangent to the spectrum (for example, curve *5* in Fig. 3).

In addition, a correlation between the dependence $S = f(\delta Y)$ and the spectrum modulation manifested itself, in particular, in the fact that the spectrum was smooth approximately within an interval of 0.01 mm in the vicinity of the values of δY corresponding to the appearance of extrema (local maxima and minima) in the dependence $S = f(\delta Y)$. The alteration of the smooth and modulated spectra as δY varied was apparently caused by the effect of the feedback. However, the emission reflected from the face and involved in the feedback returned to the active region immediately after the end of pumping. Therefore, it appears that we observed a spectrum modulation that was either appreciably enhanced or already established during the emission relaxation when the feedback was "switched on."

We can now expound the aforementioned statement that the dependence $W_e = f(\delta Y)$ at $\hbar\omega_s$ = const is modulated in a different way from the dependence $S = f(\delta Y)$ if a bulge in the spectrum $W_e = f(\hbar\omega_s)$ appears at this specific value of $\hbar\omega_s$.

We note that it is difficult to attribute the modulation of spectra simply to the formation of a cavity by the sample ends. The difference between the energies of the photons for neighboring resonant modes corresponding to the distance of 5 mm between the ends of this cavity should be equal to $\varepsilon \approx 0.034$ meV. This value of ε is several tens of times smaller than, for example, the intervals between local peaks in the curves shown in the inset in Fig. 3. In addition, the intervals between the peaks are not exactly identical. For example, in the inset in Fig. 3, the interval between the local peaks in curve *6* is larger than that in curve *7* by approximately a factor 1.5. Furthermore, it remains to be shown whether the selection of resonant modes with such a narrow intermode interval ε avoids contradicting the indeterminacy principle if the time of photon emission is restricted to the lifetime of the photogenerated charge carriers (~1 ps). It should be mentioned that we obtained the above estimate for the charge-carrier lifetime during an intense emission in an earlier study [7]; apparently, this lifetime corresponds to the formation of a "hole" in the light-amplification spectrum [10, 13]. In addition to the aforementioned difficulty with explaining the spectrum modulation by possible formation of a cavity between the sample ends, the transformation of a smooth spectrum into modulated spectrum and vice versa is also difficult to interpret. For example, as the optical path in a Fabry–Perot interferometer varies, the emission-spectrum modulation does not disappear; rather, only the spectral position of the local maxima changes. Thus, the change from a smooth spectrum to a modulated spectrum indicates once again that the emission under study is an element of a nonlinear system.

The observed modulation of the emission spectrum can be considered as the emission amplification in narrow local spectral regions. For example, this modulation of the spectrum could arise as a result of the time modulation of the emission intensity with the period

$$T \approx 2\pi/\Delta\omega_s, \qquad (2)$$

where $\Delta\omega_s$ is the interval between the local maxima in the emission spectrum. Expression (2) corresponds to a situation where the temporal pattern of the total field of several modes should exhibit a pronounced periodicity with a period of $2\pi/\Delta\omega_s$ if the frequency interval is

equal to $\Delta\omega_s$ [14]. According to the experimental spectra, $\hbar\Delta\omega_s \approx 1$ meV. Therefore, the modulation period determined from expression (2) is equal to ~4 ps; i.e., it coincides with the estimate for T obtained from expression (1). We can note that this estimate is close, for example, to the period of relaxation oscillations of the stimulated emission $2\pi(\tau_{st}\tau_p)^{1/2} \approx 6$ ps [15]) if we assume that the photon lifetime $\tau_p \approx 1$ ps, which appears to be realistic, and the charge-carrier lifetime $\tau_{st} \approx 1$ ps (called "stimulated" [15]). It is worth recalling that the relaxation oscillations are the interrelated oscillations of the emission intensity and those of the level populations, and that they originate as the pump intensity varies rapidly [5]. It also follows from the results of a numerical simulation [1] that picosecond pulsations of emission can be observed during the emission relaxation. We can preliminarily note another similarity to the relaxation oscillations, although this similarity is not very important. In our experiments, we observed a spectrum modulation that was enhanced or formed as a result of feedback; at the same time, the feedback is also required for the development of intense relaxation oscillations, as numerical simulation has shown [4].

It is remarkable that the dependence of the emission energy (with the photon energy fixed) on the delay τ between two pumping pulses was found to be modulated exactly with the periodicity of ~4 ps that was estimated from expressions (1) and (2). This effect was observed when a sample was irradiated with the *e* and *p* pulses that excited the charge carriers to different energy levels. The delay time τ between the pulses was varied within the picosecond range, causing the time distribution of the pump intensity and the total pump time to be varied as well.

In order to gain preliminary insight into the variation in the modulated emission spectrum resulting from only an increase in the pump intensity (without changes in the pump duration) by adding a *p* pulse, we show the spectra $W_e = f(\hbar\omega_s)$ and $W_{e+p} = f(\hbar\omega_s)$ in Fig. 4. The latter spectrum was measured at $\tau = 0$, and W_{e+p} is the emission energy integrated with respect to time for pumping with both the *e* and *p* pulses. A comparison of these spectra and a comparison of the corresponding plots $(W - W_B)/W_B = f(\hbar\omega_s)$ in the inset in Fig. 4 showed that the spectrum remained modulated in the case of synchronous pumping with the *e* and *p* pulses. The degree and the spectral shape of the modulation represented by the plot $(W_{e+p} - W_B)/W_B = f(\hbar\omega_s)$ differed only slightly from the case of the modulation caused by pumping with only an *e* pulse. Still, it is worth noting that, as a result of adding a *p* pulse, the emission energy increases nonuniformly over the spectrum, as was confirmed by the spectrum of relative variations in the emission energy $W_{e+p}/W_e = f(\hbar\omega_s)$ in Fig. 4. In this respect, the spectrum modulation was also found to be a nonlinear process. It should be noted that we observed this inhomogeneous variation in the energy over the spectrum in a previous experiment [16] and that it was an important factor in stimulating this study.

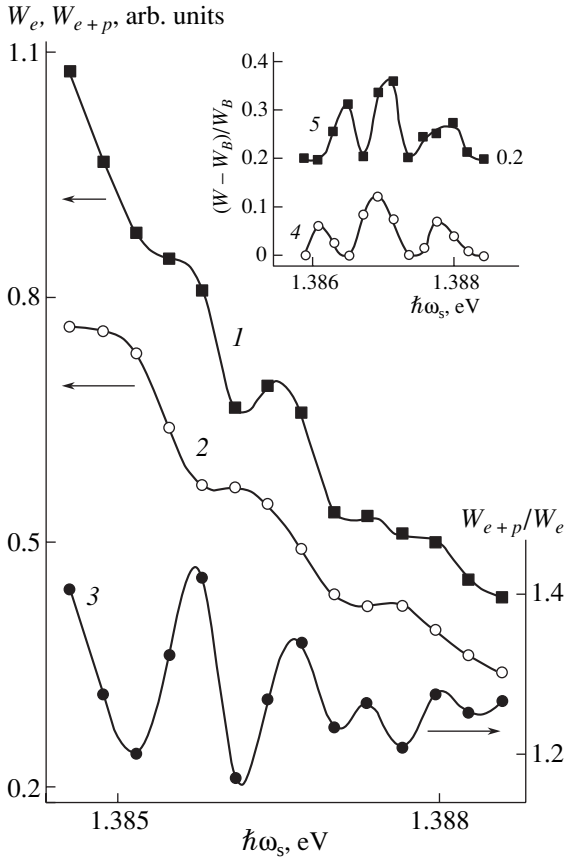

Fig. 4. Emission spectra (*1*) $W_{e+p} = f(\hbar\omega_s)$ at $\tau = 0$, (*2*) $W_e = f(\hbar\omega_s)$, and (*3*) $W_{e+p}/W_e = f(\hbar\omega_s)$ ($\delta Y = 0.26$ mm). The inset shows the spectra of relative modulation: (*4*) $(W_e - W_B)/W_B = f(\hbar\omega_s)$ in the case of pumping with a single pulse and (*5*) $(W_{e+p} - W_B)/W_B = f(\hbar\omega_s)$ in the case of pumping with two pulses and $\tau = 0$. Curve *5* is shifted upward by the value written at the curve.

It was later found that, even at an invariable total energy of the two-pulse pumping, a variation in the delay τ of the *p* pulse with respect to the *e* pulse within the picosecond range led to an appreciable variation in the modulation of the emission spectrum. This behavior can be illustrated by comparing the spectra $W_{e+p} = f(\hbar\omega_s)$ measured at $\tau = 0$ and $\tau = 2$ ps (Fig. 5). It can be seen that the emission energy changed to the greatest extent at the photon energies that corresponded to protuberances in the curve and, to much lesser extent, in the regions where there were crevasses. Figure 6 illustrates, in detail, the dependence of the emission energy on the delay τ at two photon energies corresponding to a protuberance (case 1) and a crevasse (case 2) in the spectra measured at $\tau = 0$ in additional experiments. It was found that the dependence of the emission energy on the delay time $W_{e+p} = f(\tau)$ was clearly modulated by local maxima in case 1 (curve *1*) and was nearly smooth

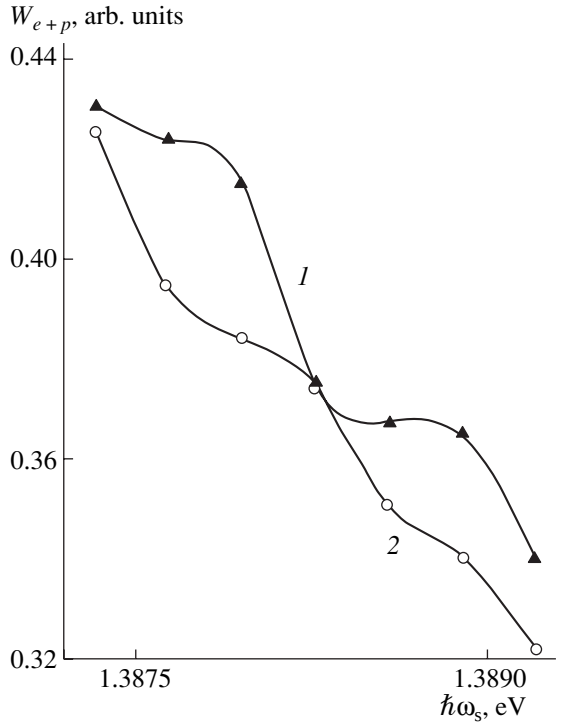

Fig. 5. Emission spectra measured at the delay times τ = (*1*) 0 and (*2*) 2 ps. δY = 0.19 mm.

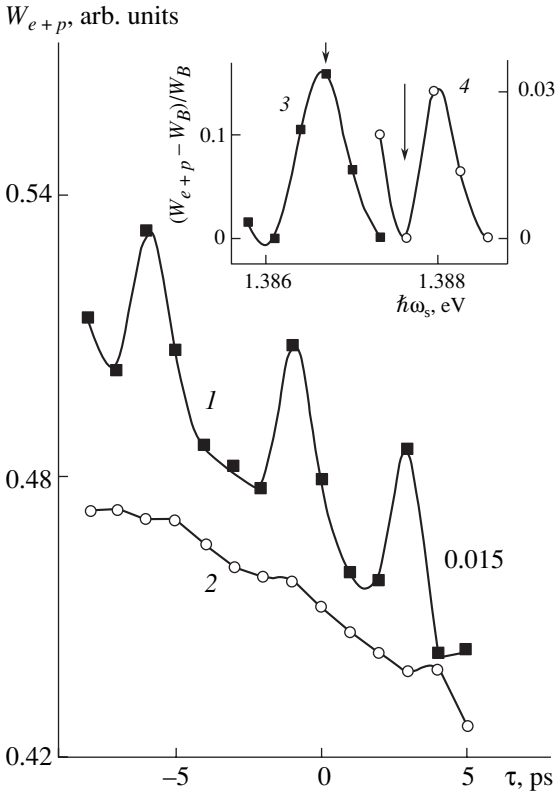

Fig. 6. Dependence of the emission energy W_{e+p} on the delay time τ at the photon energy $\hbar\omega_s$ = (*1*) 1.38671 eV (δY = 0.29 mm) and (*2*) 1.38765 eV (δY = 0.15 mm). Curve *1* is shifted upward with respect to its true position by the value written on the right of the curve. In the inset, we show fragments of the spectra $(W_{e+p} - W_B)/W_B = f(\hbar\omega_s)$ measured at (*3*) δY = 0.29 mm (the left-hand vertical axis) and (*4*) δY = 0.15 mm (the right-hand vertical axis). The arrows indicate the spectral points at which curves *1* and *2* were measured.

in case 2 (curve *2*). Specifically, the dependence $W_{e+p} = f(\tau)$ was modulated (smooth) for an emission with photon energy that corresponded to a local maximum (minimum) in the spectrum $(W_{e+p} - W_B)/W_B = f(\hbar\omega_s)$ measured, in the same experiment, for τ = 0 and illustrated by curve *3* (curve *4*) in the inset in Fig. 6. The interval between the local maxima in the dependence $W_{e+p} = f(\tau)$, represented by curve *1* in Fig. 6, was found to be equal to $\Delta\tau \approx 4$–5 ps. Specifically, the interval Δτ is approximately equal to the above estimates of the time modulation period for the emission; those estimates were based on the modulation of the dependence $S = f(\delta Y)$ and on the spectrum modulation:

$$\Delta\tau \approx T. \quad (3)$$

The modulation of the dependence $W_{e+p} = f(\tau)$ suggests that, once this modulation is observed (i.e., when the emission spectrum is modulated), the generation of the emission is pulsatory and features an interval between pulses equal to Δτ.

Thus, in this study, we observed the modulation of the characteristics of intense picosecond emission from GaAs pumped by ultrashort optical pulses. The modulation parameters were found to be interrelated by the expression

$$4\Delta Y/c_g \approx 2\pi/\Delta\omega_s \approx \Delta\tau \approx T. \quad (4)$$

This expression and qualitative variations in modulation described above suggest that the modulation of the characteristics is caused by a common (not yet identified) mechanism of self-modulation of the emission spectrum. This mechanism develops under the conditions of an ultrafast nonlinear interaction of highly photoexcited semiconductor with pumping light and stimulated recombination emission. Experiments show that an important role in the spectrum modulation and therefore in the origination of self-modulation mechanism is played by feedback via the emission reflected from the face and also the delay time τ in the case of pumping with two pulses. The experiments also provide good grounds for believing that, if the feedback is properly tuned, the effect of the self-modulation mechanism on the generation of stimulated emission is pulsatory and features an interval between pulsations *T* that is within the picosecond time region. Apparently, this mechanism should give rise to an amplitude modulation of the emission by picosecond pulsations, since the concept of such modulation corresponds to experimentally obtained relation (4).

This study was supported by the Russian Foundation for Basic Research, project no. 04-02-17146.

We are indebted to V.I. Perel', Yu.D. Kalafati, L.V. Levkin, Yu.K. Chmarovskiĭ, and G.N. Shkerdin for their participation in a detailed discussion of the results and their helpful comments.

REFERENCES

1. Yu. D. Kalafati, V. A. Kokin, H. M. Van Driel, and G. R. Allan, in *Hot Carriers in Semiconductors*, Ed. by K. Hess (Plenum, New York, 1996), p. 587.
2. L. W. Casperson, J. Appl. Phys. **48**, 256 (1977).
3. E. O. Goebel, O. Hildebrandt, and K. Lohnert, IEEE J. Quantum Electron. **13**, 848 (1977).
4. K. Y. Lau and A. Yariv, Appl. Phys. Lett. **40**, 452 (1982).
5. P. G. Eliseev, *Introduction to the Physics of Injection Lasers* (Nauka, Moscow, 1983), p. 294 [in Russian].
6. K. Lau and A. Yariv, in *Semiconductors and Semimetals*, Vol. 22: *Lightwave Communications Technology*, I: *Semiconductor Injection Lasers*, Ed. by W. T. Tsang (Academic, New York, 1985; Radio i Svyaz', Moscow, 1990).
7. N. N. Ageeva, I. L. Bronevoĭ, E. G. Dyadyushkin, *et al.*, Solid State Commun. **72**, 625 (1989).
8. Yu. D. Kalafati and V. A. Kokin, Zh. Éksp. Teor. Fiz. **99**, 1793 (1991) [Sov. Phys. JETP **72**, 1003 (1991)].
9. I. L. Bronevoĭ and A. N. Krivonosov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 13 (1999) [Semiconductors **33**, 10 (1999)].
10. N. N. Ageeva, I. L. Bronevoĭ, A. N. Krivonosov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 144 (2002) [Semiconductors **36**, 136 (2002)].
11. D. T. F. Marple, J. Appl. Phys. **35**, 1241 (1964).
12. A. Rivlin, *Dynamics of Emission from Semiconductor Lasers* (Sovetskoe Radio, Moscow, 1976), p. 175 [in Russian].
13. G. S. Altybaev, I. L. Bronevoĭ, and S. E. Kumekov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **38**, 674 (2004) [Semiconductors **38**, 648 (2004)].
14. D. Bradley, in *Ultrashort Light Pulses*, Ed. by S. Shapiro (Springer, Heidelberg, 1977; Mir, Moscow, 1980).
15. K. Lau, J. Lightwave Technol. **7**, 400 (1989).
16. N. N. Ageeva, I. L. Bronevoĭ, A. N. Krivonosov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **37**, 546 (2003) [Semiconductors **37**, 526 (2003)].

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 658–660. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 689–691.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Recombination Mechanism of the Piezophotoresistive Effect in Compensated Semiconductors

B. M. Pavlyshenko^ and R. Ya. Shuvar

Franko National University (Physics Department), ul. Kirila i Mefodiya 6, Lviv, 79005 Ukraine

^e-mail: pavlsh@yahoo.com

Submitted November 2, 2004; accepted for publication November 15, 2004

Abstract—Numerical simulation methods are used to study the recombination mechanism involved in the piezophotoresistive effect, taking into account the role of compensating impurities. It is shown that the amplitudes of the alternating concentrations of free carriers induced by dynamic deformation of a semiconductor can be increased using steady-state photoexcitation. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The piezophotoresistive effect (PPRE) arises under the combined influence of alternating deformation and steady-state photoexcitation in a semiconductor. This effect consists in an increase in an alternating conductivity component induced by dynamic deformation in a semiconductor as a result of steady-state photoexcitation [1, 2]. One of the possible mechanisms involved in the appearance of the PPRE may be a recombination mechanism caused by the influence of deformation on the recombination rates of nonequilibrium photogenerated charge carriers [3]. Recently, a great deal of attention has been paid to particular cases of the Shockley–Read–Hall recombination theory [4–7]. In the context of this theory, Karazhanov [4] studied the influence of compensating impurities on the photoelectric properties of semiconductors. It was found [4] that the resistivity, lifetime of charge carriers, and photoconductivity in a semiconductor increase drastically as the concentration of deep-level impurities increases.

In this paper, we report the results obtained from a study of the recombination mechanism of the PPRE using theoretical methods and numerical simulation and the influence of compensating impurities on this mechanism.

2. THEORETICAL MODEL

For the purposes theoretical and numerical analysis, we choose a model semiconductor with a simple energy-band structure, a single deep level of acceptor recombination centers in its band gap, and a single shallow level of donor impurities. The effect of deformation is taken into account phenomenologically by considering the modulation of the band edges for the conduction (E_c) and valence (E_v) bands:

$$E_{c,v}(t) = E_{c,v} + \Delta E_{c,v}\cos\omega t. \quad (1)$$

Here, ω is the frequency of alternating deformation and $\Delta E_{c,v}$ are the amplitudes of the alternating deformation of the band edges. Photoexcitation is taken into account by considering the optical-generation rate for nonequilibrium charge carriers.

In the context of the Shockley–Read–Hall recombination theory [8], the continuity equations for free and localized charge carriers are written as

$$\begin{aligned}\frac{dn}{dt} &= G + C_{nr}N_r n_1 - C_{nr}P_r n,\\ \frac{dp}{dt} &= G + C_{pr}P_r p_1 - C_{pr}N_r p,\end{aligned} \quad (2)$$

where

$$n_1 = N_c\exp\left[-\frac{E_c - E_r}{kT}\right];\quad p_1 = N_v\exp\left[-\frac{E_r - E_v}{kT}\right];$$
$$N_r + P_r = M_r;$$

n and p are the concentrations of free electrons and holes, respectively; C_{nr} and C_{pr} are the coefficients of charge-carrier capture by the recombination level; N_r and P_r are the concentrations of electrons and holes localized at the recombination centers; M_r is the concentration of recombination centers; G is the rate of optical generation of nonequilibrium charge carriers; E_r is the energy position of the recombination level; N_c and N_v are the effective densities of states in the conduction and valence bands, respectively; T is temperature; and k is the Boltzmann constant. Equations (2) are supplemented with an electroneutrality equation. We analyzed these equations for the first harmonic using the complex-amplitude method in the region of steady-state photoexcitation in a linear approximation with respect to the strain tensor. As a result, we obtained the

following equations for the complex amplitudes of alternating nonequilibrium-carrier concentrations:

$$i\omega\Delta n = -\frac{\Delta n}{\tau_n} + \pi_n,$$
$$i\omega\Delta p = -\frac{\Delta n}{\tau_p} + \pi_p. \quad (3)$$

Here,

$$\pi_n = \frac{N_r(N_r + p + p_1 + i\omega C_{pr}^{-1})\Delta n_1 + P_r(n + n_1)\Delta p_1}{C_{pr}^{-1}N_r + C_{pr}^{-1}(n + n_1) + C_{nr}^{-1}(p + p_1) + (C_{nr}C_{pr})^{-1}i\omega},$$

$$\pi_p = \frac{P_r(P_r + n + n_1 + i\omega C_{nr}^{-1})\Delta p_1 + N_r(p + p_1)\Delta n_1}{C_{pr}^{-1}N_r + C_{pr}^{-1}(n + n_1) + C_{nr}^{-1}(p + p_1) + (C_{nr}C_{pr})^{-1}i\omega},$$

$$\Delta n_1 = -\frac{\Delta E_c - \Delta E_r}{kT}, \quad \Delta p_1 = -\frac{\Delta E_r - \Delta E_v}{kT}p_1.$$

In the analysis that follows, we consider the quasi-steady frequency range in which the conditions $\omega \ll \tau_n^{-1}$ and $\omega \ll \tau_p^{-1}$ are satisfied. As the photoexcitation becomes more intense, the occupancy of the recombination level attains a limiting steady state and the amplitudes Δn and Δp tend asymptotically to the following values:

$$(\Delta n)_{\text{sat}} = (\Delta p)_{\text{sat}} = \frac{C_{nr}\Delta n_1 + C_{pr}\Delta p_1}{C_{nr} + C_{pr}}. \quad (4)$$

We analyzed the dependences $\Delta n(G)$ (the concentration–excitation characteristics) using a numerical simulation. We determined the steady-state concentrations of free and localized charge carriers, which appear in Eq. (3) as parameters, by numerically solving system of Eqs. (2) for the steady-state case. In these numerical calculations, we used a model of a semiconductor with the following realistic parameters: a band gap $E_g = 0.8$ eV, a recombination-level energy $E_r = 0.4$ eV, a concentration of recombination centers $M_r = 10^{16}$ cm^{-3}, coefficients of the charge-carrier capture by recombination centers $C_{nr} = 10^{-12}$ cm^3 s^{-1} and $C_{pr} = 10^{-11}$ cm^3 s^{-1}, a depth of the donor-impurity level $E_d = 0.1$ eV, an amplitude of the strain-induced variation in the energy-band edges $\Delta E = 10^{-5}$ eV, and a semiconductor temperature $T = 300$ K. In order to provide certainty, we report the results for the electron component of the conductivity; however, the general features of the dependences remain the same for the hole component as well.

3. RESULTS AND DISCUSSION

In Fig. 1, we show the calculated concentration–excitation characteristics of the piezophotoresistive effect for the semiconductor model under consideration at different concentrations of a shallow-level impurity N_d. Let us consider the mechanism of an increase in the alternating excess-carrier concentration Δn resulting from steady-state photoexcitation (these carriers appear under the effect of alternating strain) (curves *1*, *2*). As the photoexcitation intensity increases, the concentration of electrons localized at the recombination level changes. In Eqs. (2), the effect of strain is taken into account by the factors n_1 and p_1, which appear in the term related to the rate of thermal generation. As a result, the rate at which the charge carriers localized at the recombination centers are thermally excited towards the bands is modulated. As the concentration of localized charge carriers increases under the effect of steady-state photoexcitation, the thermal-generation rate described by the term $C_{nr}N_r n_1$ in Eqs. (2) becomes higher; as a consequence, the complex amplitudes of the alternating concentrations increase. At high photoexcitation levels, the amplitude Δn levels off as a result of the saturation of the recombination-level occupancy. Curve *3* in Fig. 2 shows a situation where the concentration of the compensating donor impurity equals that of the acceptor-type recombination centers, and it has a radically different shape. In this case, all the electrons from the donor centers are localized at the recombination centers in accordance with the electroneutrality condition; in addition, the thermal-generation rate, which is proportional to the concentration of localized electrons, attains a maximum. In the region of high photoexcitation intensities, as the steady-state concentrations of free nonequilibrium charge carriers increases, a certain depletion of the recombination centers and withdrawal of the electrons localized at these centers are observed; as a result, the complex amplitude decreases. In Fig. 2, we show the dependences of the

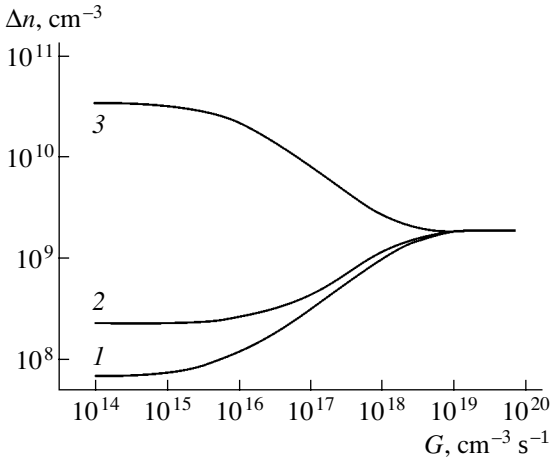

Fig. 1. Dependences of the amplitude of the alternating excess-carrier concentration (Δn) on the intensity of the nonequilibrium-carrier optical generation (G) at the following concentrations of compensating donor impurities N_d: (*1*) 10^{14}, (*2*) 10^{15}, and (*3*) 10^{16} cm^{-3}.

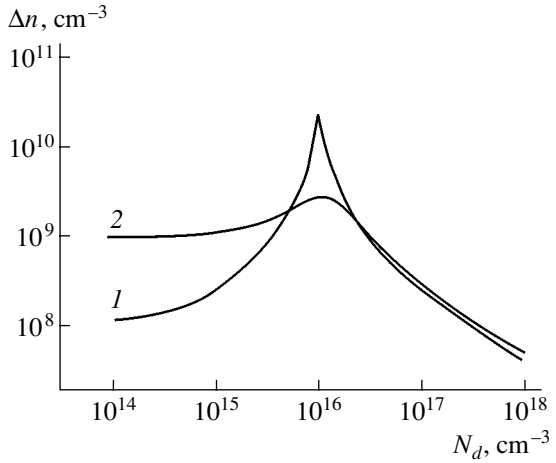

Fig. 2. Dependences of the alternating excess-carrier concentration (Δn) on the concentration of donor impurities (N_d) at the following intensities of the excess-carrier optical generation G: (*1*) 10^{16} and (*2*) 10^{18} cm^{-3} s^{-1}.

complex amplitude on the concentration of the compensating donor impurity for two photoexcitation intensities. In the curves of this figure, we can observe a single peak in the region of complete compensation. Here, the donor-impurity concentration is equal to the concentration of the acceptor-type recombination centers. The character of the curves for different photoexcitation intensities is almost identical in the region of the donor concentrations that exceed the concentration of recombination centers. In this region, photoexcitation only slightly affects the redistribution of the electrons localized at the recombination centers.

4. CONCLUSION

It is shown that, in the context of the Shockley–Read–Hall recombination model, one of the originating mechanisms of the piezophotoresistive effect can be related to the influence of alternating strain on the rate at which the charge carriers localized at the recombination centers are thermally excited towards the bands as a result of a strain-induced modulation of the energy positions of the band edges. The effect of an increase in the alternating-concentration amplitude as a result of photoexcitation is caused by the influence of steady-state photoexcitation on the occupancy of the recombination level. If the concentrations of the impurities compensate each other exactly, the alternating-concentration amplitude becomes virtually independent of the intensity of the steady-state photoexcitation, due to the complete occupation of the recombination level with charge carriers from the compensating impurity centers.

REFERENCES

1. G. D. Guseĭnov and G. B. Abdulaev, Dokl. Akad. Nauk SSSR **208**, 1052 (1973) [Sov. Phys. Dokl. **18**, 127 (1973)].
2. A. Mahapatra, P. G. Kornreich, and S. T. Kowel, Phys. Rev. B **18**, 2766 (1978).
3. Ĭ. M. Stakhira, R. Ya. Shuvar, and B. M. Pavlishenko, Ukr. Fiz. Zh. **40**, 723 (1995).
4. S. Zh. Karazhanov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 909 (2000) [Semiconductors **34**, 872 (2000)].
5. I. N. Volovichev and Yu. G. Gurevich, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 321 (2001) [Semiconductors **35**, 306 (2001)].
6. S. V. Kuznetsov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 1244 (2001) [Semiconductors **35**, 1191 (2001)].
7. A. P. Odrinskiĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **38**, 310 (2004) [Semiconductors **38**, 298 (2004)].
8. V. E. Lashkarev, A. V. Lyubchenko, and M. K. Sheĭnkman, *Nonequilibrium Processes in Photosemiconductors* (Naukova Dumka, Kiev, 1981) [in Russian].

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 661–665. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 692–696.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Spontaneous and Stimulated UV Luminescence of ZnO:N at 77 K

A. N. Georgobiani*, A. N. Gruzintsev^, E. E. Yakimov**, C. Barthou***, and P. Benalloul*****

**Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia*

***Institute of Microelectronics Technology, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

^e-mail: gran@ipmt-hpm.ac.ru

****Université P. et M. Curie, 75252 Paris Cedex 05, France*

Submitted November 10, 2004; accepted for publication November 25, 2004

Abstract—The spontaneous and stimulated UV luminescence spectra of ZnO:N samples possessing different nitrogen contents are measured at a temperature of 77 K. Luminescence peaks resulting from bound and free excitons, biexcitons, and electron recombination via the nitrogen acceptor level are identified. The optical depth (123 meV) of the N_O impurity acceptor level is determined. It is established that stimulated UV luminescence originating from inelastic exciton interaction sets in as the optical pump power increases. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Zinc oxide is a wide-gap (E_g = 3.37 eV) semiconductor that has the potential for application in the development of semiconductor lasers and LEDs operating in the UV spectral region. Its high exciton-binding energy (60 meV) permits the attainment of intense UV luminescence, which is derived from radiative exciton recombination occurring at room temperature and above (up to 550 K) [1]. The efficiency of this luminescence is so high that it is possible to obtain lasing under intense optical pumping. Depending on the degree of structural perfection and the composition of native point defects in undoped ZnO, its room-temperature luminescence spectrum is dominated either by exciton (peaking at 3.246 eV), green (with a peak at 2.43 eV), or red (with a peak at 1.91 eV) emission [1, 2]. Undoped zinc oxide features *n*-type conductivity due to the predominance of donor-type Zn_i and V_O defects, which originate from a zinc excess in the lattice. The electron conduction can be substantially enhanced by doping ZnO with gallium and aluminum donor impurities. In 2001, doping ZnO films with nitrogen was found to produce *p*-type conductivity at room temperature [3]. Further use of various acceptor impurities (As, N, P, Cu, and Ag) revealed that only the nitrogen impurity can induce *p*-type conductivity in the material, although the resulting resistivity is too low to permit its use in the development of high-efficiency optoelectronic devices. We obtained UV lasing under the optical pumping of polycrystalline nonepitaxial ZnO films deposited by magnetron sputtering on oxidized silicon substrates [4]. Stimulated luminescence was observed at room temperature in the region of electron–hole plasma recombination at 3.12 eV. As the films were polycrystalline, the threshold pump power provided by a pulsed nitrogen laser was relatively high (32 MW/cm^2). An implantation of nitrogen acceptors into these films and their subsequent annealing in the presence of oxygen radicals was observed to produce *p*-type conductivity [5]. Since, however, *p*-type conductivity in ZnO can be attained only by annealing below a certain critical temperature, this procedure had to be performed at low temperatures. The underlying reason is that, when annealed below the critical temperature (T_c = 550°C), the ZnO lattice predominantly loses zinc, while, at annealing temperatures above T_c, it is the oxygen loss that starts to prevail. As a result, donor-type V_O and Zn_i intrinsic defects are formed in the material, and these defects compensate the *p*-type conductivity caused by the nitrogen acceptor impurity. Therefore, to attain *p*-type conductivity, the implanted ZnO:N samples were annealed at temperatures below 600°C [5]. As a consequence, not all of the radiation defects were annealed, and the luminescence of the implanted films was weak. Therefore, it was impossible to confidently identify the spectral features related to the nitrogen impurity in zinc oxide and separate them from the manifestation of the radiation defects. More accurate experimental data on the luminescence lines associated with specific extrinsic defects in zinc oxide could, however, be obtained by using a very sensitive and nondestructive photoluminescence (PL) method to analyze its point defects.

In this study, we consider the effect of nitrogen impurity on the edge and impurity luminescence of zinc oxide, as well as on the generation of stimulated emission under pulsed optical excitation at different power levels. We also identify the mechanism of stimulated emission in ZnO and study its variation when this material is annealed at different temperatures in air.

2. EXPERIMENTAL

We studied zinc oxide powders prepared by pyrolysis from an aqueous solution of zinc nitrate with a concentration $Zn(NO_3)_2 \cdot 6H_2O$. The precipitate thus

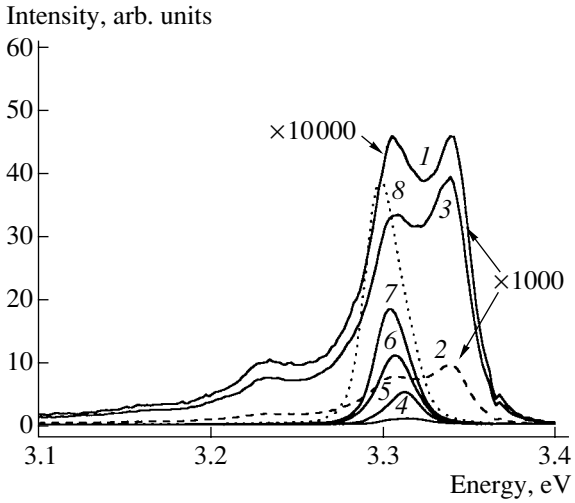

Fig. 1. Photoluminescence spectra of the initial ZnO:N samples. The excitation intensity was (*1*) 50, (*2*) 129, (*3*) 416, (*4*) 1820, (*5*) 5870, (*6*) 15500, (*7*) 23400, and (*8*) 50000 kW/cm^2. The measurement temperature T = 77 K.

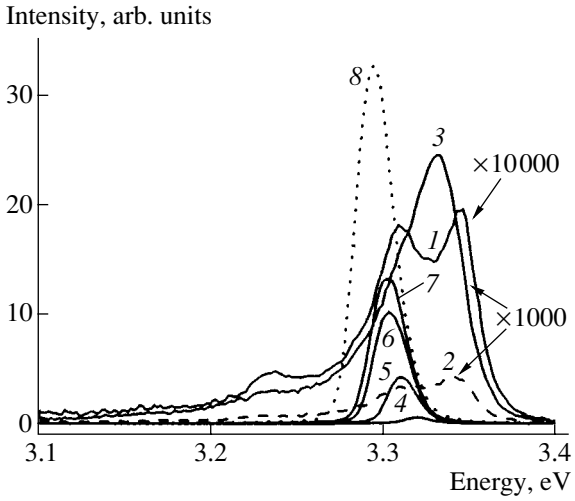

Fig. 2. Photoluminescence spectra of the ZnO:N samples annealed at 700°C. The excitation intensity was (*1*) 50, (*2*) 129, (*3*) 416, (*4*) 1820, (*5*) 5870, (*6*) 15500, (*7*) 23400, and (*8*) 50000 kW/cm^2. The measurement temperature T = 77 K.

obtained was heated in an oxygen ambient at 195°C, which is above the decomposition temperature of zinc nitrate, for three hours. Next, the initial ZnO samples were annealed in air for one hour at temperatures ranging from 400 to 1000°C (in steps of 100°C). The initial and annealed samples were analyzed using X-ray diffraction (crystallinity and phase composition), laser mass spectrometry (impurity composition), and scanning electron microscopy (morphology), as well as according to their photoluminescence (stoichiometry). The concentrations of more than one hundred analyzed impurities were found to be below 0.0001 wt %, with the exception of nitrogen, which was present in the initial samples at a concentration of 0.01 wt %. It is noteworthy that the nitrogen content in ZnO decreased under the annealing in air to 0.008 and 0.001 wt % at temperatures of 700 and 1000°C, respectively. The results of X-ray diffraction measurements of similar powders carried out at different annealing temperatures are described in [6]. It is also worth noting that, at high annealing temperatures, a narrowing of the diffraction peaks is observed to occur. This effect can be attributed to the single-crystal grains in the polycrystalline powder growing in size. This growth was also observed in the morphology studies we conducted with a JEOL-2000 scanning electron microscope.

The experiments with optically pumped luminescence were conducted using a pulsed nitrogen laser operating at a wavelength of 337.1 nm, pulse duration of 0.6 ns, and output power of 2.3 MW. The laser beam was focused to a rectangular spot 1×3 mm^2 in size, thus offering pump power densities of up to 50 MW/cm^2. However, due to the short pulse duration and low pulse-repetition frequency, no overheating or degradation of the samples was observed. The PL spectra were measured at temperatures ranging from 77 to 550 K at different excitation-intensity levels. The luminescence signal was registered using an optical waveguide (with a 0.1 rad aperture) and backscattering geometry in the direction normal to the sample surface. The luminescence was analyzed using a monochromator equipped with a CCD array that had a spectral resolution of no poorer than 0.1 nm/pixel.

3. RESULTS AND DISCUSSION

Figure 1 shows the PL spectra of the initial ZnO:N (0.01 wt %) samples measured at a liquid-nitrogen temperature for different pump power levels. At low excitation intensities, the edge luminescence curves consist of several elementary bands peaked at 3.37 eV (band *EX* related to the recombination of free excitons), 3.34 eV (band *AX* related to the recombination of acceptor-bound excitons), 3.307 eV (band *EA* related to the recombination of conduction band electrons with their transition to the acceptor level), and 3.233 eV (LO phonon replica of the *EA* 3.307-eV band) [7]. An increase in the pump power in the above-mentioned power range brings about only a slight increase in the bound-exciton band contribution (compare curves *1* and *3* in Fig. 1). Starting with pump power levels of 1820 kW/cm^2 (curve *4*), however, the PL spectrum undergoes a radical change. The edge luminescence combines to form one narrow *P* band, whose peak shifts gradually from 3.314 to 3.297 eV as the pump power increases (curves *4–8*).

Annealing the ZnO:N samples in air at 700°C for one hour substantially affects the features of the edge luminescence at low excitation intensities (Fig. 2, curves *1–3*). Note the decrease in the *EA* band contribution compared to the bound exciton band at the lowest power density of 50 kW/cm^2 (curve *1*). In addition, the

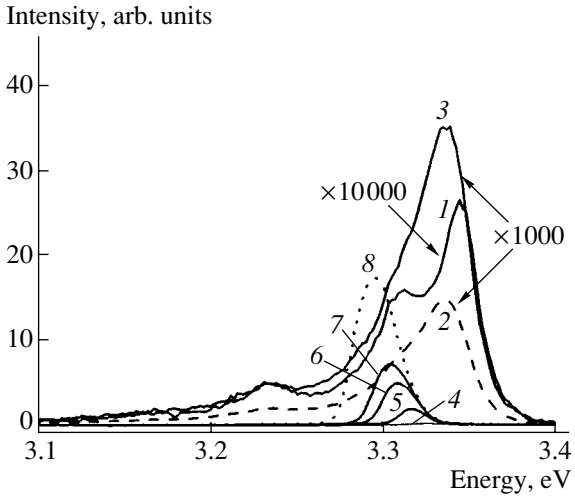

Fig. 3. Photoluminescence spectra of the ZnO:N samples annealed at 1000°C. The excitation intensity was (*1*) 50, (*2*) 129, (*3*) 416, (*4*) 1820, (*5*) 5870, (*6*) 15500, (*7*) 23400, and (*8*) 50000 kW/cm^2. The measurement temperature T = 77 K.

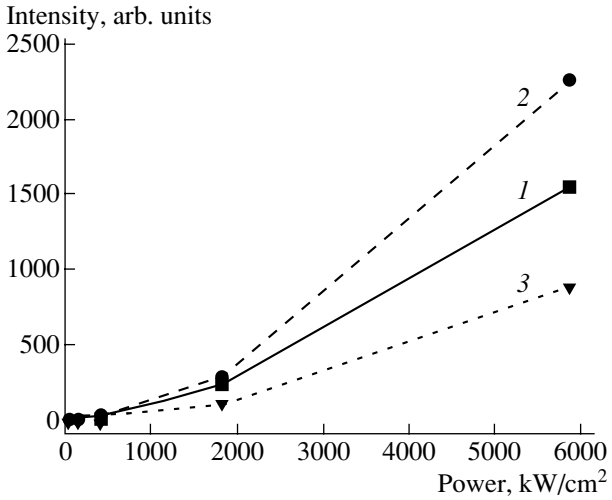

Fig. 4. Intensity of the prevailing edge-luminescence band plotted vs pump laser power density for ZnO:N samples (*1*) unannealed, (*2*) annealed at 700°C, and (*3*) annealed at 1000°C. T = 77 K.

transformation of several bands into one broad *M* band peaked at 3.333 eV sets in at a pump power density of 416 kW/cm^2 (Fig. 2, curve *3*). An increase in the temperature of the annealing in air to 1000°C further reduces the relative contribution of the *EA* band at the lowest pump intensity (Fig. 3, curve *1*). Regarding the *M* band, it appears, in this case, at a power density of 129 kW/cm^2 (curve *2*). The narrow *P* band becomes dominant in the spectra of all the samples at pump power levels above 1820 kW/cm^2 (Figs. 2, 3, curves *4–8*).

This pump power density (1820 kW/cm^2) is a threshold in the transition from a spontaneous to a stimulated mechanism of UV emission in ZnO:N. A further increase in the pump power brings about not only a sharp narrowing of the luminescence spectrum but an appreciable nonlinear increase in the luminescence intensity (Fig. 4).

We studied the effect of sample temperature on the stimulated UV luminescence of zinc oxide. Figure 5 shows the PL spectra of the initial ZnO:N samples when measured at temperatures ranging from 77 to 550 K under high-intensity nitrogen laser pumping (23400 kW/cm^2). The narrow intense *P* line of stimulated luminescence can clearly be seen to prevail until about 300 K (curves *1–5*). A further increase of temperature results in the appearance of a broad violet band of spontaneous emission (curves *6–10*). The luminescence peak shifts to longer wavelengths, from 3.3 eV at 77 K to 3.03 eV for 550 K, as the temperature increases. In Fig. 5, in order to gain more information from the luminescence curves, they are plotted on an arbitrary scale. For clarificaton of the extent to which the stimulated UV luminescence is more efficient than the spontaneous one, Fig. 6 shows the areas under the luminescence curve and the FWHM of this curve for the initial

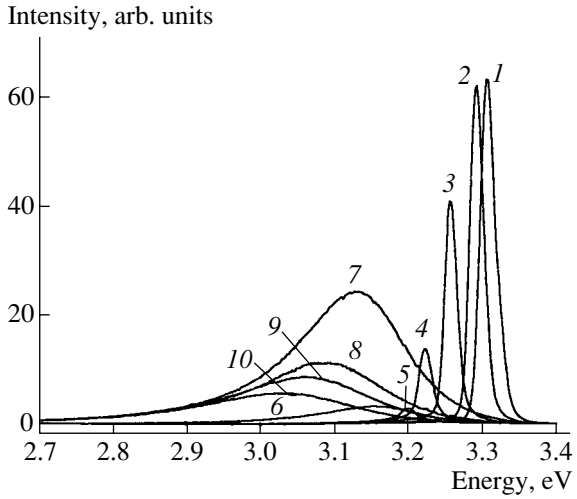

Fig. 5. Photoluminescence spectra of the initial ZnO:N samples obtained at the measurement temperatures T = (*1*) 77, (*2*) 130, (*3*) 200, (*4*) 260, (*5*) 300, (*6*) 370, (*7*) 400, (*8*) 460, (*9*) 500, and (*10*) 550 K. The pump power density was 23400 kW/cm^2.

samples in relation to the temperature. The small line half-width (3 nm) remains practically unchanged until 330 K. At higher temperatures, the half-width increases sharply, by an order of magnitude, to 30 nm. Note that the luminescence efficiency (the area under the curve) decreases by more than two orders of magnitude. It should be pointed out that the temperature behavior of the stimulated luminescence of the ZnO:N samples at high pumping levels depended little on the annealing temperature and followed, for all the samples, a pattern similar to that of the curves in Fig. 6. Furthermore, the stimulated luminescence of all the samples measured at

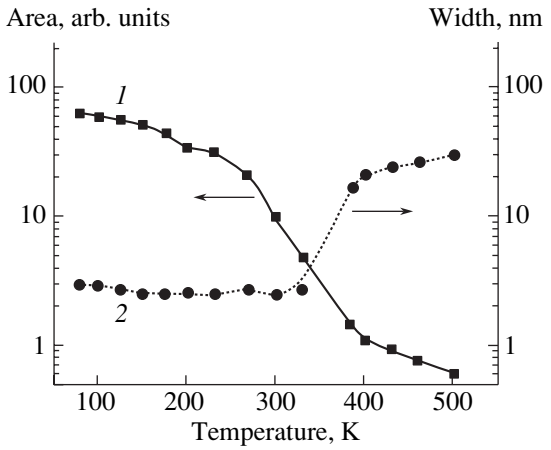

Fig. 6. Area under the luminescence band (*1*) and FWHM of the band (*2*) for the initial ZnO:N samples in relation to the measurement temperature. The pump power density was 23400 kW/cm^2.

77 K originally had the same pattern of a narrow intense *P* band.

We mentioned above, however, that the spontaneous luminescence generated at low pumping levels (less than 1820 kW/cm^2) depends heavily on the temperature at which the samples were preannealed (Figs. 1–3). We note, first of all, the decrease in the *EA* band intensity at 3.307 eV (the recombination of electrons from the conduction band with their transition to an acceptor level) as the annealing temperature increases. Recall the results of the impurity content analysis indicating a decrease in the nitrogen content in the samples from 0.01 to 0.008 and 0.001% for annealing temperatures of 700 and 1000°C, respectively. Taking these facts into consideration, the acceptor level in the electron recombination is most likely related to the N_O substitutional defects. If we know the band gap of zinc oxide at a liquid-nitrogen temperature (E_g = 3.43 eV), we can find the optical depth of the nitrogen acceptor level from the difference between these two energies (E_g and the energy of the *EA* band): $E_N = E_g - E_{EA}$ = 3.43 – 3.307 eV = 0.123 eV.

The second essential feature of the spontaneous luminescence of the annealed ZnO:N samples consists in the appearance of a fairly broad *M* band peaked at 3.333 eV, which becomes dominant as the pump power increases (Figs. 2, 3). A similar band has been observed in single-crystal, high-quality epitaxial films of zinc oxide grown by MBE on gallium nitride substrates [8]. It was found that the integrated intensity identified with the area bounded by the *M* band grew superlinearly as the pump intensity increased. This provided grounds for the identification of this band with recombination from the bound biexciton state [8]. The different pump power densities needed for biexciton formation in the samples studied is due, in our case, to differences in the structural perfection of these samples. Obviously, an increase in the annealing temperature to 1000°C gives rise to a substantial growth in grain size and facilitates the formation of biexciton states. This conclusion is consistent with the X-ray diffraction measurements, which showed that the lines become narrower as the annealing temperature increases.

The exciton kinetics in zinc oxide at 77 K can be described according to the following pattern as the optical pump power increases. At low pump powers, the number of forming excitons is small, and they have time to become localized at donor or acceptor defects before radiative recombination. As the pump power increases, the number of excitons becomes larger than that of impurity and intrinsic point defects in the material. This brings about the binding of excitons to form biexcitons, which are responsible for the *M* luminescence band in the high-quality single-crystal ZnO samples. Note that the pump power level at which the luminescence of defect-bound excitons converts to that of biexcitons depends on the actual defect density in the sample (in our case, on the annealing temperature).

As the pump power increases still further (above 1820 kW/cm^2), the kinetic energy of some of the excitons exceeds the biexciton binding energy. Therefore, a *P* band related to inelastic collisions among free excitons appears instead of the biexciton luminescence [9]. One of the colliding excitons takes away part of the energy of the other and transfers to an excited state with a quantum number $n > 1$, while the other exciton recombines, with the resulting emission of a photon. The energy of this photon can be written as [10]

$$E_n = E_{ex} - E_b^{ex}(1 - 1/n^2) - (3/2)kT, \qquad (1)$$

where n = 2, 3, 4, …, ∞; E_{ex} is the free-exciton recombination energy (3.37 eV); E_b^{ex} = 60 meV is the exciton binding energy in ZnO [11]; and kT is the thermal energy. At T =77 K, Eq. (1) yields 3.314 eV for the energy position of the E_2 peak, and 3.299 eV, for that of the peak at E_∞. As can be seen, the calculated values are in good agreement with the position of the narrow *P* band (Figs. 1–3), whose peak shifts smoothly from 3.314 to 3.297 eV as the pump power increases (curves *4–8*).

4. CONCLUSION

Thus, the stimulated UV radiation that we observed in ZnO:N samples at high optical pump powers can be accounted for by inelastic scattering and the recombination of two interacting excitons. As the pump power increases, the remaining exciton transfers to ever higher excited states and the *P* band shifts to longer wavelengths. It is the quadratic dependence of this process on the exciton density or pump intensity of zinc oxide that accounts for its dominance at high pump-power densities. This stimulated emission prevailed until temperatures became higher than 295 K (360 K) (Figs. 5, 6).

In an earlier study [4], we observed stimulated luminescence of slightly crystalline ZnO films generated by a direct recombination of electrons and holes in the electron–hole plasma of heavily pumped (above 32 MW/cm^2) zinc oxide. The poor crystallinity of the material in this case prevented us from reaching the high exciton concentration needed for the *P* band to form. Hence, a good crystal structure in this material is important for the observation of exciton-based stimulated luminescence at liquid-nitrogen and room temperatures.

The decrease in the crystal defect concentration in the ZnO:N samples under study as a result of high-temperature annealing made it possible to observe the *M* band caused by biexciton radiative recombination at $T = 77$ K. This band prevailed at moderate optical pump power levels. At the same time, low excitation intensity favored the appearance of lines related to free and point-defect-bound excitons in the edge-luminescence region, as well as of the *EA* band related to the recombination transitions of conduction-band electrons to the acceptor level of the nitrogen impurity. The optical depth of the nitrogen acceptor level measured from the valence-band top was found to be $E_N = 0.123$ eV.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 04-02-16437 and 04-02-16942), a program of the Russian Academy of Sciences "Quantum calculations," and INTAS (project no. 2002-0796).

REFERENCES

1. D. M. Bagnal, Y. F. Chen, Z. Zhu, and T. Yao, Appl. Phys. Lett. **73**, 1038 (1998).
2. D. M. Bagnal, Y. F. Chen, M. Y. Shen, *et al.*, J. Cryst. Growth **184–185**, 605 (1998).
3. M. Joseph, H. Tanaba, H. Soeki, *et al.*, Physica B (Amsterdam) **302–303**, 140 (2001).
4. A. N. Gruzintsev, V. T. Volkov, C. Barthou, *et al.*, Thin Solid Films **459**, 262 (2004).
5. A. N. Georgobiani, A. N. Gruzintsev, V. T. Volkov, *et al.*, Nucl. Instrum. Methods Phys. Res. A **514**, 117 (2003).
6. S. A. Studenikin, N. Golego, and M. Cocivera, J. Appl. Phys. **84**, 2287 (1998).
7. S. A. Studenikin, M. Cocivera, W. Kellner, and H. Pascher, J. Lumin. **91**, 223 (2000).
8. H. J. Ko, Y. F. Chen, T. Yao, *et al.*, Appl. Phys. Lett. **77**, 537 (2000).
9. P. Zu, Z. K. Tang, G. K. Wong, *et al.*, Solid State Commun. **103**, 459 (1997).
10. C. Klingshirn and H. Haug, Phys. Rep. **70**, 315 (1981).
11. W. Y. Liang and A. D. Yoffe, Phys. Rev. Lett. **20**, 59 (1968).

Translated by G. Skrebtsov

Semiconductors, Vol. 39, No. 6, 2005, pp. 666–673. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 697–704.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Increase in the Rate and Discretization of the Kinetics of Isothermal Surface Generation of Minority Charge Carriers in Metal–Insulator–Semiconductor Structures with a Planar-Inhomogeneous Insulator

A. G. Zhdan, E. I. Goldman, Yu. V. Gulyaev, and G. V. Chucheva^

Institute of Radio Engineering and Electronics, Russian Academy of Sciences (Fryazino Branch), pl. Vvedenskogo 1, Fryazino, Moscow oblast, 141190 Russia

^e-mail: gvc@ms.ire.rssi.ru

Submitted August 31, 2004; accepted for publication October 7, 2004

Abstract—Surface generation of minority charge carriers in silicon metal–oxide–semiconductor (MOS) structures is efficient only at the initial recombinationless stage. Quasi-equilibrium between surface generation centers and the minority-carrier band is established in a time $t \sim 10^{-5}$ s. In the absence of other carrier generation channels, an equilibrium inversion state at 300 K would need $t = t_\infty > 10^3$ years to become established. In fact, the time t_∞ is much shorter, due to excess-carrier generation via centers located at the SiO_2/Si interface over the gate periphery. This edge-related generation can easily be simulated in an MOS structure with a single gate insulated from Si by oxide layers of various thicknesses. At gate depleting voltages V_g, the role of the periphery is played by a shallow potential well under a thicker oxide, and the current-generation kinetics becomes unconventional: two discrete steps are observed in the dependences $I(t)$, and the duration and height of these steps depend on V_g. An analysis of the $I(t)$ curves allows determination of the electric characteristics of the Si surface in the states of initial depletion ($t = 0$) and equilibrium inversion ($t = t_\infty$), as well as the parameters of surface lag centers, including their energy and spatial distributions. The functionally specialized planar inhomogeneity of a gate insulator is a promising basis for dynamic sensors with integrating and threshold properties.

Observations of the kinetics of minority-carrier generation near semiconductor surfaces yield extensive information on the physical properties of semiconductor–insulator interfaces, in particular, on the minority-carrier thermal-generation rate G in a surface layer with nonequilibrium depletion, i.e., the most important parameter controlling the information charge storage time in metal–insulator–semiconductor (MIS) structures functioning in a permanently transient mode [1–5].

When all three (conventionally considered) minority-carrier generation channels (electron–hole pair generation via interface states, deep subsurface bulk levels in a semiconductor, and thermal diffusion of minority carriers from its electrically neutral bulk) are simultaneously active in the case of (for definiteness) an n-type semiconductor, we have (see [4, 6])

$$G = G_{ss} + G_{sv} + G_{sD}, \quad G_{sv} = n_i W/2\tau, \quad G_{sD} = D_p n_i^2/L_p N_d, \tag{1}$$

where G_{ss} and G_{sv} are the rates of minority-carrier generation via the interface states (ISs) and bulk levels, respectively; G_{sD} is the rate of thermal diffusion of holes from the electrically neutral bulk of the semiconductor; n_i is the intrinsic carrier concentration in the semiconductor; W is the width of the nonequilibrium depletion layer; τ, D_p, and L_p are the hole lifetime, diffusivity, and diffusion length, respectively; and N_d is the donor impurity concentration. It should be noted that the term $G_{sv} = n_i W/2\tau$ is inexact. The self-consistent theory of minority-carrier generation via bulk levels existing in the space-charge region (SCR) of a semiconductor, which is confirmed experimentally, yields [7, 8] $G_{sv} = N_{vg} W_g/\tau_v$, where N_{vg} is the concentration of bulk generation centers, W_g is the generation-region width, W_g is smaller than W and decreases with the observation time t, and τ_v is the electron lifetime at a given bulk level. In modern silicon MIS structures, this minority-carrier generation channel barely manifests itself at all: bulk levels located in the SCR near the midgap are either completely absent or their concentrations are rather low [4, 5]. At the same time, minority-carrier diffusion from the electrically neutral Si bulk becomes significant only at temperatures $T > 373$ K [4]: the corresponding current density of minority-carrier generation, $j_{sD} = qG_{sD} = qD_p^{1/2} n_i^2/(N_d \tau_p^{1/2}) \approx 1.9 \times 10^{-11}$ A/cm^2, is lower than the typical experimental values of qG by two to four orders of magnitude (q is the elementary charge, $\tau_p \approx 10^{-6}$ s [4], $D_p \approx 12$ cm^2/s, $n_i = 8.34 \times$

10^9 cm^{-3}, $N_d = 2 \times 10^{15}$ cm^{-3}, and $T = 293$ K). Hence, according to expressions (1), hole accumulation near a nonequilibrium-depleted Si surface at $T \approx 300$ K should be controlled by IS-related hole generation, whose rate G_{ss} (disregarding recombination), in terms of the Shockley–Read–Hall statistics [9, 10], is given by

$$G_{ss} = \int_{E_c}^{E_v} dE N_{ss}(E)/(\tau_c + \tau_v), \quad (2)$$

where $E_c = 0$ is the conduction-band bottom; E_v is the valence band top; $E_v - E_c = E_g$ is the band gap; $N_{ss}(E)$ is the spectral density of interface states; $\tau_c = \tau_{c0}\exp(E/kT)$ and $\tau_v = \tau_{v0}\exp[(E_g - E)/kT]$ are the electron and hole lifetimes at ISs with energy E, $E_c \le E \le E_v$; the preexponential factors τ_{c0} and τ_{v0} are assumed to be weakly dependent on energy E, and k is the Boltzmann constant. Taking into account the fact that the major contribution to the generation flux, and hence to integral (2), is made by ISs localized in a band of width $\sim kT$ near the midgap of these semiconductors, expression (2) can be simplified to

$$G_{ss} \approx \pi[\exp(-E_g/2kT)]N_{ss}(E_{\text{eff}})kT/2\tau_{\text{eff}}, \quad (2a)$$

$$E_{\text{eff}} = E_g/2 + [kT\ln(\tau_{v0}/\tau_{c0})]/2, \quad \tau_{\text{eff}} = (\tau_{c0}\tau_{v0})^{1/2}.$$

It follows from expressions (2) and (2a) that the generation rate G_{ss} should be independent of t and V_g if the degree of nonequilibrium depletion of the surface is sufficiently high ($|\Psi_{s0}| = |\Psi_s|_{t=0} \approx |V_g| \gg E_g/q$, where Ψ_s is the surface potential of a semiconductor and $V_g < 0$ is the depletion voltage) at the initial point in time $t = 0$.

Expressions (2) and (2a) describe only the initial generation stage at $p(t) < p^* \ll p_\infty$, i.e., under conditions far from equilibrium.[1] Here, $p(t)$ and p_∞ are three-dimensional concentrations of holes collected near the surface of a semiconductor by the time t and at the point of equilibrium, respectively; $p^* = N_v(\tau_{v0}/\tau_{c0})^{1/2}\exp(-E_g/2kT)$; and N_v is the effective density of states in the valence band. At $p(t) > p^*$, the generation rate is sharply reduced due to the attainment of quasi-equilibrium between the electrons trapped at ISs and free holes near the interface: the electron flux from the valence band to the ISs (hole generation) and the inverse flux from the ISs to the valence band (hole recombination) are almost equal. It should be noted that the characteristic bulk concentration p^* of holes at the surface of a semiconductor is low (e.g., $p^* \approx n_i = 8.34 \times 10^9$ cm^{-3} for Si at $T = 293$ K); this value corresponds to the surface concentration $p_s^* = p^*\varepsilon_0\varepsilon_s kT/q^2 N_d W$, where ε_0 and ε_s are the permittivities of free space and a semiconductor. At $T = 293$ K, $N_d \approx 10^{15}$ cm^{-3}, $\tau_{c0} = \tau_{v0}$, $W > 10^{-4}$ cm, and $p_s^* = 10^4$ cm^{-2}. Hence, the minority-carrier generation should mostly take place under the conditions of quasi-equilibrium between the electrons trapped at ISs and the free holes near the interface. In this case, the Shockley–Read–Hall statistics yields the following expression for the rate of minority carrier generation via ISs (see [7]):

$$G_{ss} = \left\{\exp\frac{E_g - F_p}{kT}\right\}\left\{\int_{E_b}^{E_g - F_p}\frac{N_{ss}(E)}{\tau_{c0}}dE\right\} \times\left\{1 - \exp\frac{E_g - F_p - F_e - q\Psi_s}{kT}\right\}. \quad (3)$$

Here, F_p and F_e are the quasi-Fermi level energies of the holes and electrons at the interface (the quasi-level F_e coincides with the bulk Fermi level F in the electrically neutral region of a semiconductor and the energy F_p is measured from the valence-band top toward the conduction band), $\Psi_s = \Psi_s(t)$ is the surface potential, and $E_b = F_p + kT\ln(\tau_{v0}/\tau_{c0})$ is the energy corresponding to the upper boundary of the generation region (E_b is measured from the level $E_c = 0$ toward the valence band). In the energy range $E_b < E < E_g - F_p$, the IS population is defined by the Boltzmann distribution, i.e., it is proportional to $\exp[(E - E_g + F_p)/kT]$; at $E < E_b$, the IS occupancy is negligible; and, at $E > E_g - F_p$, it almost attains the limiting value. The minority-carrier generation via ISs with energies $E < E_b$ and $E > E_g - F_p$ is inefficient, since its activation energies exceed $E_g - F_p$. The pattern of the sharp decrease in G_{ss} at $p > p^*$ ($p_s > p^*$) is illustrated by the energy-band diagram in Fig. 1.

Comparing expressions (2) and (3) and setting τ_{c0} and τ_{v0} equal to each other for the purpose of estimation, we obtain

$$G_{ss}|_{p>p^*}/G_{ss}|_{p>p^*} \approx \{\exp[(E_g/2) - F_p]/kT\} \times [kTN_{ss}(E_g/2)]/\int_{F_p}^{E_g - F_p} N_{ss}(E)dE. \quad (4)$$

According to expression (4), at relatively high depleting voltages corresponding to the equilibrium state of deep inversion and after accumulation of the hole concentration $p \gg p^*$ near the interface, the hole generation rate via ISs at the main stage of electron–hole pair generation ($p \gg p^*$) decreases by a factor of $\sim\exp(E_g/2kT) > 10^9$, since F_p becomes much smaller than E_g. Therefore, at a high level of silicon technology (bulk generation centers in the interface region are inefficient, the total IS density is $N_s \le 10^{10}$ cm^{-2}, τ_{c0} and $\tau_{v0} \approx 10^{-10}$ s [1, 4]), the equilibrium hole concentration near the SiO_2/Si interface during hole generation via ISs would take more than 10^4 years to become established. Since the thermodiffusion channel of minority carrier generation

[1] Nevertheless, relations (2) and (2a) are commonly used in practice to describe electron–hole pair generation until the point at which MIS structures reach equilibrium [4, 6].

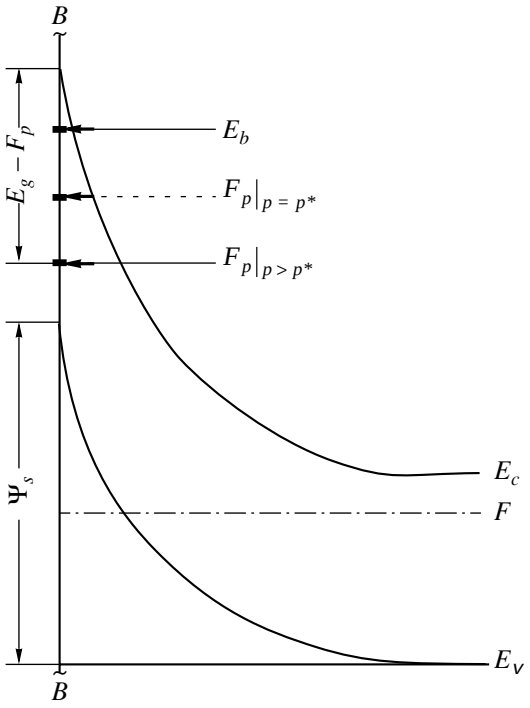

Fig. 1. An energy-band diagram of the space-charge region adjacent to an insulator over the *B*–*B* plane of an *n*-type semiconductor at the stage of the transition from highly nonequilibrium depletion to inversion: $p^* < p \ll p_\infty$ (p is the three-dimensional hole concentration near the semiconductor surface). The energy E of interface states and the upper boundary E_b of the hole-generation region are measured from the conduction-band bottom E_c toward the valence band top E_v, and the energy F_p of the quasi-Fermi level of holes is measured from the level E_v toward the conduction band (F is the bulk Fermi level). At the initial point in time ($t = 0$ and $p = 0$), the electron–hole pair generation is described by Eqs. (2) and (2a) in the region $p(t) < p^*$ at a profound depletion of the semiconductor surface. The surface hole-generation rate G_{ss} in the case shown by the diagram ($p > p^*$) obeys relation (3). The position of the quasi-Fermi level $F_p|_{p = p^*}$ of holes, corresponding to the onset of the sharp decrease in G_{ss}, is shown conditionally, since it is reached at the preceding relaxation stage under a much stronger nonequilibrium band bending.

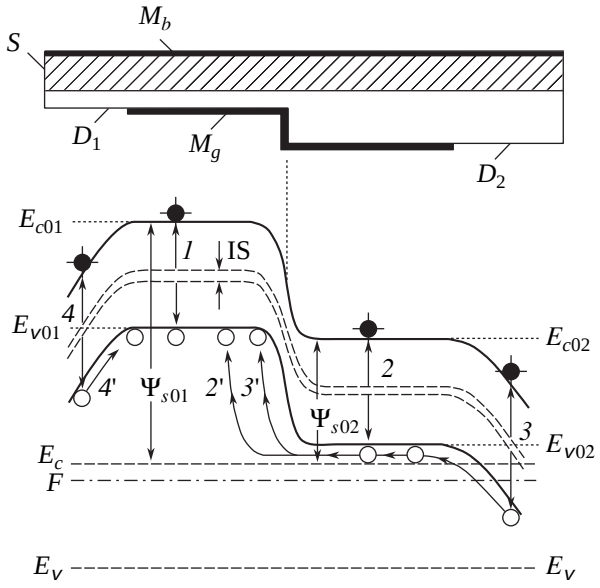

Fig. 2. An MIS structure with a planar–inhomogeneous insulator (top): M_b is the metallization layer from the side, S is the semiconductor, D_1 (thickness h_1) and D_2 (thickness h_2) are the insulator, and M_g is the field electrode. An Energy-band diagram for the state of profound nonequilibrium depletion (the semiconductor energy bands are shown from the side of the interface with the insulator) (bottom): E_c, E_v, and F are the conduction-band bottom, the valence-band top, and the Fermi level in the electrically neutral bulk of the semiconductor. E_{c01}, E_{c02}, E_{v01}, E_{v02}, Ψ_{s01}, and Ψ_{s02} are the positions of the conduction-band bottom, valence-band top, and nonequilibrium surface potential of the semiconductor at the interface under the thin and thick insulators, respectively. IS are interface states located at the semiconductor midgap. Vertical arrows *1* and *2* indicate the thermal generation of electron–hole pairs under the field electrode while *3* and *4* indicate this generation over the field electrode periphery. Generated electrons "slide down" toward the electrically neutral bulk of the semiconductor (beyond the figure plane). The holes generated in the deep potential well (the left-hand diagram region, arrow *1*) remain there, and the holes generated over the gate periphery and under the thick insulator region (arrows *2*–*4*) are "pulled" under the thin insulator to the deepest potential well (arrows *2*'–*4*').

in Si at $T \approx 300$ K is also inefficient, observations of finite times for the nonequilibrium state relaxation of silicon MOS structures at room temperature suggest that there exist alternate channels of minority carrier generation not considered in (1). These channels include electron–hole pair generation in edge fields of different types [11–13]. In particular, the main channel of thermal generation of minority carriers can give rise to an electron–hole pair over the field electrode periphery, which is accompanied by "pulling" of the minority carriers into the nonequilibrium depletion region located directly under the gate. Such a channel can be artificially "amplified" by forming a contact between the semiconductor and a planar-inhomogeneous insulator (Fig. 2). As was shown in [11], an insulating "step" under the gate significantly decreases the time of information-charge storage in dynamic memory cells that are in a nonequilibrium state of pronounced inversion. Such "steps" are inherent to modern nanoscale electronic systems with an ultrahigh packing density, combining thin ($\lesssim$100 Å) gate layers with a thick ($\gtrsim$1000 Å) insulation of the components, interconnections, etc.

The kinetics of isothermal minority-carrier generation at the periphery of MIS structures, despite the obvious physical interest and actual practical significance, has not been analyzed in detail. Below, using the example of a model MOS structure based on *n*-Si with

a nonuniform-step insulator, we study the features of the kinetics of edge-related minority-carrier generation and consider certain information related to and applied possibilities of their manifestations.

Let us consider the thermal generation of holes via ISs in such a system at a gate-insulator thickness $h_1 \ll h_2$. If voltage $V_g < 0$ is applied to the field electrode, the bands in the *n*-type semiconductor bend upward and layers of nonequilibrium depletion of various widths W arise at the initial time point $t = 0$: $W_1(h_1) > W_2(h_2)$. Two potential wells with different depths for holes arise near the semiconductor surface, with the deeper one arising under a thin insulator layer. Hole generation via ISs take place in both wells; however, hole accumulation in the shallow well is impossible at the initial stage, as the holes drain into the deep well (see the energy-band diagram in Fig. 2, arrows *2* and *2*'). Shallow wells are also formed in the voltage slope regions caused by edge fields of the gate over its periphery. A similar voltage slope also arises near the vicinity of the insulator step. The holes generated in these regions and "pulled" under the field electrode, make a certain additional contribution to the charge of the inversion layer of the deep well (arrows *3*, *3*' and *4*, *4*' in Fig. 2). The corresponding transitions at the interface of thin and thick insulators are not shown in Fig. 2. Their contribution to the total generation rate is in fact rather small, since this interface is thin in comparison with the gate perimeter.

Once quasi-equilibrium between the holes and electrons at the IS in the deep well is established ($t = t_1$), the generation rate G_{ss} decreases sharply, but not to the level to which it would fall if the insulator was uniform in thickness; G_{ss} remains high enough to be controlled by the hole generation rate in the shallow well, where quasi-equilibrium has not yet been established (free holes have not accumulated due to their continuous outflow into the deep well). This situation is retained until homogenization of the semiconductor surface under the common field electrode ($t = t_2$), i.e., until the leveling of its surface potential (band bending) under regions of thin and thick insulators: $\Psi_{s1}(h_1) \approx \Psi_{s2}(h_2)$. Thereafter ($t > t_2$), quasi-equilibrium of the holes and electrons at the ISs in both wells is established, and G_{ss} should sharply decrease to a level characteristic of a structure with a homogeneous insulator until reaching thermodynamic equilibrium. However, this equilibrium cannot occur, since the peripheral edge effect continues to maintain G_{ss} at a level high enough to prevent it, as in the case of the coexistence of shallow and deep wells under the field electrode at times $t < t_1$.

Under these circumstances, the generation current flowing in the external circuit exhibits three "steps": the first step relates to the part of the process before quasi-equilibration in the deep well between the electrons trapped at ISs and the free holes near the interface ($0 \le t \le t_1$) is established; the second step relates to the time

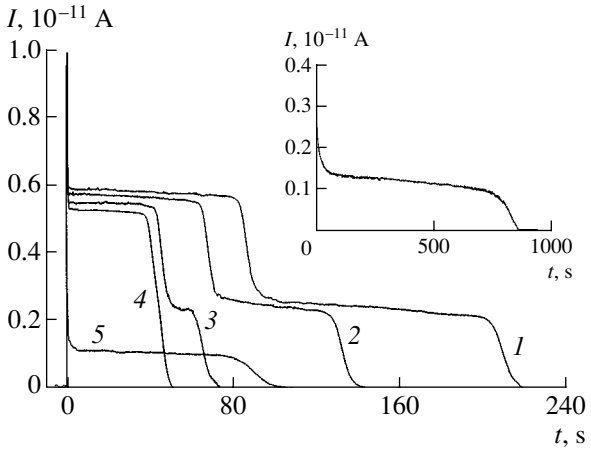

Fig. 3. Hole generation kinetics $I(t)$ at various depleting voltages V_g at the field electrode of a silicon MOS structure with a nonuniform-step gate insulator. The structure is Al–n^+-Si–SiO_2–*n*-Si:P–Al (SiO_2: $h_1 = 100$ Å and $S_1 = 4 \times 10^{-4}$ cm^2; $h_2 = 3200$ Å and $S_2 = 4.405 \times 10^{-4}$ cm^2; and *n*-Si:P has a (100) orientation, where $N_d = 2 \times 10^{15}$ cm^{-3}). V_g = (*1*) –6.19, (*2*) –4.63, (*3*) –3.05, (*4*) –2.65, and (*5*) –1.60 V. The inset shows the kinetics of the hole-generation current in an MOS structure based on *n*-Si with an oxide uniform in thickness ($h = 90$ Å); in this case, the gate area is $S = 1.6 \times 10^{-3}$ cm^2 and the depletion voltage is $V_g = -3$ V.

before the homogenization instant ($t_1 \le t \le t_2$); and the third, from the instant t_2 to the instant of complete equilibrium ($t = t_\infty$). The third step should be associated with hole generation at the field electrode periphery. In this, shallowest, well, which is always formed due to the edge effect, quasi-equilibrium between the electrons at ISs and the free holes can be attained only after the formation of an equilibrium inversion layer at the interface, due to the continuous hole drain into deeper wells under the gate. It is obvious that the number and duration of steps in curves $I(t)$ should depend on V_g and on the number of insulator thickness variations.

Figure 3 shows the typical curves of the hole-generation kinetics (in relation to V_g) in the MOS structure based on *n*-Si possessing a planar-inhomogeneous oxide and common gate (Fig. 2). The generation current $I(t)$ was measured at 293 K in a sealed and completely darkened air chamber incorporated into a digital automated system [14]. The gate areas above a thin ($h_1 = 100$ Å) pyrolytic oxide, which, in turn, was above a thick ($h_2 = 3200$ Å) thermal oxide, are $S_1 = 4 \times 10^{-4}$ cm^{-2} and $S_2 = 4.405 \times 10^{-4}$ cm^2, respectively. The table lists data characterizing the initial ($t = 0$ is shown by the subscript 0, and final equilibrium ($t = t_\infty$, $I(t_\infty) = 0$), by the subscript ∞) states of the MOS structure under thin (subscript 1) and thick (subscript 2) oxides. The calculations were carried out on the basis of the dependence

Parameters of the initial and final states of the MOS structures

Curve no. in Fig. 3	$-V_g$, V	Thin oxide, h_1 = 100 Å					Thick oxide, h_2 = 3200 Å					
		$-\Psi_{s01}$, V	W_{01}, 10^{-4} cm	$-\Psi_{s\infty 1}$, V	$W_{\infty 1}$, 10^{-4} cm	$(p_{s\infty 1} + N_d W_{\infty 1})$, 10^{13} cm^{-2}	$-\Psi_{s02}$, V	W_{02}, 10^{-4} cm	$-\Psi_{s\infty 2}$, V	$W_{\infty 2}$, 10^{-4} cm	$(p_{s\infty 2} + N_d W_{\infty 2})$, 10^{11} cm^{-2}	$-V_{g02}$, V
1	6.19	5.8783	1.9647	0.9315	0.7821	1.1041	2.7002	1.3316	0.7600	0.7064	3.9488	0.441
		5.9914	1.9835	0.9326	0.7826	1.1285	2.8132	1.3592	0.7621	0.7074	4.0979	0.6648
2	4.63	4.3439	1.6890	0.9140	0.7748	0.7720	1.7932	1.0851	0.7406	0.6974	2.8763	0.387
3	3.05	2.7950	1.3548	0.8859	0.7627	0.4379	1.0274	0.8214	0.7069	0.6813	1.8605	0.423
4	2.65	2.4041	1.2565	0.8761	0.7585	0.3542	0.8148	0.7315	0.6873	0.6718	1.5838	0.392
5	1.60	1.3824	0.9528	0.8251	0.7361	0.1390	–	–	–	–	–	–

of the semiconductor SCR width W on the insulator thickness h, V_g, p_s, and t:

$$W = \left[-\frac{2\varepsilon_0\varepsilon_s}{qN_d}(V_g + V_{g0}) - \frac{2h\varepsilon_s}{\varepsilon_i N_d}p_s(t) + \left(\frac{h\varepsilon_s}{\varepsilon_i}\right)^2\right]^{1/2} - \frac{h\varepsilon_s}{\varepsilon_i}. \quad (5)$$

We also used the relation between the equilibrium density $p_{s\infty} + N_d W_\infty$ of the surface charge and the equilibrium surface potential $\Psi_{s\infty}$ [15]:

$$p_{s\infty}(V_g) + N_d W_\infty(V_g) = \frac{\sqrt{2}\varepsilon_0\varepsilon_s kT}{qL_n}\left[e^{q\Psi_{s\infty}/kT} - \frac{q\Psi_{s\infty}}{kT} - 1 + \left(\frac{n_i}{N_d}\right)^2\left(e^{-q\Psi_{s\infty}/kT} + \frac{q\Psi_{s\infty}}{kT} - 1\right)\right]^{1/2}. \quad (6)$$

Formula (5) is the solution to a set of equations that relate the field in an oxide with the surface charge at $V_g < 0$,

$$V_i(t)/h = [V_g - \Psi_s(t) + V_{g0}]/h = -q[p_s(t) + N_d W(t)]/\varepsilon_0\varepsilon_i, \quad (7)$$

and W with the semiconductor surface potential Ψ_s,

$$\Psi_s(t, V_g) = -qN_d W^2(t, V_g)/2\varepsilon_0\varepsilon_s. \quad (8)$$

Here, $L_n = (\varepsilon_0\varepsilon_s kT/qN_d)^{1/2}$ is the Debye length, V_i is the voltage drop across the oxide, ε_i is the oxide permittivity, Ψ_s is the surface potential (in volts) measured from the conduction band bottom in the electrically neutral Si bulk ($\Psi_s < 0$ in the depletion and inversion states), V_{g0} is the voltage of "flat bands" that results from the gate–semiconductor contact potential difference and a fixed charge in the oxide.

For further calculations, it is convenient to write Eq. (7) as

$$p_s(t) + N_d W(t) = (\varepsilon_0\varepsilon_i/qh)[-V_g + \Psi_s(t) - V_{g0}]. \quad (9)$$

First, the value of V_{g01} is calculated. An integration of curve *1* (Fig. 3, V_g = –6.19 V) within $0 \le t \le t_\infty$ determines the equilibrium surface charge $Q_{s\infty}(V_g)$ in the MOS structure. Omitting the small correction to the generation current $I(t)$, which is associated with a change in the ionized donor density $N_d\Delta W$ in the SCR during relaxation, and disregarding the equilibrium surface charge $Q_{s\infty 2}$ under the thick oxide, we set $Q_{s\infty} \approx Q_{s\infty 1}$ and $p_{s\infty 1} + N_d W_{\infty 1} = (qS_1)^{-1}Q_{s\infty}$.[2] Substituting $(qS_1)^{-1}Q_{s\infty}$ into the left-hand side of Eq. (6), we calculate $\Psi_{s\infty 1}$; then, we use Eqs. (7) and (8) to calculate V_{g01} = 0.129 V and $W_{\infty 1}$. Using this value of V_{g01} at $p_s = 0$, we determine W_{01} and Ψ_{s01} from formulas (5) and (8), respectively. If V_{g01} is known, the calculation procedure for curves *2–5* is simplified. The values of $(qS_1)^{-1}Q_{s\infty}(V_g)$ are determined in a similar way, i.e., by integrating the current $I(t)$; $\Psi_{s\infty 1}(V_g)$ and $W_{\infty 1}(V_g)$ are determined using Eqs. (9) and (8); $W_{01}(V_g)$ is determined using formula (5) at $p_{s1} = 0$, and $\Psi_{s01}(V_g)$ is also determined from Eq. (8).

At the instants when the homogenization of both regions of the MOS structure occur $t = t_2(V_g)$ and $t_2(V_g)$ are times of the ends of the first wide steps in $I(t)$ curves *1–4*, $\Psi_{s1}(V_g, t_2) \approx \Psi_{s02}(V_g)$ and $W_1(V_g, t_2) \approx W_{02}(V_g)$, since the hole density in the shallow well is still negligible. The points in time t_2 were determined using the positions of the minima of the derivatives dI/dt on scale t, which were obtained by numerical differentiation of curves *1–4* in Fig. 3 (see Fig. 4). The surface charge in the deep well at the instant $t = t_2$ is $Q_{s1}(t_2) < Q_{s\infty 1}(t_\infty) \approx Q_{s\infty}$, and the disregard of the width decrease of the SCR under the thin oxide may turn out

[2] The change in the density of ionized donors is $N_d\Delta W = N_d[W_{01} - W_1(t)]$. At $t = t_\infty$, $\Delta W = W_{01} - W_{\infty 1}$, and, in the state of profound inversion, $p_{s\infty 1} + N_d W_{\infty 1} \gg N_d\Delta W|_{t = t_\infty}$. In this case, the ratio of surface charges under the thin and thick oxides is $Q_{s\infty 1}/Q_{s\infty 2} \le C_{i1}/C_{i2} = S_1 h_2/S_2 h_1 = 32$ ($S_1 \approx S_2$), where C_{i1} and C_{i2} are the capacitances of the thin and thick oxides; i.e., $Q_{s\infty 1} \gg Q_{s\infty 2}$.

to be unjustified.[3] In the range $0 \le t \le t_2$, the change in the charge density in the inversion layer is $\Delta Q_{s1}(t_2, V_g) = qS_1\{p_{s1}(t_2, V_g) - N_d[W_{01}(V_g) - W_1(t_2, V_g)]\}$ and is equal to the integral of the generation current $I(t)$ over this range or

$$p_{s1}(t_2, V_g) + N_d W_1(t_2, V_g) = (qS_1)^{-1}\int_0^{t_2} I(t, V_g)dt + N_d W_{01}(V_g). \quad (10)$$

The values of $W_{01}(V_g)$ are listed in the table. Integrating curves *1–4* (Fig. 3) within $0 \le t \le t_2$ and introducing corresponding corrections (10) into the integration results, we determine the values of $p_{s1}(t_2, V_g) + N_d W_1(t_2, V_g)$. Substituting these values and $V_{g01} = 0.129$ V into Eq. (9), we calculate $\Psi_{s1}(t_2, V_g) \approx \Psi_{s02}(V_g)$ and then determine $W_1(t_2, V_g) \approx W_{02}(V_g)$ using formula (8). Now Eq. (9) at $h = h_2$ and $p_{s2} = 0$ allows the calculation of $V_{g02}(V_g)$. The values of V_{g02} and V_{g01} are different, which, generally speaking, is not unexpected if we take into account the difference in the properties of the Si contact with the thin and thick oxides. After substituting V_{g02} into the transcendental equation that arises from equating the right-hand sides of (9) and (6), we calculate $\Psi_{s\infty 2}(V_g)$ and $p_{s\infty 2}(V_g) + N_d W_{\infty 2}(V_g)$, and then we use formula (8) to calculate $W_{\infty 2}(V_g)$. A similar analysis of the structure state under the thick oxide using curve *5* is impossible because of the absence of a homogenization transition in this structure. This analysis would provide an independent determination of Ψ_{s02}, W_{02}, and V_{g02}, which are required to calculate $\Psi_{s\infty 2}$, $W_{\infty 2}$, and $Q_{s\infty 2}$ using Eqs. (5)–(10).

The experimental data shown in Fig. 3 and listed in the table generally agree well with the considered features. At voltages $|V_g| > 2.65$ V, the current $I(t)$ features two steps, the first of which is flatter. As $|V_g|$ decreases, both steps become lower and their width is reduced; i.e., the time t_∞ required for the establishment of the equilibrium inversion state shortens. After the establishment of this state, a specific contact potential difference $\Delta V_\infty = \Psi_{s\infty 1} - \Psi_{s\infty 2}$ arises near the insulating step over the semiconductor surface, which is caused by the different positions of the valence-band top with respect to the Fermi level at the Si surface. Under the conditions of this experiment, ΔV_∞ only slightly depends on V_g and is ~0.18 V.

The "zero" step expected in the range $0 \le t \le t_1$ on this time scale cannot be observed: to an order of magnitude, the time at which this step ends is $t_1 \sim t_2 p_s^*/p_{s1}(t_2)$, where t_2, as before, is the homogenization time of the MOS structure. If $t_2 \sim 10^2$ s (see Fig. 3) and $p_s^*/p_{s1}(t_2) < 10^{-7}$ ($p_s^* \sim 10^4$ cm^{-2} and $p_{s1} > 10^{11}$ cm^{-2}), then $t_1 < 10^{-5}$ s. At the same time, if the interface between the thin oxide and Si contained hole-trapping centers with a certain density N_{s0}, the initial generation phase would be longer, since localized minority carriers, in contrast to free ones, do not recombine at V_g = const. In this situation, the time t_1 at which the zero step ended would be approximately equal to the time t' required to obtain the ultimate population of all such centers. Assuming that generation current densities in the deep and shallow wells ($(I_0 - I_2)/S_1$ and $(I_2 - I_3)/S_2$, respectively) are identical in the range $0 \le t \le t'$, we determine the height and duration of the "zero" step: $I_0 = (1 + S_1/S_2)I_2 - S_1 I_3/S_2$ and $t' = qS_1N_{s0}/I_0$ (I_2 and I_3 are the average currents at the first and second observed steps). At $S_1 \approx S_2$ and $N_{s0} = 10^{11}$ cm^{-2}, e.g., for curve *1* in Fig. 3, we have $I_0 \approx 9 \times 10^{-12}$ A and $t' \approx 0.7$ s. A similar calculation for curve *5* in Fig. 3 (the well is inactive) yields $I_0 \approx 4 \times 10^{-12}$ A and $t' \approx 1.6$ s. The measuring system makes it possible to resolve the time intervals $\Delta t \approx 0.1$ s; however, the zero step is not detected in all the $I(t)$ curves at this time resolution. This means that $N_{s0} < 10^{11}$ cm^{-2} and that the holes are mostly free in the equilibrium inversion layer of the deep well ($h = h_1$) at $p_{s\infty} \gg 10^{11}$ cm^{-2}. The condition $p_{s\infty 1} \gg 10^{11}$ cm^{-2} is satisfied for the deep well at all values of V_g (see table).

It is important that the density of holes has drifted by the homogenization time $t = t_2$ from the shallow into the deep well, $\Delta p_{s1}(t_2) \approx (qS_1)^{-1}(I_2 - I_3)t_2 \gg p_{s\infty 2}$, where $p_{s\infty 2}$

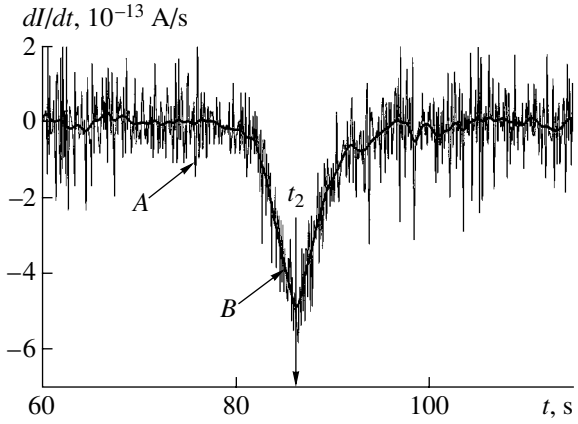

Fig. 4. Determination of the instant corresponding to leveling of the Si surface potential (homogenization instant, $t = t_2$) in the deep and shallow wells of an MOS structure from the time dependence of the derivative of the hole generation current, $dI(t)/dt$. *A* is the result of the numerical differentiation of curve *1* in Fig. 3, and *B* is the derivative $dI(t)/dt$ plotted using an adaptive algorithm developed on the basis of the Tikhonov regularization [16, 17]. The vertical arrow indicates the point $t = t_2$.

[3] Under the thick oxide in the range $0 \le t \le t_2$, the charge of the semiconductor depletion layer remains unchanged until homogenization of the surface potential of both structure components due to the continuous hole drain into the deep well; therefore, $W_2(t, V_g) = W_{02}(V_g) = \text{const}$ and $p_{s2}(V_g)|_{t=0} = p_{s2}(V_g)|_{t=t_2} = 0$.

is their equilibrium density in the shallow well. In particular, in the case of curve *1*, $\Delta p_{s1}(t_2) \approx 4 \times 10^{12}$ cm$^{-2} \gg p_{s\infty 2} \approx 2.5 \times 10^{11}$ cm^{-2}. This circumstance suggests that hole trapping is insignificant in the current generation kinetics, since the actual density of the trapping centers must be lower than $p_{s\infty 2}$. Moreover, it is known that the IS density is, on average, $N_s \ll 10^{11}$ cm^{-2} in good silicon MOS structures [4].

The dependence of the current on V_g and step flatness, which are not characteristic of minority carrier generation via ISs in homogeneous MIS structures, directly indicates that the peripheral (edge) generation channel is active: as holes are accumulated under the field electrode, the peripheral shallowest well becomes narrower and the minority carrier generation rate decreases. In contrast, the well widens as $|V_g|$ increases, the generation region widens, and the current increases.

In the region $|V_g| < 2.65$ V, the generation-current level is gradually lowered as $|V_g|$ decreases, and the dependences $I(t)$ feature a single step (Fig. 3, curve *5*). This step should be related to hole generation only over the gate periphery under the thin oxide ($S = S_1$), i.e., exclusively to the edge effect. This conclusion is confirmed by an $I(t)$ dependence similar to curve *5*, which was obtained for an MOS structure based on *n*-Si with an insulator uniform in thickness (Fig. 3, inset). Applying the considered calculation algorithm to this dependence, we obtain $\Psi_{s0} = -1.5$ V, $W_0 = 1.51 \times 10^{-4}$ cm, $\Psi_{s\infty} = 0.883$ V, $W_\infty = 0.762 \times 10^{-4}$ cm, $p_{s\infty} + N_d W_\infty = 4.26 \times 10^{12}$ cm^{-2}, and $V_{g0} = 0.27$ V. In both cases, i.e., at low depleting voltages, the region of the initial transition to the slow stage of minority-carrier generation is significantly wider than in curves *1–4*. Apparently, this region is still simply the rapidly damping initial generation stage described by Eq. (3). Under the conditions of pronounced initial depletion, this stage is obscured by a high generation current caused by the existence of a shallow well whose area is much larger than the area of the peripheral generation region.

We should emphasize that the shallow well in fact simulates the structure periphery at large $|V_g|$ (Fig. 3, curves *1–4*); however, under other conditions, the geometric parameters of the shallow well are unchanged, and the rate G_{ss} of electron–hole pair generation via ISs at the interface of Si and the thick oxide remains unchanged until the instant at which the homogenization of both components ($t \approx t_2$) occurs. In other words, the holes in the shallow well are generated in the recombinationless mode described by relations (2) and (2a). The corresponding generation current is obviously equal to the difference of the first and second step heights, which is observed in the $I(t)$ curves: $\Delta I = I_2 - I_3 \approx 3 \times 10^{-12}$ A (Fig. 3, curve *1*). Using set (2a), we obtain $\Delta I = qS_2G_{ss} = (\pi/2)qS_2[\exp(-E_g/2kT)]N_{ss}^* kT/\tau_{\text{eff}}$, $N_{ss}^* \equiv N_{ss}(E_g/2)$, which allows a fairly exact calculation of the ratio $N_{ss}^*/\tau_{\text{eff}} = 4.7 \times 10^{21}$ cm^{-2} eV^{-1} s^{-1}, and $G_{ss} = \Delta I/qS_2 = 4.25 \times 10^{10}$ cm^{-2} s^{-1} ($T = 293$ K, $E_g = 1.12$ eV, and $S_2 = 4.405 \times 10^{-4}$ cm^2). The IS density N_{ss}^* near the Si midgap is sufficiently accurately determined using equilibrium capacitance spectroscopy [2, 3, 18]. The quasi-static *C–V* characteristics of the structure under study showed that $N_{ss}^* = 6.4 \times 10^{10}$ cm^{-2} eV^{-1}.[4] Then, $\tau_{\text{eff}}[v_T\sigma_{\text{eff}}(N_cN_v)^{1/2}]^{-1} = 1.5 \times 10^{-11}$ s (v_T is the thermal rate of minority-carrier generation and σ_{eff} is the effective cross section of trapping by the generating surface centers), and, at $v_T = 10^7$ cm s^{-1} and $(N_cN_v)^{1/2} = 1.72 \times 10^{19}$ cm^{-3} ($T = 293$ K), we obtain $\sigma_{\text{eff}} = 4 \times 10^{-16}$ cm^2. These values of N_{ss}^*, τ_{eff}, σ_{eff}, and G_{ss} agree well with the published data [1–4, 11].

As was shown above, G_{ss} decreases by a factor of $\exp(E_g/2kT)$ during the transition to the quasi-equilibrium generation mode (see expression (3)). Hence, in the absence of a peripheral channel of hole generation, the value $G_{ss} = 4.25 \times 10^{10}$ cm^{-2} s^{-1} should decrease to $\tilde{G}_{ss} = 4.25 \times 10^{10}/\exp(E_g/2kT) = 9.9$ cm^{-2} s^{-1} after the structure's homogenization, the equilibrium inversion layer would take $(p_{s\infty 2} + N_dW_{\infty 2})/\tilde{G}_{ss} \approx 1280$ years to form ($p_{s\infty 2} + N_dW_{\infty 2} \approx 4 \times 10^{11}$ cm^{-2}), even in the shallow well at $V_g = -6.19$ V.

In the calculations related to the structure region under the thin oxide, changes in the ionized donor density $N_d(W_{01} - W_{\infty 1})$ in the deep well and in the surface charge $q[p_{s\infty 2} + N_d(W_{02} - W_{\infty 2})]$ in the shallow well during the transition from depletion ($t = 0$) to inversion ($t = t_\infty$) were disregarded. Taking into account these factors at $V_g = \text{const}$, the charge density $(qS_1)^{-1}Q_{s\infty 1}(V_g)$ in the equilibrium layer of inversion under the thin oxide is given by

$$p_{s\infty 1} + N_d W_{\infty 1} = (qS_1)^{-1}\int_0^{t_\infty} I(t)dt + N_d(W_{01} + W_{02}S_2/S_1) - (p_{s\infty 2} + N_d W_{\infty 2})S_2/S_1. \quad (11)$$

According to Eq. (11), corrections to the initial data on can be introduced using iterations. Let us consider the corresponding results using the example of curve *1* (Fig. 3). We assume the tabulated values of the parameters to be a zeroth approximation. Then, substituting these values into Eq. (11), to the first approximation, we obtain the value $p_{s\infty 1}^{(1)} + N_d W_{\infty 1}^{(1)} = 1.1302 \times 10^{13}$ cm^{-2}, beginning with which the entire calculation algorithm is reproduced. The results of the first approximation are again substituted into (11) and so on. The iteration cycles rapidly converge, and even the third approximation, on average, differs from the second one by no more than 0.01%. The values of V_{g01} are subjected to

[4] The measurements were carried out with a linear variation of the gate voltage over time from the value corresponding to deep inversion of the Si surface ($V_g < 0$) to the value corresponding to its profound enrichment ($V_g > 0$).

the largest correction, while the values of V_{g02} are corrected to a lesser extent: $V_{g01}^{(0)} = 0.129$ V, $V_{g01}^{(3)} = 1.4286 \times 10^{-2}$ V; $V_{g02}^{(0)} = -0.441$ V, and $V_{g02}^{(3)} = -0.6648$ V. The results of the third approximation are represented by the lower values in the first row of the table. A comparison of the data in this row makes it possible to reconcile the accuracy and objective of the experiment.

Thus, observations of the minority carrier generation in MIS structures with a planar-inhomogeneous insulator show that they exhibit, on the one hand, unconventional discretized current kinetics and allow, on the other hand, an accurate and detailed characterization of semiconductor surface regions under thick and thin insulators in nonequilibrium depletion and inversion states using simple and rapid measurements. Such observations also make it possible to determine the flat-band voltages $V_{FB} \equiv V_{g01}$ and V_{g02}, as well as to estimate the parameters of the centers and rates of actual and "virtual" surface generation. Using Eqs. (5), (7)–(9), we can easily reconstruct the relaxation kinetics of the surface potential $\Psi_s(t)$ and the semiconductor SCR width $W(t)$ from the kinetics of the generation current $I(t)$. In principle, the relaxation kinetics reflects the energy distributions of both the ISs and the minority carrier generation centers, and also reflects the doping profiles. The $I(t)$ dependences for MIS structures with an insulator containing a large number regular planar inhomogeneities should detect an ordered ladder of current steps with the number $n + 1$ (not counting the zero "rapid" step), where n is the number of insulator thickness gradations. The parameters of such a ladder can be controlled by varying the height and/or width of the insulating steps, as well as their number. A sequential displacement in time of the generation activity region from the deepest (the thinnest insulator) to the shallowest well allows detection of the planar distribution of a wide set of the electric characteristics of the heterointerface, which are determined within this approach, including inhomogeneities stimulating electron–hole pair generation [8, 12].

The edge generation rate of minority carriers is profoundly affected by various external factors: light, ambient medium composition, radiation background, and others. For example, even scattered daylight increases the rate of electron–hole pair generation by tens of thousands times. These factors can be used to develop a wide range of highly sensitive sensor systems, in particular, gas sensors based on monitoring of the edge generation rate of minority carriers. In this case, it is expedient to suppress the technologically indeterminate generation activity of the peripheral region by heavy doping of the semiconductor with a major impurity along the field-electrode perimeter. This means that this activity can be replaced by a thoroughly controlled generation-active equivalent, i.e., a shallow well, whose design features should comply with the expected physical signal. The essential and practically important feature of such structures is their capability to integrate external excitation, which give them threshold properties. Finally, the analysis carried out shows that the nonequilibrium depletion state can be very long-lived in high-quality MOS structures with a suppressed channel of peripheral minority-carrier generation; i.e., such structures can potentially be used as elements of data storage systems.

ACKNOWLEDGMENTS

We would like to thank N.F. Kukharskaya, V.G. Naryshkina, and E.L. Novikova for their assistance with the experimental data processing.

REFERENCES

1. V. G. Litovchenko and A. P. Gorban', *Fundamentals of the Physics of Microelectronic Metal–Insulator–Semiconductor Systems* (Naukova Dumka, Kiev, 1978) [in Russian].
2. E. H. Nicollian and I. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982).
3. *VLSI Technology*, Ed. by S. M. Sze (McGraw-Hill, New York, 1983; Mir, Moscow, 1986), Vol. 2.
4. *Charge-Coupled Devices and Systems*, Ed. by M. J. Howes and D. V. Morgan (Wiley, Chichester, 1979; Énergoizdat, Moscow, 1981).
5. W. D. Brown and J. E. Brewer, *Nonvolatile Semiconductor Memory Technology* (IEEE Press, New York, 1998).
6. S. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981; Mir, Moscow, 1984).
7. E. I. Gol'dman and A. G. Zhdan, Mikroélektronika **23**, 3 (1994).
8. E. I. Gol'dman, A. G. Zhdan, and A. M. Sumaroka, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 2048 (1992) [Sov. Phys. Semicond. **26**, 1152 (1992)].
9. W. Shockley and W. Read, Phys. Rev. **87**, 835 (1952).
10. R. Hall, Phys. Rev. **87**, 387 (1952).
11. V. A. Gergel', V. A. Zimoglyad, N. V. Zhukov, and V. V. Rakitin, Mikroélektronika **17**, 406 (1968).
12. E. I. Gol'dman, Fiz. Tekh. Poluprovodn. (St. Petersburg) **27**, 269 (1993) [Semiconductors **27**, 150 (1993)].
13. E. I. Gol'dman, A. G. Zhdan, and A. M. Sumaroka, Pis'ma Zh. Éksp. Teor. Fiz. **57**, 783 (1993) [JETP Lett. **57**, 797 (1993)].
14. E. I. Gol'dman, A. G. Zhdan, and G. V. Chucheva, Prib. Tekh. Éksp., No. 6, 677 (1997) [Instrum. Exp. Tech. **40**, 841 (1997)].
15. C. G. B. Garrett and W. H. Brattain, Phys. Rev. **99**, 376 (1955).
16. E. I. Gol'dman and V. A. Ivanov, Preprint No. 22 [551], IRÉ RAN (Inst. of Radio Engineering and Electronics, USSR Academy of Sciences, Moscow, 1990).
17. Yu. V. Gulyaev, A. G. Zhdan, and V. G. Prikhod'ko, Preprint No. 46 [418], IRÉ RAN (Inst. of Radio Engineering and Electronics, USSR Academy of Sciences, Moscow, 1984).
18. A. G. Zhdan, N. F. Kukharskaya, and G. V. Chucheva, Prib. Tekh. Éksp., No. 2, 120 (2002) [Instrum. Exp. Tech. **45**, 256 (2002)].

Translated by A. Kazantsev

Semiconductors, Vol. 39, No. 6, 2005, pp. 674–678. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 705–709.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Thermal–Field Forward Current in GaN-Based Surface-Barrier Structures

T. V. Blank^, Yu. A. Goldberg, E. E. Zavarin, O. V. Konstantinov, and N. M. Shmidt

Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia

^e-mail: tblank@delfa.net

Submitted August 2, 2004; accepted for publication November 15, 2004

Abstract—The voltage and temperature dependences of the capacitance and forward current in surface-barrier Ni–*n*-GaN structures are experimentally studied. The results are compared with the Padovani–Stratton thermofield emission theory. It is established that, in a temperature range of 250–410 K, the forward current of the Ni–*n*-GaN surface-barrier structures (the electron density in GaN is ~10^{17} cm^{-3}) is caused by a thermofield emission of electrons, whose energy is ~0.1 eV below the potential-barrier top. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

At present, *n*-GaN-based surface-barrier structures are widely used in photodetectors of ultraviolet (UV) radiation [1–5], since GaN is a direct-gap semiconductor and its band-gap energy (E_g = 3.39 eV) is close to the photon energy separating the visible and UV spectral regions. Such photodetectors are of practical interest for UV astronomy, ozone layer monitoring, flame sensors, and water and air purity monitoring.

However, not all the observed properties of GaN-based surface-barrier structures have been physically explained in an unambiguous manner. The available publications contain no generally accepted viewpoint on the forward current mechanism, which is, to a certain extent, accounted for by a significant potential-barrier height (~1 eV). In a number of studies [1, 6, 7], it has been shown that the experimental Richardson constant differs from the theoretical one (24 A cm^{-2} K^{-2}) by several orders of magnitude, which was attributed to the effect of tunneling and inhomogeneities at the metal–semiconductor interface. Furthermore, there is no accepted viewpoint on the reverse current mechanism. For Ni–*n*-GaN and WSi_x–GaN surface-barrier structures, it has been found that, taking into account the influence of image forces on the potential barrier height, the reverse current significantly exceeds the theoretical value corresponding to thermionic emission [7–9]. In [7, 9], it was also assumed that the reverse current at low temperatures is caused by electron tunneling from the metal into the semiconductor and that, at temperatures higher than 275 K, the reverse current results from leakages over dislocations.

In this paper, we consider the experimental voltage and temperature dependences of the forward current and capacitance for Ni–*n*-GaN surface-barrier structures with a low dislocation density in their initial material in comparison with the Padovani–Stratton thermofield emission theory [10].

2. EXPERIMENTAL

We studied 3-μm-thick *n*-GaN epitaxial layers grown on (0001) sapphire substrates by epitaxy from organometallic compounds. The electron concentration in the layers, determined from capacitance–voltage (*C–V*) measurements using a mercury probe and Hall effect measurements using the Van der Pauw method, was 8×10^{16} cm^{-3}. The electron mobility was 600 and 1800 cm^2/(V s) at room temperature and 125 K, respectively. Schottky barriers were formed by Ni/Au deposition in high vacuum using an electron beam. The total thickness of the metallization layer in the transparent electrode and contact area regions was 150 and 1000 Å, respectively. The Schottky barrier area was 5×10^{-3} cm^2.

As has previously been shown, structural features have a significant effect on the parameters of a photodetector with GaN-based Schottky barriers [11]. In particular, it has been shown that a random distribution of the charged centers associated with domain walls in the mosaic structure typical of nitrides of Group III elements results in low Schottky barrier heights, high leakage currents, and persistent photoconductivity. In this study, in order to produce and study Schottky barriers, we used GaN epitaxial layers with a well ordered mosaic structure and a dislocation density of 3×10^8 cm^{-2}. Such layers exhibit coherent domain matching in their mosaic structure and the formation of dilatational interfaces [12], which leads to high mobilities and classical forms of mobility and conductivity temperature dependences. This indicates the absence of a high concentration of scattering centers.

The dependence of the differential capacitance *C* on voltage *V* was measured in the temperature range *T* = 150–450 K. It was found that the capacitance is independent of the measuring signal frequency *f* at *f* < 1 MHz and significantly decreases at higher frequencies; therefore, the *C–V* characteristics are given at *f* = 0.465 MHz.

1063-7826/05/3906-0674$26.00 © 2005 Pleiades Publishing, Inc.

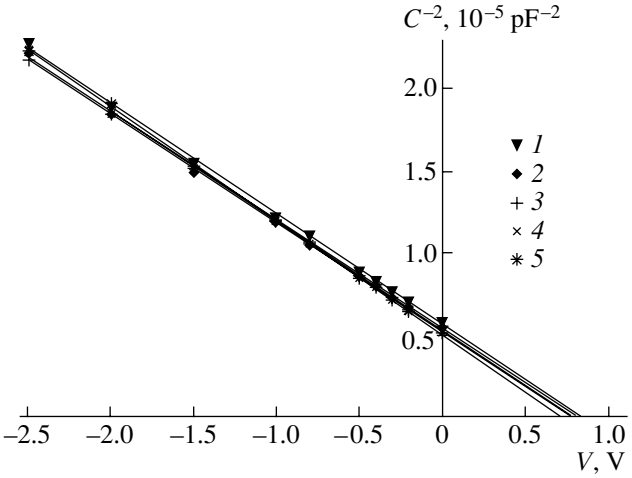

Fig. 1. Dependence of the differential capacitance C of the Ni–n-GaN surface-barrier structure on voltage V at the temperatures T = (*1*) 256, (*2*) 281, (*3*) 350, (*4*) 360, and (*5*) 400 K.

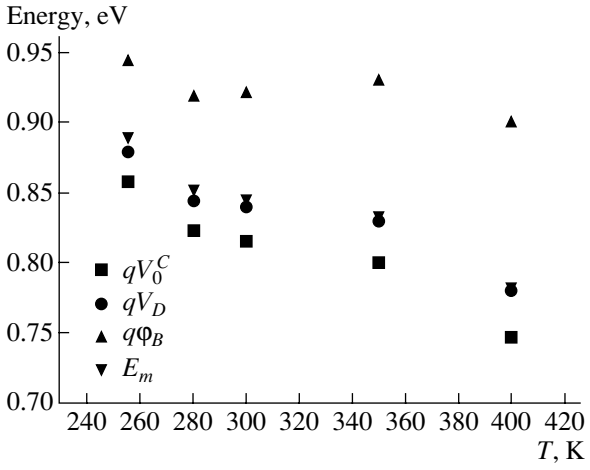

Fig. 2. Temperature dependences of the cutoff voltages of the C–V characteristics V_0^C, contact potential difference V_D, potential-barrier height $q\varphi_B$, and thermofield electron emission E_m from the semiconductor into the metal for the Ni–n-GaN surface-barrier structure ($N_d \approx 10^{17}$ cm^{-3}).

The dependence of the forward current on voltage was measured in the current range $I = 10^{-10}$–2×10^{-5} A at 150 K < T < 450 K.

3. EXPERIMENTAL RESULTS AND COMPARISON WITH THEORY

3.1. The dependence of the differential capacitance C on voltage V is linear in the coordinates $1/C^2 = f(V)$ at various temperatures (Fig. 1), which corresponds to the Schottky theory for metal–semiconductor structures. The ionized donor concentration in n-GaN, determined from the slope of these straight lines, is $N_d \approx 10^{17}$ cm^{-3} at 300 K, which is close to the electron concentration determined in the initial material. At 300 K, the cutoff voltage (determined as the point of intersection of these dependences with the horizontal axis) was $V_0^C = 0.81$ V and the contact potential difference was $V_D = V_0^C + kT/q = 0.84$ V (k is the Boltzmann constant and q is the elementary charge). The Fermi level energy in GaN with respect to the conduction-band bottom was calculated using the formula $\mu = -kT/\ln(N_d/N_c)$, where the density of states in the conduction band is N_c [cm^{-3}] = $4.3 \times 10^{14}T^{3/2}$ [13]. For the ionized donor concentration $N_d = 10^{17}$ cm^{-3}, the Fermi level energy was 0.08 eV, and the potential barrier height (in energy units) was $q\varphi_B = qV_D + \mu = 0.92$ eV. We note that φ_B depends only slightly on temperature (Fig. 2) in the range T = 250–400 K. This value of φ_B is close to that in the published data: at 300 K, $q\varphi_B$ = 0.9–1.0 eV for Ni–GaN [14], 0.87–1.03 eV for Au–GaN [15], 0.91 eV for Pd–GaN [6], and 1.03 eV for Pt–GaN [6]. In a number of papers, it has been noted that when the barrier height $q\varphi_B$ is determined from C–V characteristics, it is much higher than when it is determined from the dependence of the forward current on voltage and temperature using formulas for thermionic emission. Specifically, the differences found were ~0.2 [14] and ~0.3 eV [16] for Ni–n-GaN structures; ~0.1 [17] and ~0.2 eV [16] for Au–n-GaN structures; and ~0.06 eV [18] for Cu–n-GaN structures.

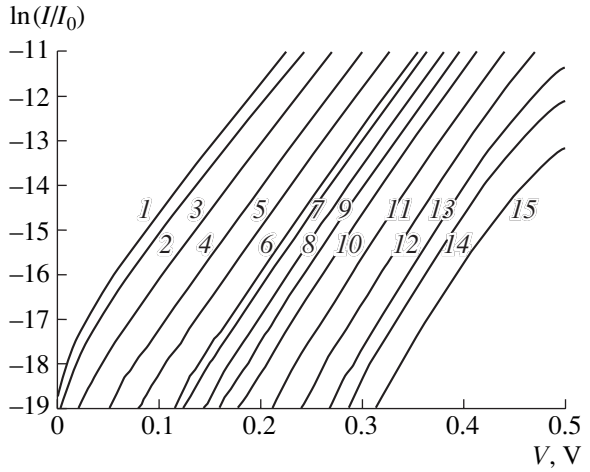

Fig. 3. Dependence of the forward current I on voltage V for the Ni–n-GaN surface-barrier structure at the temperatures T = (*1*) 411, (*2*) 400, (*3*) 380, (*4*) 360, (*5*) 342, (*6*) 320, (*7*) 310, (*8*) 299, (*9*) 288, (*10*) 275, (*11*) 250, (*12*) 224, (*13*) 202, (*14*) 183, and (*15*) 164 K. I_0 = 1 A.

3.2. The dependence of the forward current on voltage for the Ni–n-GaN structures is shown in Fig. 3. Let us consider the current flow mechanism.

Depending on the semiconductor electron concentration and temperature, three basic mechanisms of current flow can be distinguished in surface-barrier structures [19, 20]. At high temperatures ($kT \gg E_{00}$), the

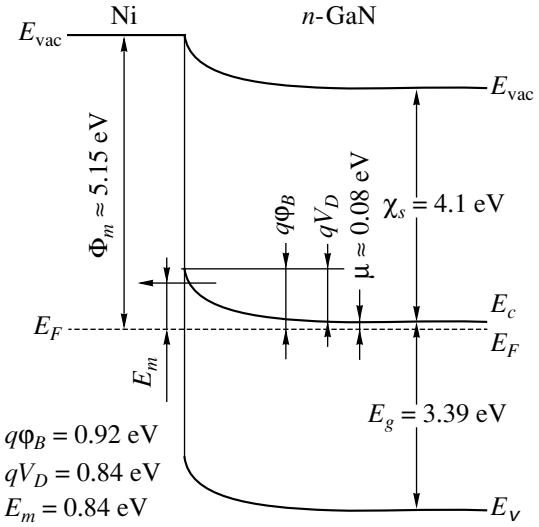

Fig. 4. An energy-band diagram of the Ni–*n*-GaN surface-barrier structure ($N_d \approx 10^{17}$ cm^{-3}) showing the level E_m of thermofield emission from the semiconductor into the metal. E_{vac} is the level of vacuum, Φ_m is the electron work function for metal, χ_s is the electron affinity of the semiconductor, E_c is the conduction-band bottom of the semiconductor, and E_v is the valence-band top of the semiconductor. T = 300 K.

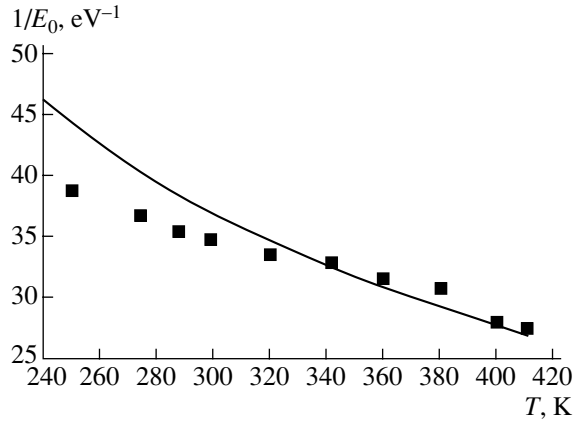

Fig. 5. Theoretical [10] (solid line) and experimental (squares) temperature dependences of the parameter $1/E_0$.

basic mechanism is thermionic emission, in which case electrons move above the potential barrier; at intermediate temperatures ($kT \approx E_{00}$), Padovani–Stratton thermofield emission prevails; and, at low temperatures ($kT \ll E_{00}$), field (tunnel) emission is dominant. Here, E_{00} is the Padovani–Stratton parameter given by

$$E_{00} = \frac{\hbar}{2}\sqrt{\frac{N_d}{\varepsilon_s m^*}} = 18.5 \times 10^{-15}\sqrt{\frac{N_d}{\varepsilon_{sr} m_r}},$$

where $m^* = m_r m_0$ is the electron effective mass in a semiconductor ($m_r = 0.2$ for GaN [13] and m_0 is the free electron mass); $\varepsilon_s = \varepsilon_{sr}\varepsilon_0$ is the semiconductor permittivity ($\varepsilon_{sr} = 8.9$ for GaN [13] and ε_0 is the permittivity of free space); N_d is the ionized donor concentration in the semiconductor, and $\hbar$ is the Dirac constant. For the structures under study ($N_d \approx 10^{17}$ cm^{-3}), the parameter value is $E_{00} = 0.0044$ eV.

According to [19], thermofield emission, i.e., electron transfer from a semiconductor into a metal through the barrier at certain energy E_m above the Fermi level (Fig. 4), takes place in a certain temperature range below and above which the current should have a tunneling and thermionic origin, respectively. For the diffusion potential differences (0.78–0.88 eV) obtained from the capacitance cutoff voltage V_0^C and the Fermi level energies E_F in the initial semiconductor with respect to the conduction-band bottom E_c, calculated for the electron concentration ~10^{17} cm^{-3} (0.06–0.12 eV), this temperature range should be 200–300 K.

According to [10], the voltage dependence of the thermofield forward current is given by

$$I = I_s \exp\left(\frac{qV}{E_0}\right),$$

where $E_0 = E_{00}\coth(E_{00}/kT)$, and the saturation current I_s should depend on temperature as follows:

$$I_s = \frac{AST\sqrt{\pi E_{00}(q\varphi_B - qV + \mu)}}{k\cosh(E_{00}/kT)}\exp\left(\frac{\mu}{kT} - \frac{q\varphi_B + \mu}{E_0}\right).$$

Here, $A = 4\pi q m_r k^2/\hbar^3$ is a Richardson constant equal to $120 m_r$ A cm^{-2} K^{-2} and S is the structure area.

An analysis of the formula for the thermofield dependence shows that

(i) the voltage dependence of the forward current should be exponential;

(ii) at each temperature, the slope of this dependence on the semilog scale should be equal to $1/E_0$, and this quantity at a given temperature depends on intrinsic semiconductor parameters rather than on barrier properties;

(iii) the cutoff on the vertical axis, obtained by extrapolation of the linear dependence $I(V)$ to $V = 0$ on the semilog scale, should yield the saturation current I_s, and the dependence $I_s\cosh(E_{00}/kT)/T$ on $1/E_0$ on the semilog scale should be linear and have a slope corresponding to the height of the metal–semiconductor potential barrier.

The voltage dependence of the forward current for the Ni–*n*-GaN structures was found to be exponential at all temperatures (250–410 K). The slope of this dependence on the semilog scale decreases with temperature in the range 250–410 K (Fig. 5), and it is close to the theoretical value of $1/E_0$.

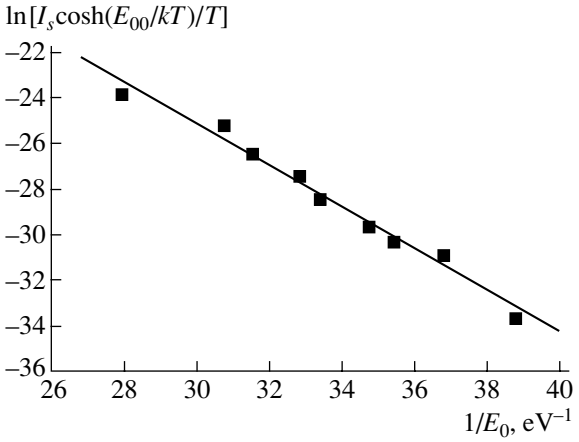

Fig. 6. Dependence of $\ln[I_s \cosh(E_{00}/kT)/T]$ on $1/E_0$ on the semilog scale for determining the potential barrier and Richardson constant; the dependence was plotted on the basis of experimental data. The current and temperature units are A and K.

The saturation current was determined from the *I–V* characteristics for each temperature. The dependence of $I_s \cosh(E_{00}/kT)/T$ on $1/E_0$ on the semilog scale is linear in the temperature range 250–400 K (Fig. 6). The potential-barrier height $q\varphi_B$ determined from the slope of this dependence is 0.9 eV. The value of $q\varphi_B$ determined from the capacitance measurements is in the range 0.90–0.94 eV (400–250 K). Since $q\varphi_B$ only slightly varies with temperature (Fig. 2), the linear temperature dependence of $q\varphi_B$ can be assumed, and its value at $T = 0$ K will be 0.95 eV.

Thus, the values of the potential-barrier height, determined from the *C–V* and *I–V* characteristics, can be considered as close to each other.

The value of the Richardson constant was estimated for GaN from the cutoff on the vertical axes, which was obtained by extrapolation of the dependence $I_s \cosh(E_{00}/kT)/T$ on $1/E_0$ on the semilog scale (Fig. 6) to $1/E_0 = 0$. This value appeared to be within the range 1–10 A cm^{-2} K^{-2}, which can be considered as close to the theoretical value $A = 24$ A cm^{-2} K^{-2}.

Thus, the forward current-flow mechanism in the Ni–*n*-GaN surface-barrier structures in the temperature range of 250–410 K corresponds to thermofield emission.

3.3. For forward current, thermofield emission implies that the maximum energy distribution of the electrons emitted from a semiconductor into a metal corresponds to a certain energy E_m that is higher than the Fermi level energy μ in the semiconductor and lower than the potential-barrier height $q\varphi_B$.

According to [19], this value, when measured from the conduction-band bottom of the semiconductor, is $qV_D\sqrt{\cosh(qE_{00}/kT)}$ = 0.74–0.76 eV for the temperatures 250–400 K.

The maximum energy distribution of emitted electrons E_m at 300 K corresponds to an energy that is 0.84 eV above the Fermi level. The potential-barrier height determined from the capacitance measurements is ~0.1-eV higher. Apparently, this circumstance explains why, in a large number of studies, the potential-barrier height determined from the capacitance characteristics exceeds that determined from the *I–V* characteristics if they are analyzed on the basis of formulas for thermionic emission.

4. CONCLUSION

The voltage and temperature dependences of the forward current and capacitance for Ni–*n*-GaN surface-barrier structures with a low dislocation density in their initial material were measured. A comparison of the experimental data with the Padovani–Stratton thermofield emission theory suggests that the forward current in the temperature range 250–410 K is controlled by thermofield electron emission from the semiconductor into the metal. This means that electrons overcome the surface barrier due to tunneling from a level located ~0.1-eV below the barrier tops. In this case, the Richardson constant is of the same order of magnitude as the theoretical value, and the potential barrier height determined from the temperature and voltage dependences of the forward current is close to the barrier height determined from capacitance measurements.

REFERENCES

1. Q. Chen, J. W. Yang, A. Osinsky, *et al.*, Appl. Phys. Lett. **70**, 2277 (1997).
2. Necmi Biyikli, Tolga Kartaloglu, Orhan Aytur, *et al.*, Appl. Phys. Lett. **79**, 2838 (2001).
3. Ching-Wu Wang, Appl. Phys. Lett. **80**, 1568 (2002).
4. E. V. Kalinina, N. I. Kuznetsov, A. I. Babanin, *et al.*, Diamond Relat. Mater. **6**, 1528 (1997).
5. Jong Kyu Kim and Jong-Lam Lee, J. Electrochem. Soc. **151**, G190 (2004).
6. J. D. Guo, M. S. Feng, R. J. Guo, *et al.*, Appl. Phys. Lett. **67**, 2657 (1995).
7. E. J. Miller, E. T. Yu, P. Waltereit, and J. S. Speck, Appl. Phys. Lett. **84**, 535 (2004).
8. Jihyun Kim, F. Ren, A. G. Baca, and S. J. Pearton, Appl. Phys. Lett. **82**, 3263 (2003).
9. E. J. Miller, D. M. Schaadt, E. T. Yu, *et al.*, J. Appl. Phys. **94**, 7611 (2003).
10. F. A. Padovani and R. Stratton, Solid-State Electron. **9**, 695 (1966).
11. N. M. Shmidt, W. V. Lundin, A. V. Sakharov, *et al.*, Proc. SPIE **4340**, 92 (2000).
12. A. V. Ankudinov, A. I. Besyulkin, A. G. Kolmakov, *et al.*, Physica B (Amsterdam) **340–342**, 462 (2003).

13. *Properties of Advanced Semiconductor Materials*, Ed. by M. Levinshtein, S. Rumyantsev, and M. Shur (Wiley, New York, 2001).

14. Q. Z. Liu, L. S. Yu, F. Deng, *et al.*, J. Appl. Phys. **84**, 881 (1998).

15. T. Mori, T. Kozawa, T. Ohwaki, *et al.*, Appl. Phys. Lett. **69**, 3537 (1996).

16. M. Sawada, T. Sawada, Y. Yanagata, *et al.*, in *Proceedings of Second International Conference on Nitride Semiconductors* (Tokushino, Japan, 1997), p. 706.

17. P. Hacke, T. Detchprohm, K. Hiramatsu, and N. Sawaki, Appl. Phys. Lett. **63**, 2676 (1993).

18. Wei-Chih Lai, Meiso Yokoyama, Chun-Yung Chang, *et al.*, in *MRS Spring Meeting: Abstracts of Symposium on Y:Wide-Bandgap Semiconductors for High-Power, High-Frequency, High-Temperature Applications*, Ed. by S. Binari, A. Burk, M. Melloch, and C. Nguyen (San Francisco, Calif., 1999), Y5.8.

19. E. H. Rhoderick, *Metal–Semiconductor Contacts* (Clarendon, Oxford, 1978; Radio i Svyaz', Moscow, 1982).

20. T. V. Blank and Yu. A. Gol'dberg, Fiz. Tekh. Poluprovodn. (St. Petersburg) **37**, 1025 (2003) [Semiconductors **37**, 999 (2003)].

Translated by A. Kazantsev

Semiconductors, Vol. 39, No. 6, 2005, pp. 679–684. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 710–715.

LOW-DIMENSIONAL SYSTEMS

The Effect of Adsorbed Molecules on the Charge-Carrier Spectrum in a Semiconductor Nanowire

V. A. Lykakh[1]^ and E. S. Syrkin[1, 2]^^

[1]*Kharkov Polytechnical Institute (National Technical University), Kharkov, 61002 Ukraine*

^e-mail: lykah@ilt.kharkov.ua, lukah@kpi.kharkov.ua

[2]*Institute for Low-Temperature Physics and Engineering, Kharkov, 61103 Ukraine*

^^e-mail: syrkin@ilt.kharkov.ua

Submitted June 30, 2004; accepted for publication October 8, 2004

Abstract—A semiconductor quantum nanowire with adsorbed organic molecules is considered. It is shown that a shift of the quantum-confinement levels in the wire includes both a linear contribution (determined by the orientation of the molecular dipoles and the sign of the charge carrier) and a nonlinear contribution (determined by the deformation of the molecular layer). In the case of a long nanowire, longitudinal quantization of the charge carriers is described self-consistently by a nonlinear Schrödinger equation with boundary conditions. For all values of the nonlinear-interaction parameter, the spectrum is determined by a set of transcendental equations. It is shown that the role of nonlinear interaction is greater for lower energy levels and increases with an increase in the mass of the charge carriers and decrease in the rigidity of the molecular layer. Carrier localization, which manifests itself in the experiment as an increase of the resistance, is possible. The processes considered may be important in relation to chemisorption sensors, chips based on nanotubes and DNA, and other structures with adsorbed organic layers. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The creation of new types of mesoscopic objects and the prospect of their application in nanoelectronics stimulate interest in studies of the fundamental properties of such objects. The optical properties and conductance of a nanoobject are determined by the set of quantum energy levels of its charge carriers [1]. The effect of the energy-level structure on conductance has been observed for metal nanowires [2] and nanotubes [3]. A quantum nanowire can be a conventional intrinsic semiconductor or a carbon nanotube in which the mean free path of the charge carriers exceeds 10 μm [4], which is an important factor enabling quantization along the nanowire axis [5]. The physical properties of a nanowire are affected considerably by the medium that is in contact with it. For example, nanowire conductance is extremely sensitive to the presence of adsorbed layers of NH_3 molecules [6] and more complex molecules forming Langmuire–Blodgett films [7]. The potential for creating chips based on nanotubes and DNA [8] or surface self-organizing layered organic structures [9] have been discussed.

In this paper, we would like to call attention to the fact that the interaction of a nanowire with a soft medium composed of organic molecules leads to a modification of the charge-carrier energy spectrum, which is a fundamental characteristic of a nanoconductor. It appears that the spectrum is extremely sensitive to the state of the molecular subsystem. We consider the effect of the interaction of the uncompensated charge carried by an electron or hole in a quantum nanowire with the neighboring medium, which has low mechanical rigidity and consists of molecules possessing an intrinsic electric-dipole moment, and derive nonlinear nonlocal equations describing such a system. Taking into account the deformation of the molecular layer, we calculate the linear and nonlinear contributions to a shift of the charge-carrier energy levels, which depends on the polarization of the molecules. For the case of a long nanowire surrounded by a thin molecular layer, the problem of longitudinal quantization is reduced to solving the spectral problem for a nonlinear Schrödinger equation. In the solution obtained, the normalization of the wave function manifests itself tangibly, which is a radical difference from the known solution for nonlinear classical oscillations in a finite chain [10]. Calculation of the nonlinearity parameter and the energy of a given quantum level of a charge carrier, expressed via the parameter of the nonlinear interaction of the carrier with the molecules, is reduced to solving a set of two transcendental equations. The analysis indicates that localization of the charge-carrier motion along the nanowire is possible. Physically, the situation considered represents one of the manifestations of the polaron effect; however, there are specific features related to the confinement of the charge carriers and the one-dimensional character of their motion.

1063-7826/05/3906-0679$26.00 © 2005 Pleiades Publishing, Inc.

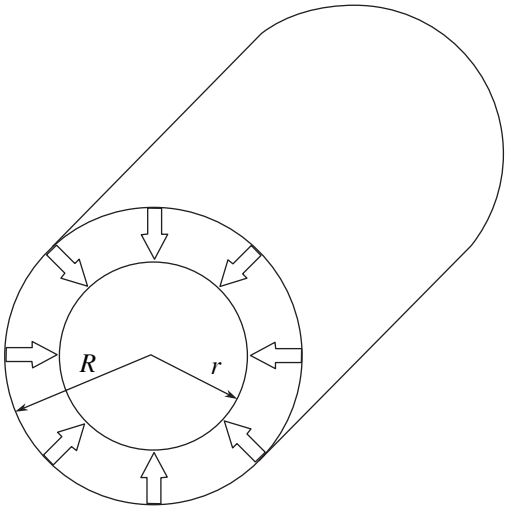

Fig. 1. A cylindrical quantum nanowire with adsorbed polar molecules. The arrows indicate the direction of the electric-polarization vector.

2. MODEL OF THE SYSTEM

The time-independent Schrödinger equation for an extra carrier in an intrinsic-semiconductor nanowire can be written as [1, 11]

$$-\frac{\hbar^2}{2m_{\text{eff}}}\Delta\psi + U(\mathbf{r})\psi = W\psi. \qquad (1)$$

Here, $\psi \equiv \psi(\mathbf{r})$ is the wave function, m_{eff} is the charge-carrier effective mass, Δ is the Laplace operator, W is the total energy, $U(\mathbf{r})$ is the potential energy, and $\mathbf{r}(x, y, z)$ is the radius vector of the particle. A deep potential well can be approximated by a well of infinite depth: $U(\mathbf{r}) \equiv U_0(\mathbf{r})$, where $U_0(\mathbf{r}) = 0$ inside the semiconductor and $U_0(\mathbf{r}) = \infty$ outside the semiconductor. If the interaction with the surrounding medium $U_{\text{int}}(\mathbf{r})$ is taken into account, the potential is given by

$$U(\mathbf{r}) = U_0(\mathbf{r}) + U_{\text{int}}(\mathbf{r}). \qquad (2)$$

This interaction can be important if molecules possessing an intrinsic dipole moment $\mathbf{d}$ are adsorbed at the surface of the nanowire. Summing the contributions from all the dipoles, we obtain the interaction potential energy $U_{\text{int}} = e\phi$ at a point $\mathbf{r}_0$ within the nanowire:

$$U^e_{\text{int}}(\mathbf{r}_0) = e\int d\mathbf{r}' n(\mathbf{r}')\mathbf{d}(\mathbf{r}')\frac{\mathbf{r}_0 - \mathbf{r}'}{\varepsilon|\mathbf{r}_0 - \mathbf{r}'|^3}. \qquad (3)$$

Here, $n(\mathbf{r}')$ is the volume number density of the molecules, e is the carrier charge, $\mathbf{d}(\mathbf{r}')$ is the mean value of a dipole moment at the point $\mathbf{r}'$, and ε is the relative permittivity of the medium. The interaction potential can be also written in terms of the electric-polarization vector $\mathbf{P}$ [12] if we substitute $n(\mathbf{r}')\mathbf{d}(\mathbf{r}')$ by $\mathbf{P}(\mathbf{r}')$ in (3). In turn, the strength of the electric field $\mathbf{E}(\mathbf{r}')$ created by a charge carrier determines the potential energy of the interaction between this carrier and an individual molecular dipole:

$$U^d_{\text{int}}(\mathbf{r}') = -\mathbf{d}(\mathbf{r}')\int d\mathbf{r}_0 \frac{e|\psi(\mathbf{r}_0)|^2(\mathbf{r}' - \mathbf{r}_0)}{\varepsilon|\mathbf{r}' - \mathbf{r}_0|^3}. \qquad (4)$$

Set (1)–(4) is completed with the material equations

$$n(\mathbf{r}') = n(\mathbf{E}(\mathbf{r}')); \quad \mathbf{d}(\mathbf{r}') = \mathbf{d}(\mathbf{E}(\mathbf{r}')). \qquad (5)$$

The interaction described by (3) and (4) is nonlocal and depends substantially on the configuration of the system. Schrödinger equation (1) with potentials (2)–(4) can be transformed into a nonlinear integro-differential equation. Such equations can be solved only by using approximation methods [11]. It is reasonable to assume that all of the dimensions of the nanowire and the adsorbed layer are much smaller than the nanowire length, i.e., $r, R \ll 2L$; this condition is satisfied under experimental conditions (see the photograph in [7]). Let us consider a system with cylindrical geometry (Fig. 1), in which, under the assumption made, the separation of the coordinate variables is possible: $\psi(\mathbf{r}) = \psi(x)\psi_\perp(y, z)$ and $W = W_x + W_\perp$ (here, x is the coordinate along the wire axis). An enhancement of the carrier tunneling into the region occupied by adsorbed molecules, as compared to the case of a vacuum surrounding, may result in a modification of $\psi_\perp(y, z)$ and $W_\perp$. We also assume that any variations in $\psi(x)$ occur on a length scale on the order of L. The type of configuration where the predicted effects are most pronounced can be described as follows. The effect of the charge-carrier field is strongest if the molecules possess an intrinsic electric-dipole moment $\mathbf{d}$ and the molecular system is soft. However, calculations can be carried out more readily for a molecular system where only some of the degrees of freedom are soft, for example, in the case of layered smectic A liquid crystals [13]. We assume that the molecular system is rigid in the direction over the nanowire surface (a close-packed layer of long linear molecules) and soft in the radial direction (elastic molecules or elastic coupling between the layers) [13, Chapter 10], so that the molecule axes are oriented normally to the layer [13, Chapter 5; 9]. The electric-dipole moment in these molecules exists due to the presence of atomic groups that break the charge symmetry [14; 13, Chapter 10; 9]. For simplicity, we assume that the dipole moment of a molecule is oriented along its axis. Ferroelectric and antiferroelectric ordering and disordered phases are possible [14], and a nonzero polarization can appear due to the flexoelectric effect [13, Chapter 6].

In the approximation of a long nanowire, integral contributions can be reduced to those of a local nature. Let us express integral (3) in cylindrical coordinates. The integration limits r and R coincide with the outer radii of the nanowire and the molecular layer (see Fig. 1). In the absence of charge carriers or in the case of a rigid molecular system, the density in (3)–(5) is constant ($n(\mathbf{r}') = n_0$) and, along with $\mathbf{d}$, can be kept out of the

integration sign. Integration with respect to x (the coordinate along the nanowire axis) is reduced to an integration within infinite limits, even for $x_0 - x' \geq 3R$. Using [15], we write the potential energy of the interaction of a charge carrier with the dipole subsystem as

$$U_{\text{int}}^{e0} = -4\pi dne(R - r). \quad (6)$$

Next, we obtain an approximate analytical expression for the potential energy of the interaction between a charge carrier and an individual molecular dipole. Instead of carrying out a spatial integration in (4), we calculate the fluxes, thus making it unnecessary to specify the shape of the radial distribution; in other words, we replace the local value of the radial component of the field strength with a value calculated for an infinitely long wire taking the local value of the wave function:

$$U_{\text{int}}^{d}(x) = -\mathbf{Ed} = -\frac{2\tau d}{\varepsilon r'}, \quad (7)$$

$$\tau(x) = e|\psi(x)|^2 F_\perp, \quad F_\perp = \int |\psi(y, z)|^2 dydz. \quad (8)$$

Here, $\tau(x)$ is the local linear charge density. The fact that the positive direction of a dipole moment **d** coincides with the direction towards the center of the nanowire, as indicated by (3), is taken into account. The error resulting from such an approximation can be found using the Ostrogradskiĭ–Gauss theorem if we determine the "leakage" of the electric-field flux through the bases of a cylinder coaxial with the nanowire. This error can be estimated as a product of the base area $S \propto R^2$ and the axial component of the electric-field strength $E_x \propto L^{-2}$; thus, the disregarded contribution to the flux is on the order of $(R/L)^2 \ll 1$, while the flux through the side surface of the cylinder is ~1.

In order to obtain material equation (5) for a thin molecular layer, we combine the condition $\delta = 2\tau d/k\varepsilon r'^2$ for the equilibrium elastic displacement of an individual dipole pulled into (or pushed out of) the region of a stronger field $F_i = -k\delta = -\nabla U_{\text{int}}^{d}$ and the condition for the conservation of the number of molecules

$$\pi(R_0^2 - r^2)n_0 = \pi(R_1^2 - r^2)n_1. \quad (9)$$

Here, subscripts 0 and 1 correspond to the cases without and with a charge carrier, respectively, and $R_1 = R_0 + \delta$. Substituting the expression for $(R_1 - r)n_1$ obtained from (9) into Eq. (6) in place of $n(R - r)$ and expanding it with respect to $\delta/R_0 \ll 1$, we obtain the potential energy of a charge carrier moving in a self-consistent field of elastically displaced dipoles:

$$U_{\text{int}}^{e} = U_{\text{int}}^{e0} + U_{\text{int}}^{e1}, \quad (10)$$

$$U_{\text{int}}^{e1} = -G|\psi(x)|^2, \quad (11)$$

$$G = \frac{8\pi n_0 d^2 e^2 F_\perp}{k\varepsilon R^3}(R_0 - r). \quad (12)$$

In the case of a film consisting of several layers arranged along the cylinder surface of the nanowire, the volume free-energy density can be written in the continuum approximation [16] as

$$w = \frac{1}{2}\left[B(u_{r'})^2 + K_{11}\left(\frac{1}{r^2}u_{\phi\phi} + u_{xx}\right)^2\right] - \mathbf{PE}. \quad (13)$$

Here, $u_{r'}$, u_{xx}, and $u_{\phi\phi}$ are the spatial derivatives of the layer displacement from the equilibrium and B and K_{11} denote the elastic moduli of the liquid crystal with respect to variation in the interlayer spacing and to layer bending, respectively. The displacement along the layer is negligible [16], and $K_{11}u_{xx}^2/B \sim \lambda^2 u^2/L^4 \ll u_{r'}^2 \sim u^2/r^2$, where $\lambda \sim 50$ Å, i.e., $\lambda \sim r \ll L$ (see [16]). In an axially symmetric case, $u_{\phi\phi} = 0$. Thus, below, we disregard the term with K_{11}. Let us follow the reasoning used in the derivation of the barometric height formula: the pressure change that results from incrementing the layer radius by $\Delta r'$ equals $\Delta p = -\nabla(EP)\Delta r'$. Using the expressions for E and P and the relationship between the uniaxial pressure and the density $n = n_0(1 + p/B)$, we perform an integration and obtain

$$n(r') = n_0 \exp[r_d(r'^{-1} - R_1^{-1})], \quad (14)$$

$$r_d = 2\tau n_0 d/\varepsilon B, \quad (15)$$

where R_1 is the outer radius of the molecular layer deformed by the electric field of the charge carriers and $|r_d|$ is the effective radius of the deformed layer. Here, we use the boundary condition $n(R_1) = n_0$, which implies that the pressure at the outer surface vanishes. To obtain further analytical estimates, we consider the case of a small effective radius ($|r_d| \ll r, R_1$). Then,

$$n(r') \approx n_0[1 - r_d(r'^{-1} - R_1^{-1})]. \quad (16)$$

After inserting (16) into (3) and performing an integration, we substitute R_1 found from the condition of the conservation of the number of molecules (similar to (9)) and expand the resulting expression for the potential up to linear terms in r_d to obtain

$$G = -\frac{8\pi(n_0 de)^2}{\varepsilon B}F_\perp\left[2 - \frac{r}{R_0} - \left(\frac{r}{R_0}\right)^2 + \ln\frac{R_0}{r}\right]. \quad (17)$$

Then, Eq. (1) can be rewritten to describe the one-dimensional motion of the charge carriers in the nanowire as follows:

$$-\frac{\hbar^2}{2m_{\text{eff}}}\frac{\partial^2\psi(x)}{\partial x^2} + [U_{\text{int}}^{e0} - G|\psi(x)|^2]\psi(x) = W_x\psi(x). \quad (18)$$

Here, $-L < x < L$. U_{int}^{e0}, given by (6), and G, given by (12) or (17), determine the parameters of the linear and nonlinear interaction of the charge carriers with the elastic molecular subsystem. The sign of the linear interaction parameter depends on the sign of the charge and

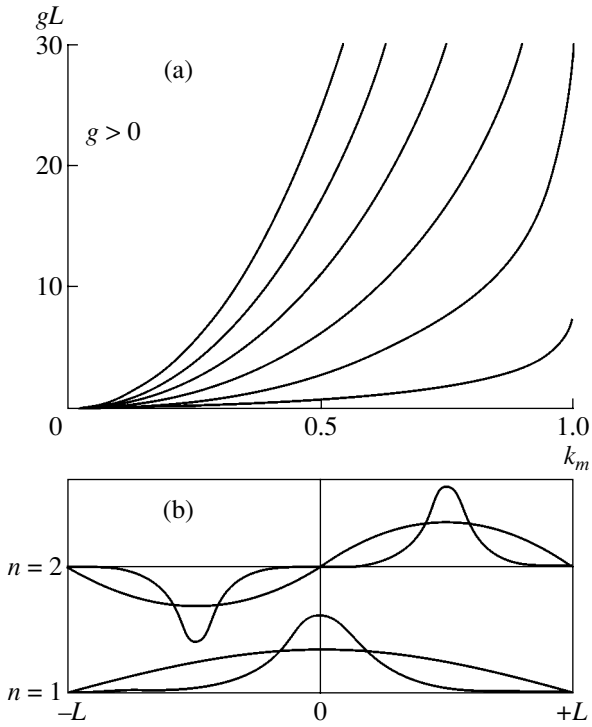

Fig. 2. A graphical representation of Eq. (27). (a) The relationship between the dimensionless interaction parameter gL and the elliptic modulus $k_m^{(n)}$. The plotted curves correspond to n = 1–6 (from bottom to top). (b) The shape of the charge-carrier wave functions for the first and second levels according to (21). The sinusoids correspond to the elliptic modulus $k_m = 0$ ($g = 0$), and the bell-shaped curves correspond to $k_m = 0.99$ (nonlinear interaction with elastic molecules is present).

the orientation of the dipoles. The nonlinear interaction always leads to a decrease in the energy of the system, due to the elastic response of the molecular layer; thus, the parameter G is even-numbered in d and e.

3. ANALYSIS OF THE ENERGY SPECTRUM

Let us introduce the following variables, which are conventional in quantum-mechanical problems [11]:

$$k_p^2 = (W_x - U_{\text{int}}^{e0})\frac{2m_{\text{eff}}}{\hbar^2}, \quad 2g = G\frac{2m_{\text{eff}}}{\hbar^2}; \qquad (19)$$

here, k_p is the component of the particle wave vector along the quantum conductor and g is the renormalized interaction parameter. Then, Eq. (18) assumes the form

$$\psi''(x) + k_p^2\psi(x) + 2g|\psi(x)|^2\psi(x) = 0. \qquad (20)$$

A first-order integral of this equation exists that makes it possible to separate the variables. Further integration, carried out using the properties of elliptic functions [17], yields even- and odd-numbered solutions to the Schrödinger equation:

$$\psi_{\text{even}} = b\,\text{cn}(\kappa x), \quad \psi_{\text{odd}} = k_m' b\,\text{sd}(\kappa x). \qquad (21)$$

Here, $k_m' = \sqrt{1 - k_m^2}$ and $\text{sd}(y) = \text{sn}(y)/\text{dn}(y)$. The amplitude b and the modulus k_m of the elliptic functions are related to the integration constant κ by the equations

$$b^2 = \frac{1}{2g}(\kappa^2 - k_p^2), \qquad (22)$$

$$k_m = b\sqrt{g}/\kappa. \qquad (23)$$

The boundary conditions $\psi(\pm L) = 0$ result in the following equation for k_m and κ:

$$MK(k_m) = \kappa L. \qquad (24)$$

Here, $M = 2m + 1$ ($m = 0, 1, 2, \ldots$) for even-numbered solutions and $M = 2m$ ($m = 1, 2, \ldots$) for odd-numbered solutions. The normalization condition $\int|\psi(x)|^2\,dx = 1$ results in the following equation for k_m and b (we use [18] and (24)):

$$\frac{2Lb^2}{k_m^2 K(k_m)}[E(k_m) - k_m'^2 K(k_m)] = 1. \qquad (25)$$

Here, $E(k_m)$ is a complete elliptic integral of the second kind. Equations (22)–(25) form a closed set with respect to the parameters b, κ, k_m, and k_p. Eliminating b and κ, we obtain the solution for k_p:

$$k_p^2 = (1 - 2k_m^2)\frac{K^2(k_m)}{L^2}n^2. \qquad (26)$$

Here, k_m is a root of the equation

$$\frac{2}{gL}K(k_m)[E(k_m) - k_m'^2 K(k_m)] = \frac{1}{n^2}. \qquad (27)$$

The parameter k_p^2 (the square of the quasi-momentum) controls the charge-carrier energy spectrum (19), which can be studied experimentally.

In order to analyze the dependence of k_p^2 on the parameters of the system, we make use of a graphical representation. Equation (27) describes the relationship between the parameter k_m, which characterizes the nonlinearity, and the parameter of nonlinear interaction gL (this relationship is shown in Fig. 2). The shape of this dependence suggests that (i) an increase in gL results in a larger nonlinearity parameter $k_m^{(n)}$ and (ii) the nonlinear interaction affects the lowest levels the most profoundly.

Let us consider the manifestation of these features in the behavior of energy spectrum (19). Figure 3a shows the dependence of the energy levels on the modulus of

the elliptic integral. For $k_m = 1/\sqrt{2}$, the energy levels cross the zero value. In Fig. 3b, the level energies are plotted as functions of the interaction parameter. The nonlinear character of the first level can clearly be seen.

In the limit of a weak interaction ($gL \longrightarrow 0$ and $k_m \longrightarrow 0$), the nonlinearity vanishes and $K(k_m) \longrightarrow \pi/2$. Then, (26) is reduced to a known solution for a rectangular potential well of infinite depth [11]: the wave vector $k_p^2 = \pi^2 n^2/4L^2$. The limit of a strong interaction can be realized even for moderate values of g, provided the nanowire is sufficiently long ($Lg \longrightarrow \infty$ and $k_m \longrightarrow 1$). Using the known asymptotic behavior of elliptic functions, [16] ($K(k_m) \longrightarrow \infty$, $E(k_m) \longrightarrow 1$ and $k_m'^2 K(k_m) \longrightarrow 0$), we find that, in this case, set of Eqs. (27) and (26) transforms into

$$K(k_m)n/L = g/2n, \quad k_p^2 = -(g/2n)^2, \qquad (28)$$

and wave function (21), taking into account (22), (23), and (28), assumes the soliton-like shape characteristic of localized states for $n = 1$:

$$\psi(x) = \frac{\sqrt{g}}{2}\frac{1}{\cosh(gx/2)}. \qquad (29)$$

This solution can be obtained directly from (20). The effect of the walls of the potential well vanishes when $1/g \ll L$: the energy becomes independent of L, and ψ is the same for any choice of the origin. The appearance of such localized states may be responsible for a reduction of the conductance in chemisorption sensors [6] and nanowires coated with Langmuir–Blodgett films [7].

For the purposes of estimation, let us use the following values for the parameters [16]: $B \sim (10^6$–$10^8)$ J/m^3, $d = el$ and $l = 2 \times 10^{-10}$ m, the volume occupied by a molecule $1/n_0 = 5 \times 5 \times 20 \times 10^{-30}$ m^3, $F_\perp \sim 1$, $\varepsilon \approx 2$, and $L \sim (1$–$10) \times 10^{-6}$ m. Then, according to (15), the effective radius of the deformation of the molecular layer $r_d \sim (10^{-8}$–$10^{-10})$ m, and, according to (19) and (17), the dimensionless interaction parameter $gL \sim \mu(1$–$10^3)$; here, $\mu = m_{\text{eff}}/m_e \sim (10^{-2}$–$10^2)$ is the ratio of the charge-carrier effective mass to the free-electron mass. Thus, the nonlinear-interaction energy may vary in a wide range, from ~10^{-4} eV for $gL = 10$ to ~1 eV for $gL \sim 10^3$. The nonlinear interaction is stronger and the extent of charge-carrier localization is greater for "softer" coatings, longer nanowires, and heavier carriers.

Localization is enhanced if the displacement of the molecules along the surface of the nanowire and rotation of the dipole groups are possible, whereas localization is reduced if the molecules do not possess an intrinsic dipole moment, this moment is compensated, or if the temperature is increased. It is also necessary to take into account the relationship between the lifetime of the excited state τ_e and the relaxation time of the molecular system τ_M. In Fig. 3b, the spectrum for $\tau_M \ll \tau_e$ is shown. If $\tau_M \gg \tau_e$ (the adiabatic approximation), an excited carrier occupies a ground-state level determined by the deformation. For $\tau_M \sim \tau_e$, the charge-carrier excitation and tunneling processes should be considered simultaneously with the molecular-system excitations.

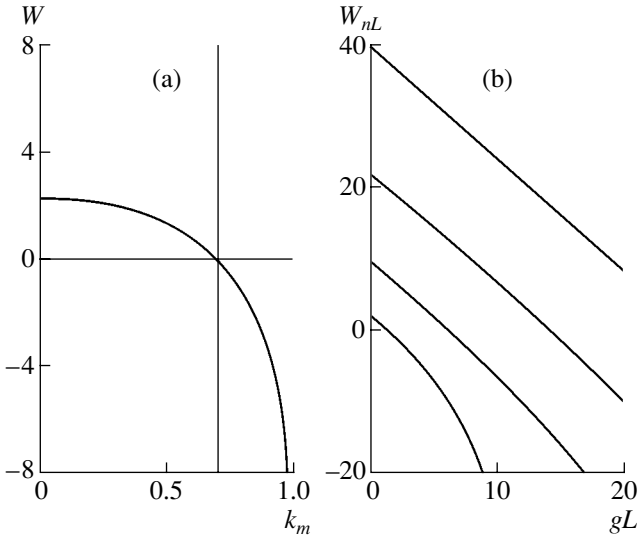

Fig. 3. (a) Renormalized level energy $W = k_p^2 L^2/n^2$ as a function of the elliptic modulus k_m, and (b) $W_{nL} = k_p^2 L^2$ as a function of the interaction parameter gL according to (26).

4. CONCLUSIONS

Thus, it is shown that the problem of calculating the longitudinal-quantization spectrum in a long nanowire coated with an adsorbed film of liquid crystal is reduced to solving a nonlinear Schrödinger equation with boundary conditions. For the interaction parameters, which vary in a wide range, the charge-carrier wave functions and energies are determined. They change from those characterizing a particle in a rectangular box (in the case of a rigid molecular layer) to those typical of a completely localized charge carrier with a soliton-like wave function (in the case of a "soft" molecular layer and a heavy carrier). The localization of the charge carriers induced by coating a nanowire with complex organic molecules may be responsible for the experimentally observed reduction in conductance [7]. Depending on the sign of the carrier charge and of the molecule polarization, the linear interaction and the corresponding shift of the levels change sign.

The charge-carrier energy spectrum depends most strongly on the rigidity of the adsorbed molecular system. Crystallization of the liquid-crystal film leads to a sharp increase in the rigidity and to a corresponding drop in the nonlinear-interaction parameter; as a result, the localization of the longitudinal motion of a carrier is disrupted and a jump in the temperature dependence of the conductance is observed. Thus, a nanowire can be used as a sensor for the state of the molecular sys-

tem. The processes considered should be taken into account in the design of chips based on nanotubes and DNA or layered organic surface structures [8, 9].

REFERENCES

1. D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge Univ. Press, Cambridge, 1997).
2. A. I. Yanson, I. K. Yanson, and J. M. van Ruitenbeek, Phys. Rev. Lett. **87**, 216805 (2001).
3. D. Orlikowski, H. Mehrez, J. Taylor, *et al.*, Phys. Rev. B **63**, 155412 (2001).
4. P. Poncharal, C. Berger, Yan Yi, *et al.*, J. Phys. Chem. B **106**, 12104 (2002).
5. C. Dekker, Phys. Today **52**, 22 (1999).
6. H. Dai, Phys. World **13** (6), 43 (2000).
7. N. P. Armitage, J.-C. P. Gabriel, and G. Gruner, cond-mat/0307712.
8. E. Buzaneva, A. Gorchynskyy, G. Popova, *et al.*, in *Frontiers of Multifunctional Nanosystems*, Ed. by E. Buzaneva and P. Scharff (Kluwer Academic, Dordrecht, 2002), NATO Adv. Study Inst. Ser., Ser. 2, Vol. 57, p. 191.
9. O. Neilands, in *Molecular Low Dimensional and Nanostructured Materials for Advanced Applications*, Ed. by A. Graja *et al.* (Kluwer Academic, Dordrecht, 2002), NATO Adv. Study Inst. Ser., Ser. 2, Vol. 59, p. 181.
10. A. S. Kovalev, Teor. Mat. Fiz. **37**, 135 (1978).
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Nauka, Moscow, 1975; Pergamon, New York, 1977).
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 2: *The Classical Theory of Fields*, 6th ed. (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
13. A. S. Sonin, *Introduction to the Physics of Liquid Crystals* (Nauka, Moscow, 1983) [in Russian].
14. R. Blinc and B. Zeks, *Soft Modes in Ferroelectrics and Antiferroelectrics* (North-Holland, Amsterdam, 1974; Mir, Moscow, 1975).
15. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986).
16. L. M. Blinov, *Electro- and Magneto-optics of Liquid Crystals* (Nauka, Moscow, 1978) [in Russian].
17. E. Jahnke, F. Emde, and F. Lösch, *Tafeln Höherer Funktionen*, 7th ed. (Teubner, Stuttgart, 1968; McGraw-Hill, New York, 1960; Nauka, Moscow, 1968).
18. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1986; Gordon and Breach, New York, 1989), Vol. 3.

Translated by M. Skorikov

Semiconductors, Vol. 39, No. 6, 2005, pp. 685–696. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 716–728.

LOW-DIMENSIONAL SYSTEMS

Local Tunneling Spectroscopy of Silicon Nanostructures

N. T. Bagraev*, A. D. Bouravlev*, L. E. Klyachkin*, A. M. Malyarenko*, W. Gehlhoff, Yu. I. Romanov***, and S. A. Rykov*****

**Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia*

***Institut für Festkörperphysik, Technische Universität, D-10623 Berlin, Deutschland*

****St. Petersburg State Technical University, Politekhnicheskaya ul. 29, St. Petersburg, 195251 Russia*

Submitted October 7, 2004; accepted for publication October 26, 2004

Abstract—The recharging of many-hole and few-electron quantum dots under the conditions of the ballistic transport of single charge carriers inside self-assembled quantum well structures on a Si (100) surface are studied using local tunneling spectroscopy at high temperatures (up to room temperature). On the basis of measurements of the tunneling current–voltage characteristics observed during the transit of single charge carriers through charged quantum dots, the modes of the Coulomb blockade, Coulomb conductivity oscillations, and electronic shell formation are identified. The tunneling current–voltage characteristics also show the effect of quantum confinement and electron–electron interaction on the characteristics of single-carrier transport through silicon quantum wires containing weakly and strongly coupled quantum dots. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The fabrication of semiconductor structures with self-assembled quantum wells (QWs) and superlattices is one of the most important problems of modern semiconductor physics and nanoelectronics. In recent years, this problem has become especially important in the context of the need to create fundamentally new nanoelectronic and optoelectronic devices, such as single-electron transistors, single-electron memory cells, and lasers that operate via intraband transitions [1–4]. Accordingly, the development of new technologies for semiconductor nanostructures raises an increasingly greater number of questions related to the microscopic nature of self-assembled QWs and the two-dimensional barriers separating them, as it is upon these that the fabrication of quantum wires (QWrs) and quantum dots (QDs) using the electrostatic confinement of charge carrier motion is based.

QDs are zero-dimensional systems that are obtained by fully confining the motion of charge carriers in QWrs, QWs, or in bulk crystals. Since the charge carriers in a QD are confined in all directions, the corresponding energy spectrum is completely discrete, just as it is for an isolated atom. For this reason, QDs are often called artificial atoms, although each QD consists of thousands, or even hundreds of thousands, of real atoms. Naturally, in this case, charged QDs are implied. Empty QDs cannot be considered as analogues of real atoms, but they are of special interest for studying charge-carrier resonant tunneling through zero-dimensional systems [5]. In turn, like a real atom, a charged QD (an artificial atom) can contain one or several free charge carriers exhibiting confinement effects and the effects of electron–electron interaction during recharging. The relative contributions of these effects are determined by the dot size and by the characteristics of its boundary [2, 6].

Like QWs and QWrs, QDs can be obtained both using molecular-beam epitaxy in combination with nanolithography and selective etching [1, 2, 6, 7] and electrostatically using deposited metal microcontacts [2, 8]. In the first case, taking AlGaAs/GaAs as an example, the formation of QDs inside the heterostructure starts from the deposition of masks onto the surface of the wide-gap semiconductor (AlGaAs). Then, the entire AlGaAs layer and part of the GaAs layer are removed by deep etching. Electrons produced by ionization of the shallow donor centers in AlGaAs concentrate in the zero-dimensional GaAs dots that appear. Therefore, the number of charge carriers localized in a QD is determined by the donor concentration and by the self-compensation resulting from the formation of DX centers [9]. This self-compensation gives rise to a lot of problems in the study of ballistic carrier transport, but the disadvantage can be eliminated by applying an electrostatic method in which one-dimensional and zero-dimensional systems are created by confining the motion of the charge carriers in QWs (Figs. 1a, 1b). The main advantage of this method is that it makes it possible to increase the number of electrons or holes in a QD to several hundred, thus allowing the observation of certain interesting phenomena such as the Coulomb blockade and Coulomb oscillations arising due to the enhancement of the electron–electron interaction [2, 8].

One of the unresolved problems of practical nanoelectronics consists in finding the relative contributions made by quantum interference and electron–electron interaction to the ballistic transport in electrostatically induced QWrs under the conditions of the elastic backscattering of charge carriers by internal δ-shaped barri-

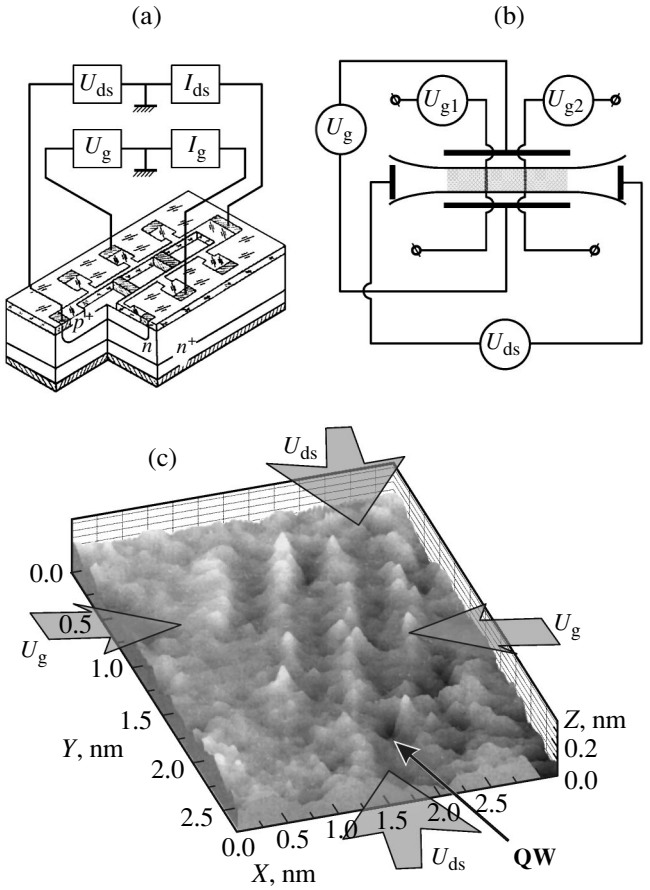

Fig. 1. (a) A planar p^+–n structure with Hall geometry (schematic diagram). The structure is designed for studying the Coulomb blockade, Coulomb oscillations, and Coulomb staircase of quantum dot conductance during the transport of individual charge carriers. (b) A schematic representation of the split-gate (U_g) used for inducing modulated QWrs inside QWs. The voltages U_{g1} and U_{g2} are applied to the finger gates intended for producing quantum dots. (c) A 3D split-gate image obtained by scanning tunneling microscopy in the vicinity of a QW on the Si (100) surface.

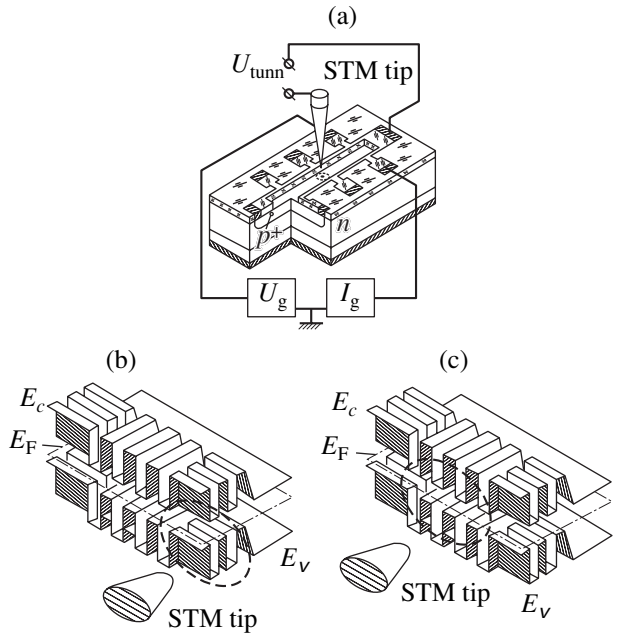

Fig. 2. (a) A planar structure containing a p^+ diffusion profile with an approaching tunneling microscope tip. (b, c) 3D band diagrams of a longitudinal *p*-type QW at the *n*-Si surface, which contains a quantum dot representing a multitunneling junction under the conditions (b) $U_{tunn} = U_{ds}$ and (c) $U_{tunn} = U_g$. The dashed circle shows the region of the point of contact of the tip.

ers. This type of modulation in the coherent transport of single charge carriers can occur as a result of the effect of residual impurity centers distributed over the boundaries of a QWr and also due to a nonuniform distribution of the gate voltage along this QWr [10, 11]. By varying the source–drain voltage (U_{ds}) and the gate voltage (U_g), we can not only enhance the effect of the random electrostatic δ barriers but also create a QD inside the QWr using the split-gate technique. The transport characteristics of the QD are controlled by finger gates, whereas the central gate voltage controls the number of charge carriers in the QD (Fig. 1). Thus, an electrostatically generated charged QD represents an artificial atom. Moreover, it is possible to change the number of charge carriers in this atom using an external electric field.

The characteristics of the QDs inside QWs can be monitored by local tunneling spectroscopy (Fig. 2a). In this case, the voltage U_{tunn} applied between the tip of the scanning tunneling microscope (STM) and the point of contact at the boundary of the planar structure, which is equal to the potential difference between the ends of an electrostatically induced quantum wire and is the sum of the source–drain voltage and the gate voltage, is

$$U_{tunn} = U_{ds} + U_g.$$

Depending on the number of charge carriers in a QD, we can distinguish between many-electron and few-electron artificial atoms. It should be noted that the number of charge carriers determines the effects that can be observed in relation to ballistic transport. For example, the Coulomb oscillations in the conductivity that appear when changing the central gate voltage (Fig. 1a) are exclusively of a many-electron character [6, 8]. Few-electron QDs, however, exhibit many properties typical of real atoms, such as orbital degeneracy and shell formation [12, 13]. In addition, the detection of the Kondo effect and Fano resonances in low-dimensional systems containing few-electron QDs provides yet more evidence of the similarity between real and artificial atoms [14, 15].

In this study, we use local tunneling spectroscopy to investigate different modes of the ballistic transport of single holes through many-electron and few-electron

QDs inside self-assembled QWs on a Si (100) surface. We focus on the detection of one-hole recharging at high temperatures (up to room temperature). The detection is carried out using a connected series of several QDs with a capacitance of ~10^{-19} F.

In the following sections, the characteristics of *p*-type silicon QWs formed between self-assembled layers of microdefects, which were obtained by a preliminary oxidation and subsequent boron diffusion on a Si (100) surface, are briefly analyzed. We present data obtained from studies of the effects of Coulomb oscillations and the Coulomb blockade. These effects arise during the single-hole recharging of a QD, which is electrostatically induced inside a QWr using a split gate built into the plane of the self-assembled Si QWs (SASQWs). Next, the local tunneling spectra, showing the effects of the Coulomb blockade and Coulomb oscillations when a single hole travels through weakly coupled many-electron QDs inside the SASQW structures (SASQWSs), are discussed. Finally, the corresponding tunneling current–voltage characteristics of strongly coupled few-electron QDs are described. These characteristics indicate various scenarios for electronic shell formation that depend on the number of electrons in a QD and on its shape.

2. SELF-ASSEMBLED SILICON QUANTUM WELLS

It is known that the formation of oxide layers on the surface of monocrystalline Si facilitates the generation of excess fluxes of intrinsic interstitial atoms and vacancies. Moreover, it is known that these fluxes have a preferential crystallographic direction along the ⟨111⟩ and ⟨100⟩ axes, respectively [16–19]. At the initial stage of oxidation, the formation of a thin oxide layer is accompanied by the generation of excess interstitial Si atoms, which can form small microdefects, whereas the fluxes of vacancies in the opposite direction result in the annihilation of these defects (Fig. 3a). Since the sources and sinks of excess interstitial atoms and vacancies on the oxidized Si(100) surface are located close to positively and negatively charged reconstructed silicon dangling bonds, respectively [19], it can be assumed that the size of the microdefects, consisting of intrinsic interstitial atoms and pyramidal in shape, is about 2 nm. Therefore, the distribution of the microdefects created at the initial stage of oxidation apparently represents a modification of a Sierpinski gasket-type fractal and includes a built-in longitudinal QW (Figs. 3b and 4a).

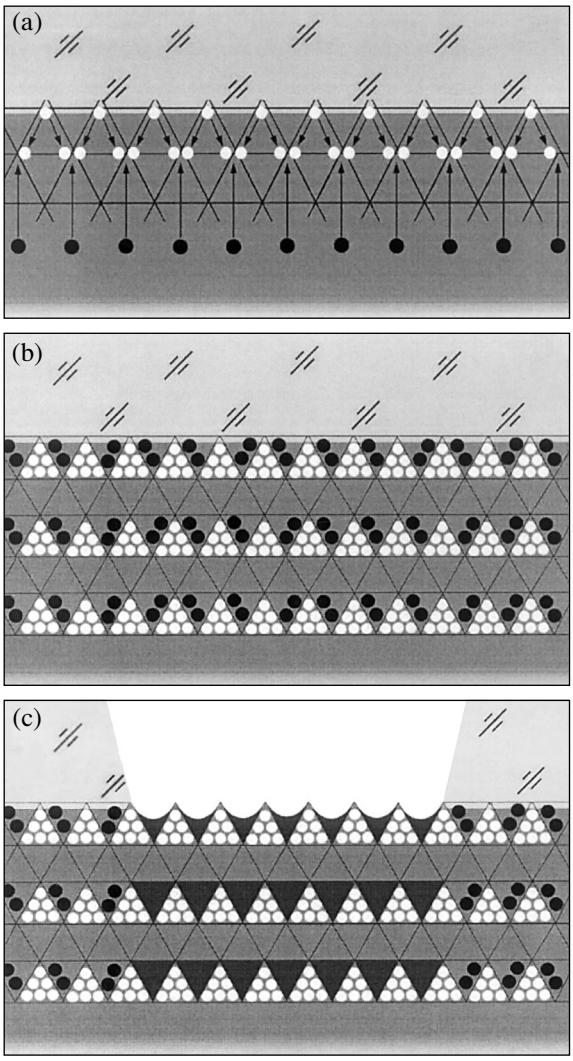

Fig. 3. Diagrams illustrating the fabrication of a self-assembled silicon quantum well structure (SASQWS) on the Si (100) surface under the conditions of injection of intrinsic interstitial Si atoms (open circles) and vacancies (filled circles) during (a, b) preliminary oxidation and (c) subsequent boron diffusion: (a) The excess fluxes of intrinsic interstitial Si atoms and vacancies that are generated during preliminary oxidation of the Si (100) surface and are crystallographically oriented along the [111] and [100] axes, respectively. (b) A diagram of a longitudinal SASQWS that appears between the layers of the microdefects formed from intrinsic interstitial Si atoms and vacancies at the stage of preliminary oxidation of the Si (100) surface. (c) A diagram of a longitudinal SASQWS formed by the subsequent passivation of the microdefects by the vacancy mechanism (dark regions) under the conditions of a short period of boron diffusion using planar silicon technology.

It should be noted that, during further oxidation of the Si (100) surface, the fractal distribution of the microdefects is reproduced and the dimensions of separate microdefects nucleated at P_b centers [20] increase [21]. However, the growth of thick oxide layers results in the predominant generation of vacancies by the oxidized surface and, hence, to the disappearance of microdefects and the self-assembly of transverse QWs [10, 21].

Although the structures of both the longitudinal and transverse silicon QWs introduced into the fractal system of microdefects are of great interest in relation to their use as a basis for optically and electrically active microcavities in optoelectronics and in relation to nano-

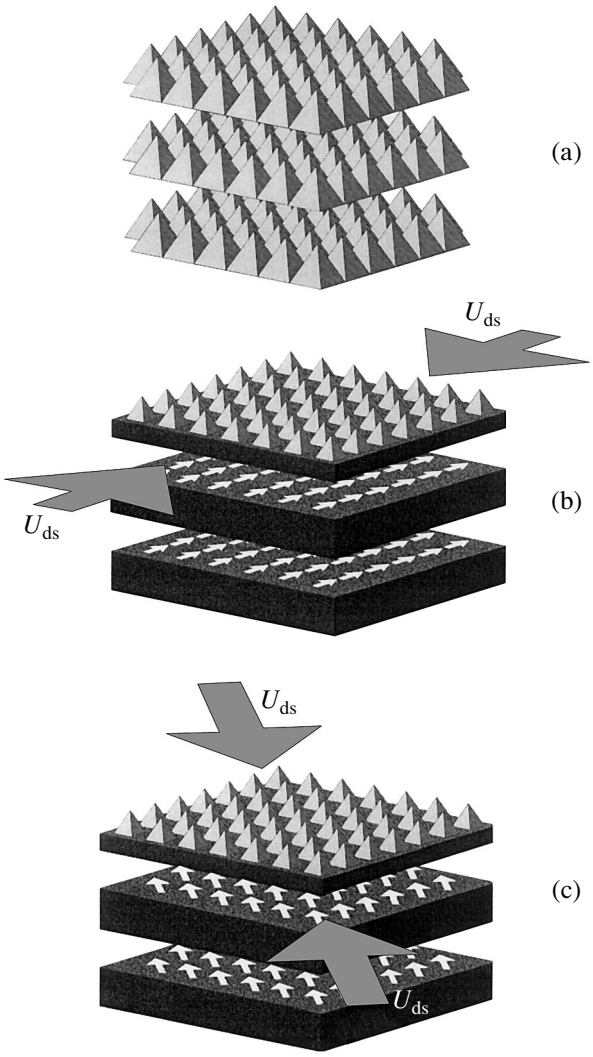

Fig. 4. A 3D image of the longitudinal SASQWSs (a) formed between the layers of microdefects, which are then (b, c) transformed into neutral δ barriers after passivation by the vacancy mechanism under a short period of boron diffusion. The white arrows show the direction of the ordering of the boron impurity dipoles inside the δ barriers under the conditions of a source–drain voltage U_{ds} applied along the crystallographic axes (b) [001] and (c) [011].

electronics, the presence of dangling bonds at their boundaries is a deleterious factor for the practical implementation of such structures. Therefore, in order to passivate dangling bonds and other defects created during the preliminary oxidation of Si (100) substrates, it is reasonable to implement a subsequent short period of boron diffusion, which can transform the layers of microdefects into neutral δ-shaped barriers bounding the SASQWSs (Figs. 3c, 4b, and 4c).

To this end, we used *n*-Si (100) substrates that were 0.35 mm thick and had resistivities of 500 and 20 Ohm cm. The substrates were preliminarily oxidized at a temperature of 1150°C in an atmosphere of dry oxygen containing CCl_4 vapors. The thickness of the oxide layers depended on the oxidation time, which was varied from 20 min to 24 h. Below, for the purposes of the fabrication and study of longitudinal SASQWSs, we mainly focus on samples with a thin oxide layer (0.22 μm), in which windows in the Hall geometry were formed photolithographically in order to perform a short period of boron diffusion from the gas phase at T_{dif} = 900°C. It is known that, at this temperature, an equilibrium between different diffusion mechanisms is established [10, 16–18]. This equilibrium results in a sharp retardation of the rate at which boron is diffused into silicon and facilitates the optimum passivation of the layers of microdefects bounding the SASQWSs (Figs. 3c, 4b, and 4c). It should be noted that the pyramidal microdefects consisting of intrinsic interstitial atoms, which become apparent at the stage of chemical etching [18], did not disappear after the short period of boron diffusion (Fig. 1c).

An earlier analysis of the resulting ultrashallow boron concentration profiles using secondary-ion mass spectrometry (SIMS) [22] has shown that their depth does not exceed 7 nm [10, 17, 18, 21]. Thus, the depth of the diffusion profiles virtually corresponds to the expected vertical size of the self-assembled silicon nanostructure consisting of longitudinal SASQWSs bounded by δ-shaped barriers (Figs. 3c, 4b, and 4c). In addition, the presence of a *p*-type QW at the *n*-Si (100) surface was verified using the four-probe method under the conditions of layer-by-layer etching. The application of SIMS to this verification encountered certain difficulties related to the limited resolution of this method and to the smoothing of the ultrashallow diffusion profile as a result of the diffusion of impurity atoms under the action of an ionic beam [17, 18].

The characteristics of the single *p*-type SAQW were determined from the angular dependence of the cyclotron resonance of electrons and holes when the magnetic field was rotated in the {110} plane perpendicular to the plane of the boron diffusion profile on the Si (100) surface [23, 24]. The quenching and shift of the cyclotron resonance lines recorded by an ESR spectrometer (*X*-band, 9.1–9.5 GHz) were found to show a 180° symmetry when the magnetic field was oriented parallel to the plane of the obtained ultrashallow boron profile, which clearly indicated that the QW was oriented parallel to the (100) planes.

The cyclotron resonance spectra (Fig. 5) consist of unusually narrow lines, indicating a substantial increase in the spin–lattice relaxation time of the nonequilibrium electrons and holes in the SASQWSs as compared to the corresponding times for the bulk Si samples. The relaxation times estimated from the width of the cyclotron resonance spectral lines for the electrons ($\tau \simeq 7 \times 10^{-10}$ s), light holes ($\tau \simeq 5 \times 10^{-10}$ s), and heavy holes ($\tau \gtrsim 5 \times 10^{-10}$ s) directly indicate a high mobility of two-dimensional charge carriers: $\mu \gtrsim 200$ m^2/(V s). This result is rather unexpected in view of the boron doping level of the δ-shaped barriers between which the

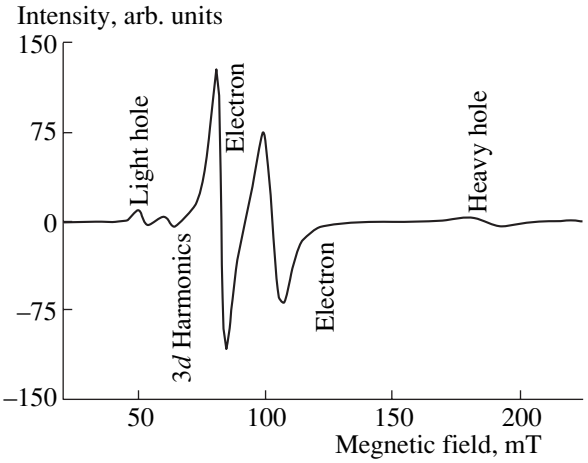

Fig. 5. The cyclotron resonance spectrum for a *p*-type SASQWS formed between the δ barriers at the *n*-Si (100) surface. The magnetic field is directed perpendicularly to the SASQWS planes in the plane {110}: $B \parallel \langle 100 \rangle + 30°$. $T = 4.0$ K and $\nu = 9.45$ GHz.

QWs are formed. In order to explain this fact, we used ESR data, which indicates that heavily doped δ-shaped barriers have ferroelectric properties, since they consist of trigonal impurity B^+–B^- dipoles [10, 25]. In this model, the reconstruction of shallow boron acceptors, resulting in the formation of neutral dipole centers ($2B^0 \longrightarrow B^- + B^+$), must be accompanied by the appearance of a correlation energy gap in the density of states of the degenerate two-dimensional hole gas. The determination of this energy gap from the temperature dependence of the Seebeck coefficient (Fig. 6) makes it possible to explain the above values of the charge carrier mobility, which are retained in SASQWSs up to 77 K [26].

The electrostatic ordering of the reconstructed impurity dipoles inside the δ-shaped barriers by an external electric field applied along different crystallographic axes in the plane of the SASQWSs (Figs. 4b, 4c) creates a transverse confinement of the charge carriers, which, it would seem, results in the angular dependence of the conductivity (Fig. 7). The conductivity maxima observed when the external electric field is oriented along the [010], [001], and [011] axes correspond to the crystallographic directions that are most energetically favorable for a reconstructed deep center under the conditions of the quadratic Stark effect [27, 28]. Therefore, the presence of δ-shaped barriers with ferroelectric properties allows, in particular, the observation of a quantum conductivity staircase using a split-gate configuration (Fig. 1) both at zero and nonzero gate voltages [10, 29]. Furthermore, the electrostatic ordering of impurity dipoles, which gives rise to the transverse confinement of charge carriers moving along the plane of the QWs, allows us to use local tunneling spectroscopy to study quasi-one-dimensional charge carrier transport. In this case, the external electric field $U_{tunn} = U_{ds} + U_g$ applied in the plane of the QWs (Fig. 2a), on the one hand, gives rise to transverse confinement due to the ordering of the impurity dipoles (U_g) and, on the other hand, is responsible for the transport of individual charge carriers (U_{ds}).

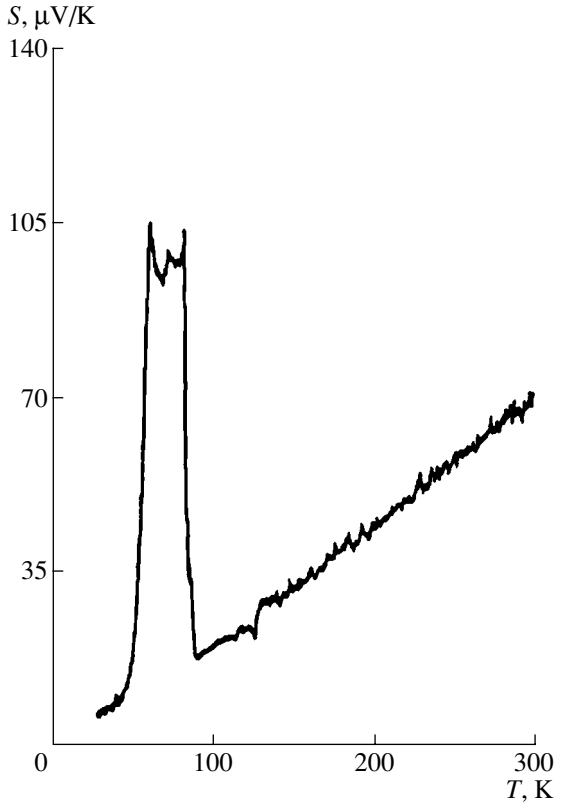

Fig. 6. The temperature dependence of the thermoelectric power (Seebeck coefficient) for a *p*-type SASQWS formed between the δ barriers at the *n*-Si (100) surface.

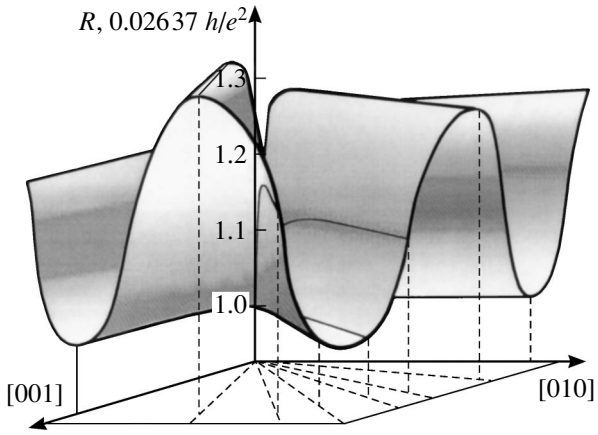

Fig. 7. The crystallographically dependent change in the resistance in the plane of a *p*-type SASQWS formed between the δ barriers at the *n*-Si (100) surface ($T = 77$ K).

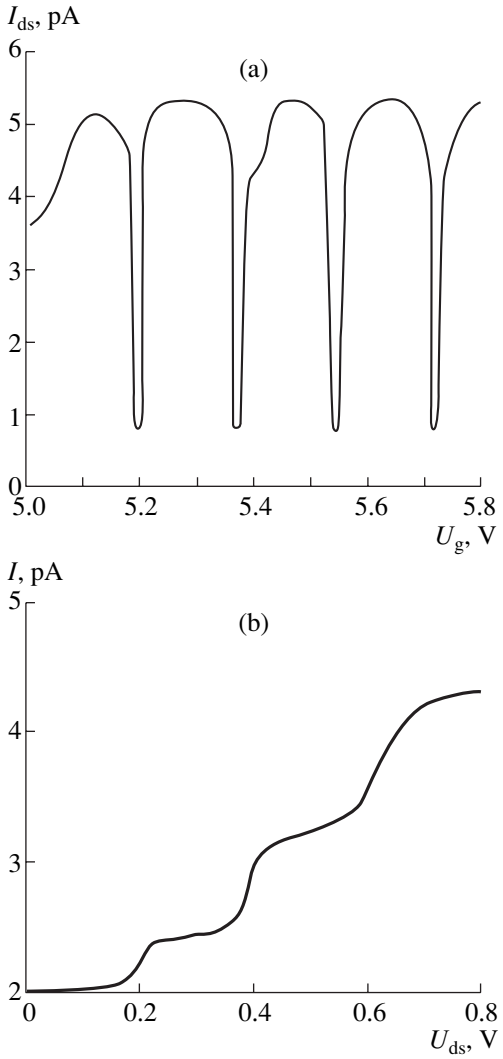

Fig. 8. (a) Coulomb oscillations and (b) the Coulomb staircase observed during the transit of single holes through a weakly coupled quantum dot inside a QWr formed using the split-gate technique in the plane of a *p*-type SASQWS formed between the δ barriers at the *n*-Si (100) surface (T = 77 K).

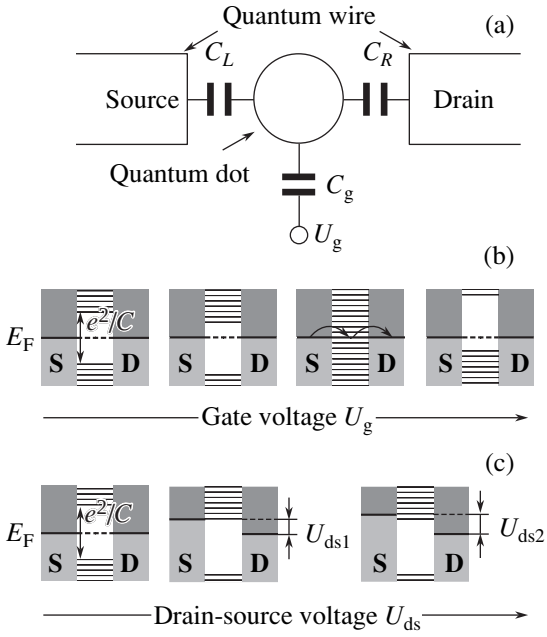

Fig. 9. (a) An equivalent circuit for a weakly coupled quantum dot inside a quantum wire with an applied gate voltage U_g. The circuit corresponds to a single charge-carrier transistor circuit. (b, c) A band diagram of a quantum dot in relation to the charge state and charge-carrier kinetic energy, which can be adjusted by changing (b) the gate voltage U_g and (c) the source–drain voltage U_{ds}.

3. MANY-ELECTRON ARTIFICIAL ATOMS

3.1. The Coulomb Blockade and Coulomb Oscillations

Figures 8a and 8b show the current–voltage (*I–V*) characteristics of a QD formed in an SASQWS using the split-gate technique (Fig. 1c). The dependence $I_{ds} = f(U_g)$ was measured at a small potential difference U_{ds} between the source and the drain, which only slightly exceeded the value required for the measurement of the tunneling conductivity between them. Nevertheless, the value of U_{ds} was sufficient for the formation of a one-dimensional channel in the SASQWS because of the electrostatic ordering of the impurity dipoles inside the δ-shaped barrier. In this case, it is not necessary to use finger gates, since the split-gate voltage is entirely localized at a QD formed near a pair of unreconstructed impurity dipoles [10]. Depending on the gate voltage, the current oscillates with a characteristic period that corresponds to the voltage required for introducing one charge carrier into the QD [6, 8]. Therefore, the Coulomb oscillations shown in Fig. 8a represent the current–voltage curve of a single-electron transistor that operates by "switching" the QD on and off each time a single charge carrier enters the QD in the process of quasi-one-dimensional transport.

The behavior of Coulomb oscillations can be considered using a model of the Coulomb blockade that arises under the conditions of weak coupling between the QD and the QWr containing it [30, 31]. This model describes the recharging mechanism of the QD during carrier tunneling through it inside the QWr; again, the QD is assumed to be electrically neutral. The energy required for introducing the charge Q into the QD is $Q^2/2C$, where

$$C = C_L + C_R + C_g$$

is the total capacitance of the QD, C_L and C_R are the capacitances between the QD and the QWr containing it, and C_g is the gate capacitance (Figs. 9a, 9b). Since the minimum charge added to the QD corresponds to the charge of a single electron (hole), it follows that, for the current to flow through it, the energy $e^2/2C$ is required. Thus, charge quantization produces an energy gap e^2/C in the spectrum of the QD states (Figs. 9b, 9c).

If the temperature is fairly low, i.e., if $kT < e^2/2C$, neither an electron nor a hole can tunnel inside the QWr containing the QD; i.e., we have a situation corresponding to the Coulomb blockade.

By varying the gate voltage U_g, we can change the energy required for adding charge to the QD (Fig. 9b). In this case, the gate voltage U_g is applied between the gate and the source. However, if the voltage between the source and the drain is very small, then the drain, the source, and the QD have the same potential. Therefore, in the presence of voltage at the gate, the electrostatic energy of the QD is changed [6, 8]:

$$E = QU_g + \frac{Q^2}{2C}. \quad (1)$$

In the case of a negative charge Q, the first term in the equation describes the attractive interaction between the charge Q and the positively charged gate, and the second term describes the electrostatic repulsion between the particles inside the QD. Equation (1) shows that the energy is at a minimum at $Q_0 = -C_gU_g$. By changing U_g, we can choose any value of Q_0 that minimizes the energy in Eq. (1) if the charge is not quantized. However, since the real charge is quantized, the variation in energy is discrete. As $Q_0 = -Ne$, the total number of charge carriers N corresponds to the minimum of the energy E, and the Coulomb interaction induces changes in the energy $e^2/2C$ as N either increases or decreases by unity. For any other value of Q_0, except for $Q_0 = -(N + 1/2)e$, there exists a smaller, but nonzero, energy level required for the addition or removal of a charge carrier. Under such circumstances, there is no current at low temperatures. However, if $Q_0 = -(N + 1/2)e$, then the state with $Q_0 = -Ne$ and the state with $Q_0 = -(N + 1)e$ are degenerate, and the charge varies between these two values even at zero temperature. Hence, the energy gap in the tunneling spectrum disappears and the current can flow (Fig. 9b). Conductivity peaks appear when the condition $C_gU_g = Q_0 = -(N + 1/2)e$ is satisfied. These peaks are periodic and are separated by a gate voltage equal to e/C_g (Fig. 8a).

Thus, the energy gap in the tunneling spectrum exists for all U_g except for the points of charge degeneracy (Figs. 8a, 9b). Closer discrete levels, shown outside this energy gap, exist due to the excited states of the charge carriers in the QD. When U_g continuously grows, the energy gap shifts downwards with respect to the Fermi energy until it reaches the point of charge degeneracy. When the gap crosses this point, the tunneling spectrum changes abruptly: the energy gap disappears and then appears again with an upward shift of e^2/C. At the same time, the charge of the QD increases by e and the process is repeated. The point of charge degeneracy and the conductivity peak are reached each time the voltage increases by e/C_g; moreover, this value is necessary for addition of one charge carrier to the QD.

It should be noted that the model of the Coulomb blockade takes the charge quantization into account but disregards the energy quantization that arises due to the small size of the QD. At low energies, the confinement makes the distances between the energy levels in the QD relatively large. Taking the confinement into account somewhat changes the pattern of Coulomb recharging of the QD [3, 8, 30]. In particular, the Coulomb oscillations become aperiodic as the number of charge carriers in the QD decreases. Therefore, in relation to tunneling through a few-electron dot, the distance between the two neighboring peaks is

$$\Delta U_g = \frac{e}{C} + \frac{\varepsilon_{N+1} - \varepsilon_N}{e}; \quad (2)$$

thus, this distance depends on the peak number, in contrast to the case of many-electron dots, for which the Coulomb oscillations are strictly periodic (Fig. 8a). To calculate the positions of the peaks of the Coulomb oscillations more accurately, we have to numerically calculate the energy eigenvalues for the real shapes of the QDs [12, 32].

The energy spectrum of a QD can also be investigated by measuring the dependence of the tunneling current on the voltage U_{ds} at fixed values of U_g (Fig. 8b), e.g., at those that satisfy the condition $Q_0 = -(N + 1/4)e$. In this case, the Fermi level of the source is raised with respect to its position in the drain, according to the increase in U_{ds}. Furthermore, the Fermi level E_F rises with respect to the confinement levels (Fig. 9c). The current begins to flow as soon as the Fermi level of the source rises above the first confinement level. When the Fermi level rises further, the higher energy levels begin to contribute to the current. The positions of the confinement levels can be determined from the conductivity curves corresponding to the Coulomb staircase (Fig. 8b) by measuring the voltages at which the current increases or, equivalently, the voltages corresponding to the peaks in the dI/dU_{ds} curve [6, 8]. At high U_{ds}, the second, third, etc. confinement levels are seen (Fig. 8b). Furthermore, the confinement levels are shifted by eU_g as the voltage U_g increases. It will be shown below that these shifts result in the corresponding asymmetry of the curves $I = f(U_{ds})$ and dI/dU_{ds}.

Thus, the Coulomb blockade mode is characterized not only by the fact that the condition $kT < e^2/2C$ is satisfied but also by the appearance of periodic Coulomb peaks that originate from the one-electron recharging of a QD. The values of the conductivity, even at the maxima of the Coulomb peaks, are much smaller than the conductivity quantum e^2/h [33–35].

3.2. Local Tunneling Spectroscopy of Many-Hole Silicon QDs in the Coulomb Blockade Mode

The technique of local tunneling spectroscopy is based on the use of a scanning tunneling microscope (STM). In this study, we used an STM with an electro-

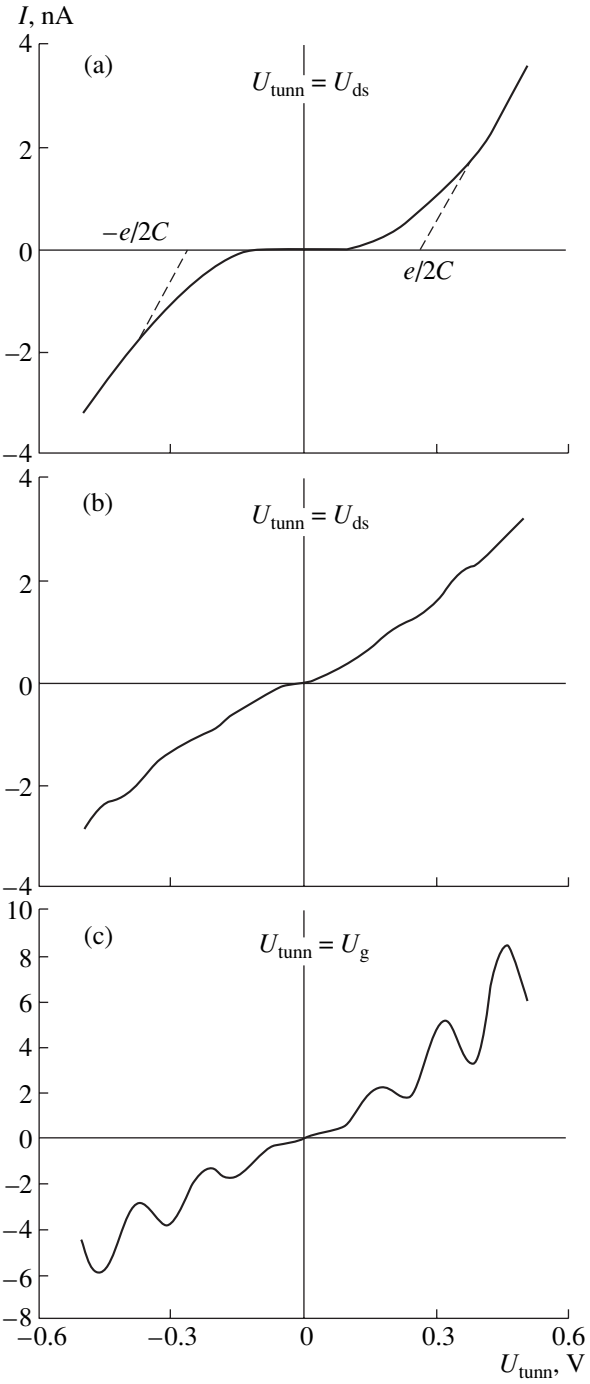

Fig. 10. Local tunneling current–voltage characteristics showing (a) the Coulomb blockade, (b) the confinement effect, and (c) Coulomb oscillations at the transit of individual holes through a weakly coupled many-hole silicon dot inside a *p*-type SASQWS at the Si (100) surface (T = 4.2 K).

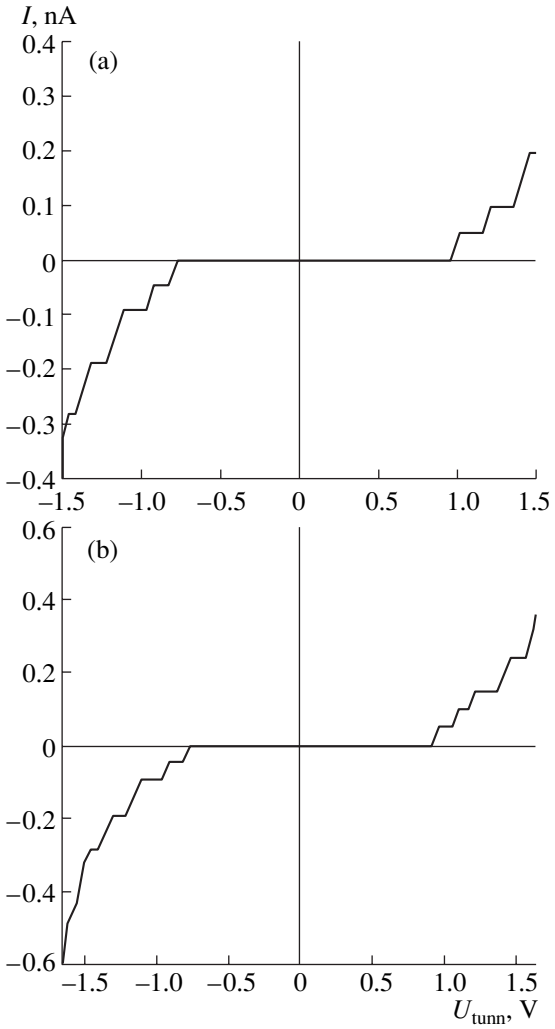

Fig. 11. Local tunneling current–voltage characteristics corresponding to the mode of the Coulomb blockade at the gate voltages U_g (a) 0.7–1 and (b) 0.6–0.9 V. The blockade arises at the transit of individual holes through a weakly coupled many-hole quantum dot inside a *p*-type SASQWS at the Si (100) surface (T = 300 K).

magnetic system characterized by a rough approach of the tip to the sample in the temperature range 4.2–300 K. Using this technique, the largest area that could be scanned at 300 K was 4×4 μm^2 and the largest area at 4.2 K was 0.6×0.6 μm^2. The STM measurements were performed using mechanically sharpened steel and tungsten tips. Local tunneling current–voltage curves $I = f(U_{tunn})$ were recorded at T = 4.2 and 300 K during study of the processes involved in the transport of individual holes in the SASQWSs described above. The tunneling current was measured as a function of the voltage applied between the STM tip and the contacts at the boundaries of the planar structure formed in the Hall geometry (Fig. 2a).

Figures 10, 11, and 12 show the tunneling current–voltage characteristics that correspond to different points of the boron diffusion profile in the Si (100) surface representing a δ-shaped barrier under which the SASQWSs are located. Although the local tunneling curves are appreciably different, all of them show the presence of Coulomb interaction and quantum confinement effects in the transport of individual holes in the

QWrs, which are induced electrostatically in some regions of the SASQWSs by voltage U_{tunn} being applied between the STM tip and the contact at the boundary of the planar structure (Fig. 2a). As was noted above, these QWrs appear due to the electrostatic ordering of impurity dipoles in the δ barriers. This ordering is induced by the STM tip as it approaches the surface of the diffusion profile. The orientation of the QWrs is determined by the relative positions of the tip and the contact at the boundary of the planar structure. The current and the shape of the local current–voltage characteristic depend on the mode of the ballistic transport of individual holes. The main characteristics of the QWr responsible for the shape of the local current–voltage curves are its smoothness, which varies due to fluctuations in the distribution of the alloying impurity, and the presence of isolated QDs. Furthermore, it is important how close these QDs are to the region where the tip directly affects the surface of the diffusion profile (Figs. 2b, 2c). This closeness determines the voltage drop across a QD, which is equivalent to the gate voltage U_g in a circuit with a field-effect transistor. If an isolated QD formed inside a dynamic QWr is located outside the region of the direct effect of the tip, then $U_{tunn} = U_{ds}$. However, when the tip is close to a zero-dimensional defect, we have $U_{tunn} = U_g$. Hence, the local tunneling current–voltage curves (see Figs. 10, 11, 12) that characterize the transport of individual carriers in electrostatically induced Qrs are representative of the standard $I_{ds} = f(U_{ds})$ and $I = f(U_g)$ curves for a single-electron transistor (see Figs. 8b, 8c).

The shape of the local tunneling current–voltage characteristics shown in Figs. 10a, 11a, and 11b corresponds to the mode of the Coulomb blockade, which arises due to the presence a weakly coupled many-hole QD inside a QWr. These *I–V* curves are indicative of the charge quantization that leads to an energy gap in the spectrum of states for charge carrier tunneling through a QD (Fig. 9b). According to the model of the Coulomb blockade, for an electron to tunnel through a QD, the electron thermal energy (kT) and kinetic energy (eU_{tunn}) must exceed the Fermi energy by $e^2/2C$. Accordingly, the hole energy must be lower than the Fermi energy by the same quantity. When measuring local current–voltage characteristics, the threshold voltage for overcoming the Coulomb blockade is determined by the hole kinetic energy ($U_{tunn} > e/2C$). For the structures under study, this voltage is 0.15 and 1.0 V, respectively, both in the forward and reverse current–voltage characteristics. These values correspond to the QD capacitance values of 5×10^{-19} and 8×10^{-20} F. Thus, the parameters of the structures under consideration indicate that the Coulomb blockade can be observed at room temperature, as was demonstrated virtually for the first time when studying the recharging of silicon many-hole QDs (Fig. 11a). The result obtained is probably related to the complicated connecting structure of the QDs, which takes the form

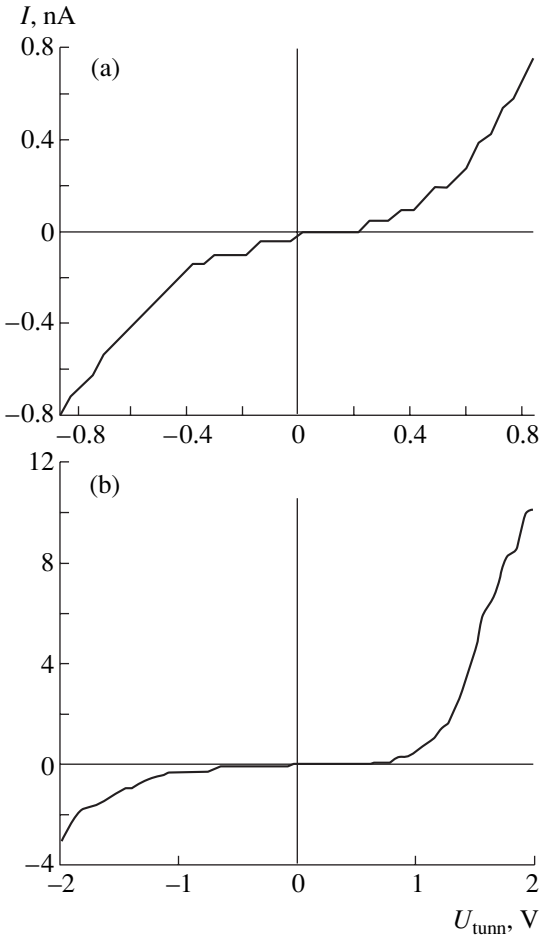

Fig. 12. Local tunneling current–voltage characteristics showing (a) the effect of quantum confinement at the transit of individual holes through a weakly coupled many-hole quantum dot inside a *p*-type SASQWS at the Si (100) surface (T = 330 K) and (b) the different modes of Coulomb blockade at $U_{tunn} > 0$ and $U_{tunn} < 0$ resulting from the asymmetry in the location of a weakly coupled many-hole quantum dot between the microscope tip and the contact at the periphery of the planar structure (T = 300 K).

chains of multitunneling silicon junctions (Figs. 2b, 2c) separated by high-energy barriers. In this case, the total capacitance of the QDs decreases, since a QD itself is a series connection of QDs. As the above current–voltage curves are symmetric with respect to zero applied voltage, a QWr contains at least two tunneling barriers and the main part of the applied voltage drops across these barriers. Under the conditions of the Coulomb blockade, these barriers, between which the QDs are localized, have different widths, which can be electrostatically tuned depending on the direction of the applied voltage, as is indicated by the symmetry of the forward and reverse current–voltage characteristics with respect to zero applied voltage. The similar behavior of the *I–V* curves indicates that the voltage applied between the STM tip and the contact at the boundary of the planar structure corresponds to the source–drain voltage $U_{tunn} = U_{ds}$, since a QD is located virtually in the

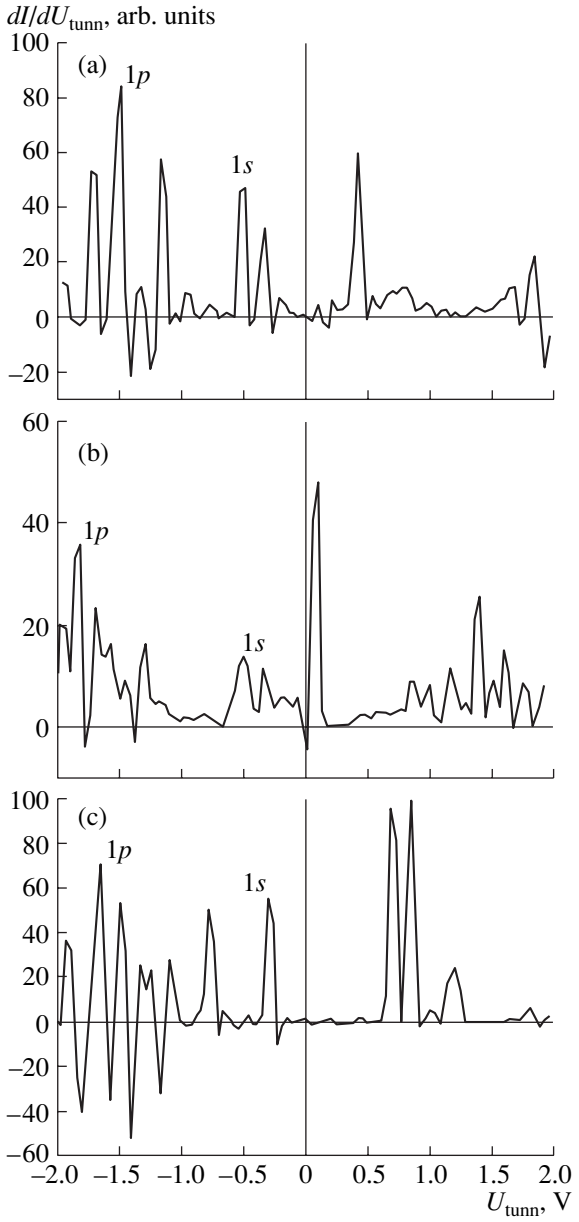

Fig. 13. The conductance of a strongly coupled few-electron quantum dot inside an SASQWS on the Si (100) surface at $T = 300$ K. Spectra (a), (b), and (c) are measured at different positions of the STM tip with respect to the QD and indicate the formation of electronic shells.

middle of the dynamic QWr containing it. However, the situation is different if we consider the local *I–V* curve in Fig. 12b, obtained at an asymmetric location of a QD with respect to the ends of the electrostatically induced QWr containing it (this curve shows the effect of the Coulomb blockade only for the forward current–voltage characteristic). We can see that, in this case, the barrier width is not an electrostatically tuned parameter.

In Figs. 10b and 12a, we can see current–voltage curves of a different type. These curves exhibit features in the shape of steps symmetrically located with respect to $U_{ds} = 0$, both at positive and at negative applied voltages U_{ds}. The appearance of steps in the $I = f(U_{tunn})$ curves can be explained using the above model of Coulomb oscillations, which are seen against a background of Coulomb blockade if we take the confinement effect into account. In other words, the $I = f(U_{tunn})$ curves seem to represent the spectra of the confinement energy levels in QDs.

As U_{ds} increases, the Fermi level of the source (tip) rises with respect to its position in the drain (an ohmic contact to the sample) and, therefore, with respect to the confinement levels. The current starts to flow as the Fermi level of the source crosses a confinement level (Fig. 9c). At a further rise in the Fermi level, the higher confinement levels contribute to the current, thus producing its stepped growth (Figs. 10b, 12a). Therefore, the positions of the steps in Figs. 10b and 12a are determined by the energies of the confinement levels, which are pinned to the levels of charge quantization (Fig. 9c). Furthermore, the presence of Coulomb steps in the local current–voltage curves indicates that, in the specific geometry of the experiment, the isolated dots are located outside the range of the direct effect of the STM tip ($U_{tunn} = U_{ds}$).

4. FEW-ELECTRON ARTIFICIAL ATOMS

4.1. The Formation of Electronic Shells

At low temperatures, few-electron QDs show a discrete energy spectrum with a characteristic gap between the confinement levels. In order of magnitude, this gap is comparable to the energy of the Coulomb interaction. In this case, the energy (U) required for placing an electron onto a QD is essentially similar to the electronic affinity of a real atom. In the presence of a limited number of charge carriers in a QD, low-energy excitations can transfer the carriers to higher confinement levels. Due to the Coulomb interaction between the charge carriers, the energy corresponding to these excitations is much smaller than U.

Thus, we can draw a parallel between charged few-electron QDs and real atoms; however, there are two fundamental distinctions. In real atoms, the Coulomb field has a spherical symmetry and is quantized in units of the elementary charge because of the discreteness of the nucleus charge. It should be noted that, in a real atom, the spectrum of energies corresponding to the addition or removal of an electron can only be weakly controlled. The electrons in a real atom interact with a fixed nucleus potential and with each other, and these two interactions determine the energy spectrum. At the same time, the Coulomb field in a QD may have an arbitrary symmetry and vary continuously (with respect to the elementary charge). Thus, it is possible to totally change the energy spectrum of an artificial atom by changing its geometry or structure [12, 32].

When filling the discrete energy levels in highly symmetric QDs by electrons, the effects of orbital degeneracy and the formation of shells characteristic of

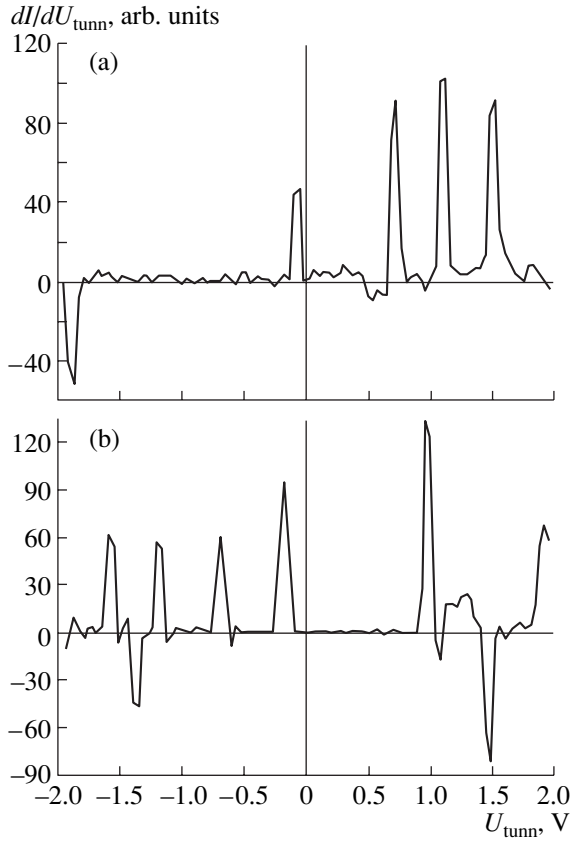

Fig. 14. The conductance of a strongly coupled few-electron QD inside an SASQWS on the Si (100) surface at T = 300 K. Spectra (a) and (b) are measured at different positions of the STM tip near the QD and indicate the suppression of the formation of electronic shells, which results from the closeness of the STM tip to the QD.

real atoms can appear. We recall that the presence of a 3D spherically symmetric potential in an atom results in the appearance of a shell structure with 1*s*, 2*s*, 2*p*, … shells. The ionization energy is at its highest for atomic numbers 2, 10, 18, …, and, up to the atomic number 23, these shells are filled sequentially according to Hund's rule. Therefore, the filling of the corresponding shells in the QDs must also depend primarily on factors such as electron–electron interaction and electron spin [12].

The identification of the electronic shells of QDs has become possible due to the use of scanning tunneling microscopy (STM). A study of the local tunneling current–voltage characteristics of separate semiconductor nanocrystalline InAs QDs, which had a characteristic size ranging from 2 to 8 nm, made it possible to show that, in particular, the first and second excited states of the conduction band had the *s* and *p* symmetry manifesting itself in the corresponding multiplet series [13].

4.2. Local Tunneling Current–Voltage Characteristics of Few-Electron Silicon QDs

Local tunneling $I = f(U_{tunn})$ curves, which represent the energy spectra of QDs, are a particularly effective

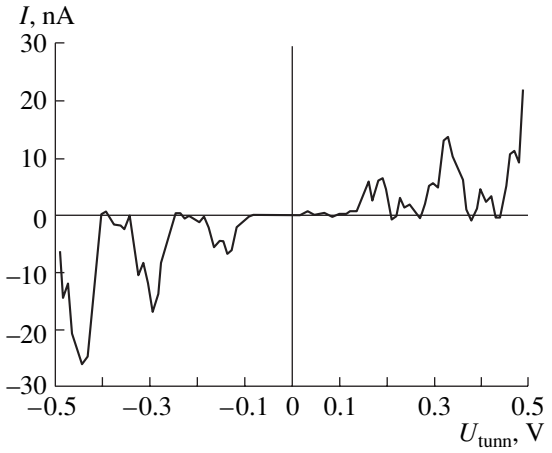

Fig. 15. A local tunneling current–voltage characteristic indicating the Coulomb oscillations under the conditions of removal of the spin degeneracy of the quantum confinement levels at the transit of individual electrons through a weakly coupled few-electron quantum dot inside an SASQWS at the Si (100) surface (T = 4.2 K).

tool for identification of the parameters of few-electron silicon QDs. In this study, we present the results of studies of few-electron small-size silicon artificial atoms for the first time. These results make it possible to determine the positions of the confinement levels, even at room temperature, using local tunneling spectroscopy data.

Figures 13–15 show the current–voltage characteristics obtained for the electron transit through QDs located in *p*-type longitudinal SASQWSs with a built-in longitudinal component in the field of the p^+–n junction [36]. As a result of this configuration, the QD conduction can be locally inverted from *p*-type to *n*-type. These current–voltage characteristics show that, due to their small size, the few-electron QDs under study have a discrete energy spectrum with a characteristic separation between their confinement levels; in order of magnitude, this separation is comparable to the energy of the Coulomb interaction. In particular, the conductance of high-symmetry dots with a characteristic size of ~2 nm (Fig. 1c) is obviously indicative of the *s* and *p* symmetry manifested in the corresponding local tunneling *I*–*V* curves by the series of multiplets (Fig. 13). In other words, when filling discrete energy levels in high-symmetry few-electron QDs with electrons, orbital degeneracy and shell formation specific to real atoms are observed. The effect of the formation of electronic shells appeared to be very sensitive to the local electric fields and virtually completely disappeared both as the STM tip approached a few-electron QD and as its symmetry was disrupted by local fluctuations in the boron distribution in the δ barriers (Fig. 14). It is interesting that, in spite of the suppression of the formation of electronic shells, the energy gaps between the confinement levels are conserved (Fig. 14) and, in addition, in some cases, their spin degeneracy is removed (Fig. 15).

5. CONCLUSIONS

Thus, local tunneling spectroscopy studies have shown that, in the SASQWSs at the surface of Si (100), the STM tip induces QWrs with many-hole weakly coupled QDs and few-electron strongly coupled QDs.

Local tunneling current–voltage characteristics show the different modes of Coulomb blockade and Coulomb conductivity oscillations occurring at the transit of single holes through many-hole weakly coupled QDs. The manifestation of these modes is caused by the Coulomb interaction and quantum confinement effects.

It is shown that, for tunneling through many-hole QDs, the Coulomb oscillations have a periodic character, whereas, for the transport through few-electron dots, the regularity of the oscillations is disrupted.

The local tunneling current–voltage characteristics of high-symmetry few-electron silicon QDs indicate the formation of shells specific to real atoms.

ACKNOWLEDGMENTS

This study was supported by programs of the International Science and Technology Center (grant no. 2136), a grant from the President of the Russian Federation for the support of young Russian scientists (no. MK-4092.2004.2), and a grant for young candidates of sciences from high schools and academy institutes in the St. Petersburg region (no. 11304).

REFERENCES

1. C. Weisbuch and B. Winter, *Quantum Semiconductor Structures* (Academic, New York, 1991).
2. T. J. Thornton, Rep. Prog. Phys. **58**, 311 (1994).
3. C. G. Smith, Rep. Prog. Phys. **59**, 235 (1996).
4. N. N. Ledentsov, in *Proceedings of ICPS-23*, Ed. by M. Scheffler and R. Zimmermann (Berlin, Germany, 1996), p. 19.
5. L. Eaves, in *Proceedings of ICPS-23*, Ed. by M. Scheffler and R. Zimmermann (Berlin, Germany, 1996), p. 43.
6. M. A. Kastner, Phys. Today **46**, 24 (1993).
7. R. Nötzel, Semicond. Sci. Technol. **11**, 1365 (1996).
8. U. Meirav and E. B. Foxman, Semicond. Sci. Technol. **10**, 255 (1995).
9. S. L. Feng and J. C. Bourgoin, Solid State Phenom. **10**, 265 (1989).
10. N. T. Bagraev, A. D. Buravlev, L. E. Klyachkin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 462 (2002) [Semiconductors **36**, 439 (2002)].
11. C. T. Liang, M. Y. Simmons, S. G. Smith, *et al.*, Phys. Rev. Lett. **81**, 3507 (1998).
12. S. Tarucha, D. G. Austing, and T. Honda, Phys. Rev. Lett. **77**, 3613 (1996).
13. O. Millo, D. Katz, Y. Levi, *et al.*, J. Low Temp. Phys. **118**, 365 (2000).
14. D. Goldhaber-Gordon, Hadas Shtrikman, D. Mahalu, *et al.*, Nature **391**, 156 (1998).
15. J. Göres, D. Goldhaber-Gordon, S. Heemeyer, *et al.*, Phys. Rev. B **62**, 2188 (2000).
16. W. Frank, U. Gosele, H. Mehrer, and A. Seeger, in *Diffusion in Crystalline Solids*, Ed. by G. E. Murch and A. S. Nowick (Academic, New York, 1984), p. 63.
17. N. T. Bagraev, W. Gehlhoff, L. E. Klyachkin, *et al.*, Defect Diffus. Forum **143–147**, 1003 (1997).
18. N. T. Bagraev, A. D. Bouravleuv, W. Gehlhoff, *et al.*, Defect Diffus. Forum **194–199**, 673 (2001).
19. J. Robertson, Adv. Phys. **32**, 361 (1983).
20. E. H. Poindexter, P. J. Caplan, B. E. Deal, and G. J. Gerardy, in *The Physics and Chemistry of SiO_2 and Si–SiO_2 Interface* (Plenum, New York, 1988), p. 299.
21. N. T. Bagraev, A. D. Bouravleuv, W. Gehlhoff, *et al.*, in *Abstracts of DIMAT2004* (Krakow, 2004), O-37.
22. P. S. Zalm, Rep. Prog. Phys. **58**, 1321 (1995).
23. W. Gehlhoff, N. T. Bagraev, and L. E. Klyachkin, Mater. Sci. Forum **196–201**, 467 (1995).
24. W. Gehlhoff, N. T. Bagraev, and L. E. Klyachkin, Solid State Phenom. **47–48**, 589 (1995).
25. N. T. Bagraev, W. Gehlhoff, L. E. Klyachkin, *et al.*, Mater. Sci. Forum **258**, 1683 (1997).
26. N. T. Bagraev, E. V. Vladimirskaya, V. É. Gasumyants, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 2133 (1995) [Semiconductors **29**, 1112 (1995)].
27. N. T. Bagraev and V. A. Mashkov, Solid State Commun. **65**, 1111 (1988).
28. N. T. Bagraev, J. Phys. I **2**, 1907 (1992).
29. N. T. Bagraev, W. Gehlhoff, V. K. Ivanov, *et al.*, Phys. Low-Dimens. Semicond. Struct. **1/2**, 37 (2000).
30. D. V. Averin, A. N. Korotkov, and K. K. Likharev, Phys. Rev. B **44**, 6199 (1991).
31. H. van Houten and C. W. J. Beenakker, Phys. Rev. Lett. **63**, 1893 (1989).
32. S. Nagaraja, Ph. Matagne, V.-Y. Thean, *et al.*, Phys. Rev. B **56**, 15752 (1997).
33. L. I. Glazman and M. E. Raĭkh, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 378 (1988) [JETP Lett. **47**, 452 (1988)].
34. L. I. Glazman and K. A. Matveev, Zh. Éksp. Teor. Fiz. **98**, 1834 (1990) [Sov. Phys. JETP **71**, 1031 (1990)].
35. M. Tewordt, V. J. Law, J. T. Nicholls, *et al.*, Solid-State Electron. **37**, 793 (1994).
36. N. T. Bagraev, A. D. Bouravleuv, W. Gehlhoff, *et al.*, in *Proceedings of ICPS-26*, Ed. by A. R. Long and J. H. Davies (Edinburgh, U.K., 2002), Inst. Phys. Conf. Ser., No. 171, G3.3.

Translated by I. Zvyagin

Semiconductors, Vol. 39, No. 6, 2005, pp. 697–702. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 729–734.

LOW-DIMENSIONAL SYSTEMS

Room-Temperature Electroreflectance and Reflectance of a GaAs/AlGaAs Single Quantum Well Structure

A. A. Herasimovich[1, 2]^, S. V. Shokhovets[1, 2], G. Gobsch[2], and D. S. Domanevskiĭ[1]

[1]*Belarussian National Technical University, Minsk, 220100 Belarus*

[2]*Institute of Physics, Ilmenau Technical University, Ilmenau, 98684 Germany*

^*e-mail: Andrei.Herasimovich@tu-ilmenau.de*

Submitted September 14, 2004; accepted for publication November 2, 2004

Abstract—The room-temperature electroreflectance and reflectance of a GaAs/AlGaAs single quantum well (QW) structure are studied. An oscillatory behavior of the electroreflectance signal as a function of the thickness of the top AlGaAs barrier layer is observed. The experimental data are analyzed using a dielectric function of QWs and the transfer matrix method for multilayer systems. This analysis allows the determination of the parameters of the QW and the barrier layers. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Semiconductor structures with quantum wells (QWs) are widely used in the fabrication of injection lasers, LEDs, photodetectors, and optical modulators. The design, optimization, and control over the technology of these devices demand accurate information on the optical properties of QWs (the electronic spectrum and matrix elements of optical transitions) and thicknesses of the structure layers. Generally, optical properties are studied at low temperatures. In various methods of study, based on analysis of absorption, reflectance, photoluminescence, and photoluminescence excitation spectra, the signal is determined by the dielectric function (DF) of a QW and the optical constants of the other layers in the structure. However, when the optical properties are studied at room temperature (which simplifies the technology involved in monitoring the structure parameters), the sensitivity of the conventional low-temperature methods decreases as a result of the thermal broadening of the optical transitions, broadening of the energy distribution of electrons and holes, and decreased contribution of excitonic effects. At the same time, the signal in modulation spectroscopy methods, such as electroreflectance (ER) and photoreflectance, is mainly determined by the derivative of the QW DF. Due to this specific feature, modulation spectroscopy methods retain their high sensitivity and spectral resolution up to room temperature.

The most important aspect in a quantitative analysis of optical spectra is knowledge of the QW DF. Semiempirical models, e.g., Lorentzian or Gaussian oscillator models, have been applied in a large number of publications [1–5]. In such an approach, a specific form of the energy spectrum of QW states can be taken into account, in particular, the strong nonparabolicity of the valence band and dependence of the optical transition matrix element on photon energy. Recently, we reported the results of a calculation of the DF of a GaAs/AlGaAs single QW as a function of the density of 2D electron and hole gases using a self-consistent solution to the Poisson and Schrödinger equations [6]. In the present study, these results are applied to an analysis of the room-temperature ER spectra of an undoped structure. Knowing the DF allows us to apply a transfer matrix formalism to multilayer systems [7] and, in the calculation of optical spectra, to take into account all the layers and interfaces in the structure. In this case, the contribution of a QW is calculated directly rather than being regarded as a adjustable parameter, in contrast to a series of studies by Klipstein *et al.*, who carried out calculations introducing an additional empirical function (see, e.g., [1]). We show that the ER spectrum of a single QW strongly depends on the thickness of the top AlGaAs barrier layer. An analysis of experimental data on ER and reflectance allows us to determine the energies of optical transitions and the thicknesses of the QW and both barrier layers. This technique can be used as a method for the nondestructive room-temperature control of structures.

2. EXPERIMENTAL

We studied a GaAs/AlGaAs single QW structure grown by MBE on a semi-insulating (100) GaAs substrate. The sample included the following sequence of layers: undoped GaAs buffer layers of 200 and 310 nm in thickness, grown at 580 and 400°C, respectively; two undoped $Al_xGa_{1-x}As$ barrier layers (bottom and top) with a GaAs QW between them; and a 17-nm-thick GaAs capping layer. The thicknesses of bottom barrier layer, QW, and top barrier layer, as determined from the experimental data (see Section 4), were 157.5, 9.61, and 214.5 nm, respectively. The Al content x in the barrier layers differed slightly, with $x = 0.308$ in the top barrier and $x = 0.316$ in the bottom.

1063-7826/05/3906-0697$26.00

Table 1. Calculated energies of the optical transitions

Conduction band	Valence band					
	hh_1	lh_1	hh_2	hh_3	lh_2	hh_4
E_1	1.4677	1.4814	1.4898	1.5248	1.5391	1.5618
E_2	1.5655	1.5793	1.5877	1.6227	1.6370	1.6596
E_3	1.6998	1.7138	1.7220	1.7569	1.7712	1.7939

Reflectance and ER were measured at room temperature in an electrolytic cell at a nearly normal incidence of light. The electrolyte was a 3% aqueous solution consisting of a 1 : 2 mixture of tartaric acid and ethylene glycol. The working and reference electrodes were platinum ring and saturated calomel electrodes, respectively. The bias voltage U_{dc} and square-wave modulating voltage, with a frequency of 64 Hz and amplitude U_{mod} (in the ER measurements), were applied with a potentiostat. It should be noted that U_{dc} is the potential of the reverse side of the sample (the GaAs substrate) relative to the reference electrode. The spectra presented below were recorded at $U_{dc} = 0$ and $U_{mod} = 1.5$ V. The spectral resolution of the monochromator was no less than 0.5 meV. A standard lock-in technique was used in the signal detection.

The thickness of the top AlGaAs barrier layer was varied by anodic oxidation of the structure in the same electrolyte. The thickness of the grown oxide was monitored by variation of reflectance at a wavelength of 372 nm. The oxide was removed in a concentrated aqueous solution of NH_3OH. At the first step of etching, the GaAs capping layer and about 10 nm of the top barrier layer were completely removed. At the ninth step, the QW was removed and the etching of the bottom barrier layer was begun.

3. CALCULATION OF THE QW DF

The band structure and wave functions of the electrons and holes in the QW were determined using a self-consistent solution to the Poisson and Schrödinger equations. The conduction band was assumed to be uncoupled to the valence band, and, in terms of the model, it was regarded as having an energy-dependent mass [11, 12]. The structure of the valence band and wave functions of the holes were calculated using the *kp* method, with a 4×4 Hamiltonian in the axial approximation [10–13]. This Hamiltonian describes the interaction between heavy-hole (*hh*) and light-hole (*lh*) subbands. As was shown in [11–13], the axial approximation yields exact values of the energies and wave functions in the center of the Brillouin zone (the electron wave vector $k_p = 0$) and satisfactory results in its vicinity. The mechanical stresses present in the structure were disregarded. The differential equations were solved using the method of finite differentials, with a constant discretization step of 0.1 nm in the direction of growth. In the solution to the Schrödinger equation, the electron and hole wave functions in the AlGaAs barrier layers were set equal to zero at a distance of 15 nm from the interfaces with the QW. In the solution to the Poisson equation, the boundary conditions were determined by the voltage applied to the surface of the structure and by the position of the Fermi level within the GaAs buffer layer at thermodynamic equilibrium. The standard parameters of the material were used in the calculations of the band structure [14]. For all the layers, except for the QW, the residual doping level for donors and acceptors was assumed to be 1×10^{14} and 5×10^{14} cm^{-2}, respectively.

The calculated energies of the optical transitions were fitted to those observed in the ER spectra of the structure under study (for more detail, see Section 4). As a result of this fitting process, a QW thickness of 9.61 nm and a 75/25 ratio between the offsets of the conduction and valence bands were obtained. The thickness of the QW appeared to be a monolayer greater than the nominal thickness (9.33 nm) preset in the growing procedure. The obtained ratio between the band offsets was slightly higher than 65/35, which is the ratio commonly used in band structure calculations, but it lies within the range of values previously reported in other publications [11, 12, 14–16]. For the obtained parameters, we identified three quantum-well states in the conduction band (E_1, E_2, and E_3) and six in the valence band (hh_1–hh_4, lh_1, and lh_2). The calculated energies of the optical transitions from a sublevel in the valence band to a sublevel in the conduction band are listed in Table 1.

The imaginary part of the QW DF was calculated in terms of a dipole approximation under the assumption of direct transitions in quasi-momentum space [17, 18]. The real part of the DF was calculated using the Kramers–Kronig relation. The matrix elements for light polarized parallel to the QW plane, which appear in the expressions for determining the DF, were calculated via the wave functions of the conduction and valence bands. Excitonic effects were disregarded, since, at room temperature, their contribution is small.

Figure 1 shows the QW DF calculated for a zero bias on the structure (the electric field within the QW is about 7 kV cm^{-1}). The arrows indicate the energies of the optical transitions that make the largest contributions to the DF. The other transitions listed in Table 1 are forbidden; therefore, their amplitude is small. Furthermore, their unambiguous identification at room temperature is problematic. Calculations and an analysis of the experimental data indicate that, as a modulation voltage of 1.5 V is applied to the structure, the electric field within the QW is changed from a near-zero value to 20 kV cm^{-1}. Mainly due to the Stark effect, which is the principal mechanism of electroreflectance for the QW under study, this change in the electric field leads to a slight modification of the dielectric function of the QW. It is necessary to note that the discussed

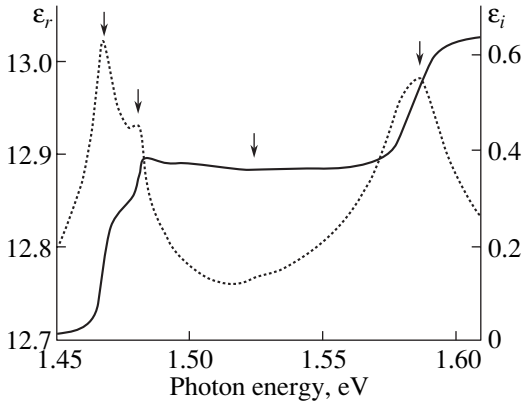

Fig. 1. Real (ε_r) and imaginary (ε_i) parts of the QW DF in the structure under study. The arrows indicate the energies of the optical transitions that make the largest contribution to the DF (from left to the right: hh_1E_1, lh_1E_1, hh_3E_1, and hh_2E_2).

modification of the DF is not seen in the scale used in Fig. 1. For example, under the conditions used, the decrease in the energy of the hh_1E_1 transition is about 1 meV. Furthermore, calculations show that, owing to the low electron density in the experimental conditions, the structure can be regarded, from the standpoint of the potential distribution, as a flat capacitor with a voltage applied across its plates.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Figures 2a and 2b show the experimental reflectance (2a) and ER (2b) spectra after the first and ninth etching steps. The arrows indicate the band gaps of the GaAs buffer and $Al_xGa_{1-x}As$ barrier layers as well as the energies of the principal optical transitions in the QW, which were found from the fitting of the calculated ER spectra to those obtained experimentally. The peaks observed at about 1.42 eV correspond to bulk GaAs (buffer layer). With regard to the QW signal, the optical transitions from the valence band to the first electron subband, hh_1E_1 and lh_1E_1, lie in the energy range 1.46–1.57 eV; transitions to E_2, hh_2E_2 and lh_2E_2 are identified in the range 1.57–1.67 eV; and the range 1.67–1.8 eV corresponds to transitions to E_3 and hh_3E_3. It can be seen that the QW signal in the ER spectrum is much stronger, which provides direct evidence of the significantly better sensitivity of the modulation technique. Two transitions corresponding to bulk $Al_xGa_{1-x}As$ with an Al content $x = 0.308$ and 0.316 are seen in the high-energy portion of the spectrum after the first etching. It can also be seen in these figures that, after the ninth etching, the signal from the QW disappears from the ER spectrum, and, instead of a double structure, only one transition, corresponding to $x = 0.316$, remains in AlGaAs range.

Figure 3 shows the ER spectra in the range of transitions between the ground sublevels of heavy and light

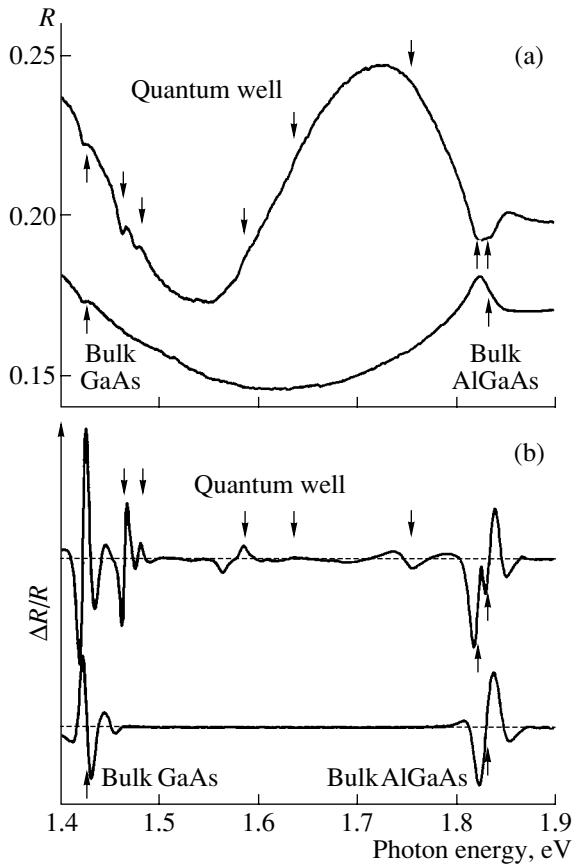

Fig. 2. (a) Reflectance and (b) ER spectra after the first and ninth etching steps. The upward-pointing arrows indicate the energies of the GaAs band gap (the buffer layer) and of the top and bottom $Al_xGa_{1-x}As$ barrier layers. The downward-pointing arrows indicate the energies of the principal optical transitions in the QW (hh_1E_1, lh_1E_1, hh_2E_2, lh_2E_2, and hh_3E_3).

holes, hh_1 and lh_1, and the ground electron sublevel E_1 after the first to eighth etching steps. It can be seen that, as the thickness of the top barrier layer decreases, the maximum of the signal is gradually transformed into a minimum and, vice versa, the minimum into a maximum. As a result, the signal shape is restored after several etching steps (see, e.g., the spectra after the second and sixth or third and seventh etching steps). In other words, the magnitude of the ER signal at a fixed photon energy is an oscillating function of the thickness of the top barrier layer. This behavior is observed for all the optical transitions in the QW, and the period of oscillations decreases as the photon energy increases.

The observed effect is accounted for by the fact that the source of the ER signal, i.e., the QW, lies within the structure and the optical path of light across the top barrier layer is reduced by etching. The calculations show that the shape of the ER signal in the spectral range of the QW transitions is also dependent on the other layers in the structure, in particular, on the thickness of the bottom barrier layer. This dependence is due to the

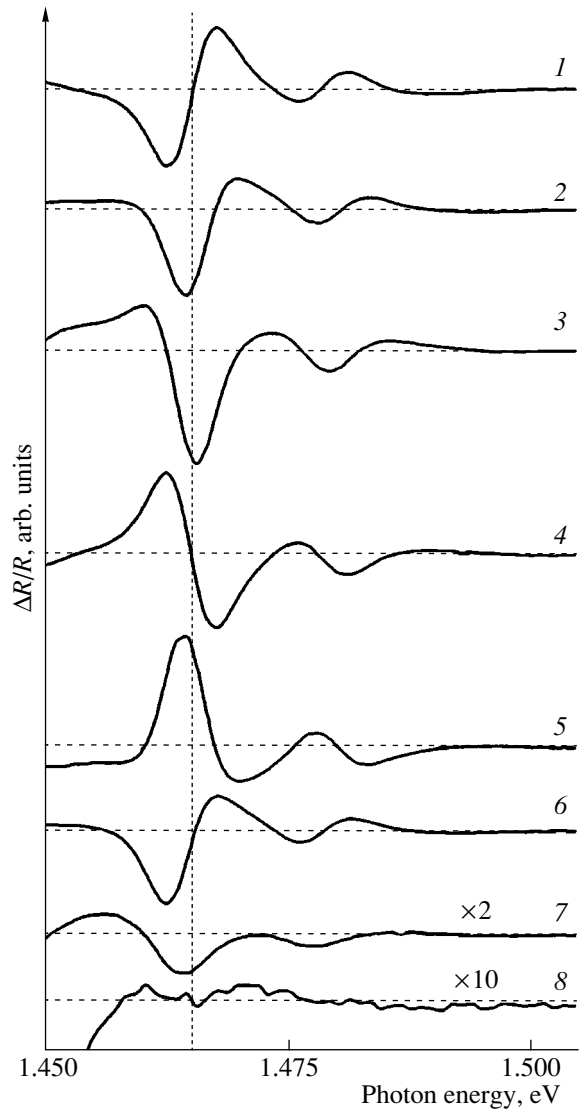

Fig. 3. ER spectra in the spectral range of the principal transitions in the QW (the number of etching step is indicated, *1–8*).

interference of light and shows that all the layers and interfaces of the structure must be considered in a quantitative analysis of the ER spectra. Naturally, this conclusion is also valid for the reflectance spectra, but, in this case, the contribution is proportional to the layer thickness. Therefore, the QW effect in the reflectance spectrum is relatively weak (see Fig. 2). Furthermore, since the Al content in the barrier layers is approximately the same, only the total thickness of barrier layers can be determined from the reflectance spectra in real practice. At the same time, a combined application of the reflectance and ER methods makes it possible to determine the thickness of both the barrier layers and the QW (see below).

The quantitative analysis was performed using the transfer matrix method for multilayer systems [7] and the QW DF, which was calculated as a function of the bias applied to the structure (as described in Section 3). The ER spectra ($\Delta R/R$) were determined from reflectance spectra corresponding to two values of the applied bias, V_1 and V_2:

$$\frac{\Delta R}{R} = \frac{1}{2}\frac{R(V_1) - R(V_2)}{R(V_1) + R(V_2)}. \quad (1)$$

Here, $R(V_1)$ and $R(V_2)$ are the structure reflectances at the biases $V_1 = U_{dc} - U_{mod}$ and $V_2 = U_{dc} + U_{mod}$, respectively. The spectra were calculated under the assumption that, in all the etching steps, starting with the first, the structure consists of a semi-infinite GaAs substrate, bottom and top $Al_xGa_{1-x}As$ barrier layers (with a thickness of d_1 and d_2, respectively), and a QW between these layers. The content of Al in the barrier layers was found by averaging the data obtained by fitting the ER spectra for all the etching steps in the spectral range, which corresponds to the absorption edge of bulk $Al_xGa_{1-x}As$. This content appeared to be somewhat different in the bottom and top barrier layers ($x = 0.316$ and 0.308, respectively). However, the effect of this difference on the energies of the optical transitions and the QW DF is insignificant, and we used the same value $x = 0.312$ for both barriers. In the calculation, we used optical constants of GaAs and AlGaAs taken from [19], which were assumed to be independent of the electric field within the structure and the refractive index of the electrolyte (1.41).

The total thickness ($d_1 + d$) of barrier layers after each etching was determined by fitting the calculated reflectance spectra to the experimental ones (see Fig. 2a). For the first etching step, we obtained ($d_1 + d_2$) = 362 nm. Then, keeping the total thickness constant and redistributing it between the top and bottom barriers, we found the best fit of the calculated to the experimental ER spectrum for the first etching in the spectral range of the QW transitions. The other adjustable parameters were the QW width, band offset, broadening of the optical transitions, and oscillator strength for the hh_1E_1 transition. It was found that the best fit can be obtained when the top barrier layer thickness d_2 = 204.5 nm and bottom barrier layer thickness d_1 = 157.5 nm, the QW width is 9.61 nm, and there is 75/25 ratio between the offsets of the conduction and valence bands. Figure 4 shows a comparison of the calculated and experimental ER spectra in the range of the first two QW transitions.

Using the determined thickness of the top barrier layer after the first etching step, d_2 = 204.5 nm, and the total thickness of barrier layers at each of the etching steps, as obtained from the fitting of the reflectance spectra (the second row in Table 2), we determined the thickness of the top barrier layer for all the etching steps. These data are listed in the third row in Table 2. Using these parameters, the ER spectra for each etching step were calculated, and good agreement with the

Table 2. Thickness of the $Al_xGa_{1-x}As$ barrier at each etching step

Etching step	1	2	3	4	5	6	7	8	9
$d_1 + d_2$, nm	362.0	342.3	322.7	301.9	279.3	233.9	202.7	181.3	155.1
d_2, nm	204.5	184.8	165.2	144.4	121.8	76.4	45.2	23.8	–

experimental data was obtained. It is noteworthy that the above-listed parameters of the QW are confirmed for all the etching steps, and the calculated ER spectra demonstrate oscillatory behavior dependent on the thickness of the top barrier layer. Figure 5 shows this behavior for a photon energy of 1.465 eV. It can be seen that the experimental data points taken from the experimental spectra at the same photon energy (see Fig. 3, the vertical dashed line) agree well with the calculated curve. The deviation of the last two points (the seventh and eighth etching steps) from the calculated dependence is related to a strong decrease in the efficiency of the electric field modulation in the QW at small thicknesses of the top barrier layer. It can be assumed that this effect is due to the presence of a surface charge on the interface between the sample and electrolyte, which we disregarded in our calculations.

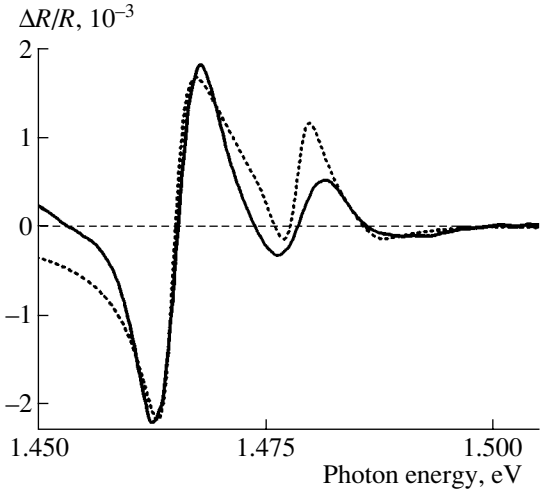

Fig. 4. ER spectra in the range of the hh_1E_1 and lh_1E_1 transitions after the first etching step. The solid line indicates the spectra of the experiment, and the dashed line, those of the calculation.

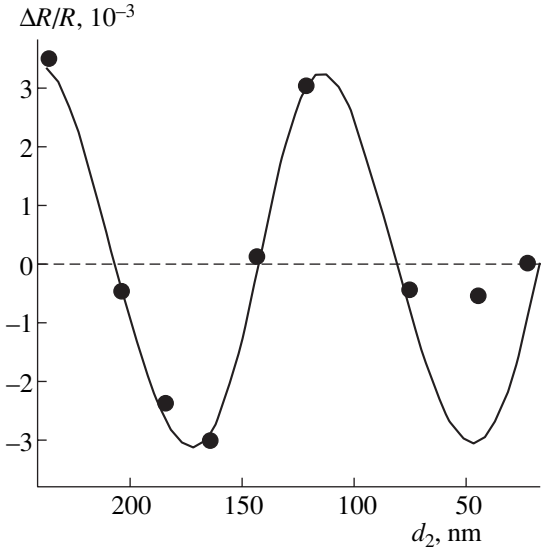

Fig. 5. Calculated (solid line) and experimental (points) magnitudes of the ER signal as function of the thickness of the top barrier layer. Photon energy = 1.465 eV.

5. CONCLUSIONS

The reflectance and electroreflectance of a GaAs/AlGaAs single QW structure was studied at room temperature. A comparison of the experimental data with the results of calculations allowed us to identify the observed transitions and determine the QW parameters (the QW width, band offset, and content of Al in the barriers) and the thickness of each of the barrier layers. This technique can be used as a nondestructive inspection method for QW structures.

ACKNOWLEDGMENTS

The authors are grateful to M. Henini and M. Chamberlain, who provided the samples used in this study; D. Johnston, B. Feldmann, and N. Stein for their participation in the experiments; V. Nakov for his assistance with the calculations; and R. Goldhahn for taking part in numerous valuable discussions.

REFERENCES

1. A. J. Shields and P. C. Klipstein, Phys. Rev. B **43**, 9118 (1991).
2. A. F. Terzis, X. C. Liu, A. Petrou, *et al.*, J. Appl. Phys. **67**, 2501 (1990).
3. R. Ferrini, M. Geddo, G. Guizzetti, *et al.*, Phys. Rev. B **59**, 15395 (1999).
4. D. Y. Lin, S. H. Liang, Y. S. Huang, *et al.*, J. Appl. Phys. **85**, 8235 (1999).
5. Y. Fu, M. Willander, Z. L. Miao, and W. Li, J. Appl. Phys. **89**, 3138 (2001).
6. A. Herasimovich, S. Shokhovets, R. Goldhahn, and G. Gobsch, Thin Solid Films **450**, 199 (2004).
7. M. Born and E. Wolf, *Principles of Optics*, 4th ed. (Pergamon, Oxford, 1969; Nauka, Moscow, 1970).

8. G. Bastard, Phys. Rev. B **24**, 5693 (1981).
9. M. G. Burt, Appl. Phys. Lett. **65**, 717 (1994).
10. M. Altarelli, U. Ekenberg, and A. Fasolino, Phys. Rev. B **32**, 5138 (1985).
11. D. A. Broido and L. J. Sham, Phys. Rev. B **34**, 3917 (1986).
12. A. Twardowski and C. Hermann, Phys. Rev. B **35**, 8144 (1987).
13. D. Ahn and S.-L. Chuang, IEEE J. Quantum Electron. **26**, 13 (1990).
14. I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).
15. D. S. Chemla, D. A. B. Miller, P. W. Smith, *et al.*, IEEE J. Quantum Electron. **20**, 265 (1984).
16. E. H. Li, B. L. Weiss, and K. S. Chan, IEEE J. Quantum Electron. **32**, 1399 (1996).
17. S. L. Chuang, *Physics of Optoelectronic Devices* (Wiley, New York, 1995), p. 385.
18. D. Ahn and T.-K. Yoo, IEEE J. Quantum Electron. **29**, 2864 (1993).
19. C. C. Kim, J. W. Garland, and P. M. Raccah, Phys. Rev. B **47**, 1876 (1993).

Translated by D. Mashovets

Semiconductors, Vol. 39, No. 6, 2005, pp. 703–708. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 735–740.

LOW-DIMENSIONAL SYSTEMS

The Optical Properties of Heterostructures with Quantum-Confined InGaAsN Layers on a GaAs Substrate and Emitting at 1.3–1.55 μm

N. V. Kryzhanovskaya*^, A. Yu. Egorov*, V. V. Mamutin*, N. K. Polyakov*, A. F. Tsatsul'nikov*, A. R. Kovsh*, N. N. Ledentsov, V. M. Ustinov*, and D. Bimberg****

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

^e-mail: kryj@mail.ioffe.ru

***Institut für Festkörperphysik, Technische Universität, D-10623 Berlin, Deutschland*

Submitted November 24, 2004; accepted for publication December 10, 2004

Abstract—Under study is the photoluminescence of two types of MBE-grown heterostructures with quantum-confined InGaAsN/GaAs layers: (1) conventional InGaAsN quantum wells (QWs) in GaAs, and (2) heterostructures with an active region consisting of a short-period GaAsN/InGaAsN superlattice that has an InGaAsN QW with a submonolayer InAs insertion at its center. At room temperature, the structures under study emit light in the range from ~1.3 to ~1.55 μm. In the second type of heterostructure, emission with a wavelength larger than 1.5 μm can be obtained at lower nitrogen and indium concentrations than in a conventional QW. This leads to a significant depression of the effects related to decomposition of an InGaAsN solid solution, thus improving the radiative efficiency of the InGaAsN QWs. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The current interest in studies of quaternary InGaAsN solid solutions is due to their specific physical properties and the possibility of fabricating emitters covering an optical fiber communication range of 1.3–1.55 μm on GaAs substrates [1–3]. Such a system can be considered a serious alternative to the conventional InGaAs/InP system. The strong nonlinearity of the energy gap dependence on the nitrogen content provides an opportunity to significantly increase the emission wavelength of InGaAsN quantum wells (QW) at only a slight increase of the nitrogen content [4]. At the present time, stripe lasers and VCSELs operating in the vicinity of 1.3 μm with an active region based on InGaAsN QWs are used. These lasers demonstrate a low threshold current, high emission power, and high characteristic temperature [1–3]. Lasing at 1.52 μm has also been obtained [5], but the use, in this case, of a rather high content of nitrogen and indium in the InGaAsN QWs in the active region leads to a significant deterioration of the laser characteristics. In order to reduce the mole fractions of indium and nitrogen, various structure designs with additional layers have been suggested, and a 1.41-μm emission was attained without a significant deterioration of the optical properties [6]. Recently, Egorov *et al.* [7] suggested an original method for a further extension of the emission wavelength to the vicinity of 1.55 μm. This method requires the introduction of thin InAs insertions into the center of InGaAsN QWs embedded in short-period InGaAsN/GaAsN superlattices.

In the present study, we perform a detailed investigation and comparative analysis of the optical properties of conventional InGaAsN QWs and InGaAsN QWs with InAs insertions placed in InGaAsN/GaAsN superlattices, which emit in 1.3–1.55 μm range.

2. EXPERIMENTAL

The structures under study were grown by MBE on (001) GaAs substrates using an Applied Epi UNI-bulb RF Plasma Source of nitrogen.

In order to study the effect of QW composition on the optical properties of structures, four structures with 6.2-nm-thick InGaAsN QWs differing in their In and N content were grown:

$In_{0.38}Ga_{0.62}As$, structure *A*;

$In_{0.38}Ga_{0.62}As_{0.974}N_{0.026}$, structure *B*;

$In_{0.38}Ga_{0.62}As_{0.965}N_{0.035}$, structure *C*;

$In_{0.41}Ga_{0.59}As_{0.95}N_{0.05}$, structure *D*.

Because of the existence of an immiscibility range, an increase in the In and N mole fractions leads to a decomposition of the InGaAsN solid solution, thus providing In- and N-enriched regions [8]. To suppress this effect, which is enhanced when the N content increases, an MBE deposition of the N-containing layers was performed at a reduced substrate temperature. It should be noted that the use of these nonequilibrium growth parameters imposes extra requirements on the purity of the initial materials and the growth chamber. In the samples under study, a QW was placed in the center of a 0.2-μm-thick GaAs layer, confined by $Al_{0.3}Ga_{0.7}As$ bar-

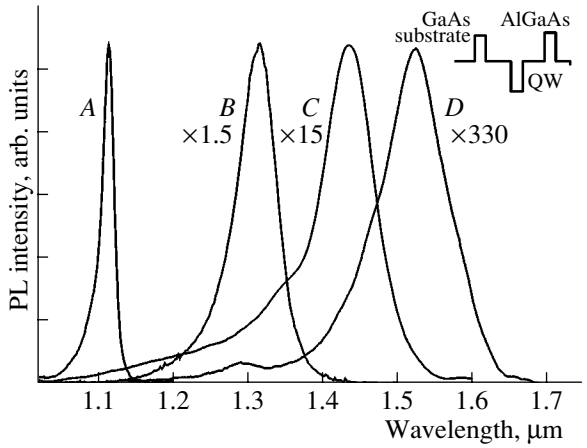

Fig. 1. PL spectra of structures *A*–*D* with $In_xGa_{1-x}As_{1-y}N_y$ QWs possessing different N (y) and In (x) contents. In structures *A*–*C*, $x = 0.38$; in *D*, $x = 0.41$. The values of y are (*A*) 0, (*B*) 0.026, (*C*) 0.035, and (*D*) 0.05. $T = 300$ K.

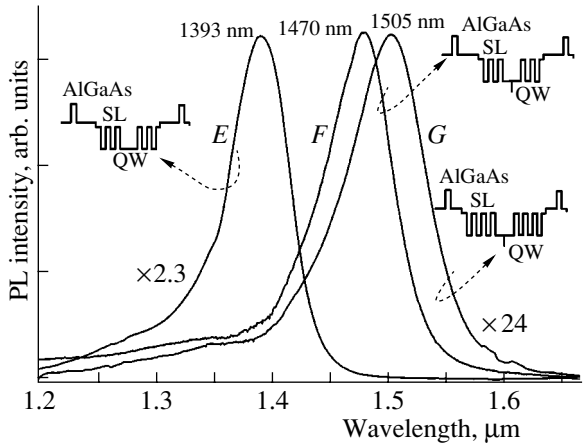

Fig. 2. PL spectra of structures *E*, *F*, and *G*. $T = 300$ K.

riers on the substrate and surface sides. During the growth of this QW, the substrate temperature was reduced to ~350°C. The other parts of the structures were grown at 500–600°C. The inset in Fig. 1 shows a schematic representation of the conduction band bottom for structures *A*–*D*.

In order to estimate the possibility of a further red shift of the emission wavelength, an additional structure, structure *E*, was grown on an MBE EP-1203 machine using the same plasma source of nitrogen as for the other samples (Applied Epi UNI-bulb RF). In structure *E*, a QW with parameters similar to those in structure *B* was placed in a $GaAs_{0.96}N_{0.04}/In_{0.38}Ga_{0.42}As_{0.974}N_{0.026}$ superlattice with a layer thicknesses of 1.3/1 nm, respectively. In structures *F* and *G*, which were based on structure *E*, an InAs layer of about one monolayer in thickness was placed in the center of an $In_{0.38}Ga_{0.42}As_{0.974}N_{0.026}$ layer. Structures *E* and *F* contained two-period superlattices positioned symmetrically on opposite sides of the QW. Structure *G* included three-period superlattices. The total thickness of the central InGaAsN QW was 6.8 nm. The insets in Fig. 2 show a schematic representation of the conduction band bottom for structures *E*–*G*.

Photoluminescence (PL) was excited in the structures either using a CW Ar^+ laser ($W = 1$–1500 W/cm^2 and $\lambda = 514$ nm) or with the second harmonic of a CW YAG:Nd laser ($W = 1500$ W/cm^2 and $\lambda = 532$ nm). Studies of the samples' PL were performed in a closed-cycle helium cryostat in the temperature range 10–300 K. The PL excitation spectra were recorded under excitation induced by the light of a halogen incandescent lamp, which was passed through a monochromator. The samples were placed in a flow-through helium cryostat. The signal was detected using a monochromator and a cooled Ge photodiode.

3. RESULTS AND DISCUSSION

Figure 1 shows the PL spectra of InGaAsN QWs with different N and In contents (structures *A*–*D*). Increasing the nitrogen mole fraction from 0 to 3.5% at an In content of ~38% (structures *A*–*C*) resulted in a shift of the PL peak from 1.1 to 1.43 μm. Further raising of the nitrogen and In content to 5 and 41%, respectively (structure *D*), allowed us to shift the PL peak to 1.55 μm. As the nitrogen and In content increased, the PL intensity decreased and the PL line was broadened, which indicates the rising inhomogeneity of the QW composition and thickness. This observation is related to the decomposition of the solid solution, which induces strong local stresses, deformation of the surface, and the formation of centers at which a 3D localization of carriers occurs [9]. At a high enough In and N content, even the transition to a 3D-growth mode with the resulting formation of QDs is possible [10]. Figure 3a shows the dependences of the PL peak positions on the optical excitation density at 10 K for structures *A*–*D*. It is well known that increasing the excitation density leads to a gradual filling of the localized states related to fluctuations in the QW composition and thickness and to a corresponding blue shift of the PL spectral peak. The absence of this shift for the PL peak of structure *A* indicates the high homogeneity of its $In_{0.38}Ga_{0.62}As$ QW. In the case of $In_{0.38}Ga_{0.62}As_{0.974}N_{0.026}$, the PL peak is shifted by $\Delta = 16$ meV as the excitation density increases by three orders of magnitude. When the nitrogen content is raised to 5% (structure *D*), the peak is shifted by $\Delta = 26$ meV, which indicates the significant inhomogeneity of the layer. This result correlates with the increase observed in the half-width of the structure's PL line (Fig. 1).

Information concerning the presence of localized states in the structures can also be obtained from PL temperature dependences. It should be taken into account that, if the excitation density is sufficient for a complete filling of the centers where the 3D localiza-

tion of carriers occurs, which are induced by potential fluctuations, the recombination of carriers localized in the QW plane becomes dominant. Therefore, in order to provide an adequate characterization of the optical and structural quality of the samples, we studied the PL temperature dependences at a low excitation density (W = 10 W/cm^2). Figure 3b shows the temperature dependences of the PL peak positions. Generally, the temperature shift of the PL peaks is described by the Varshni law, according to which the GaAs band gap decreases by ~84 meV as the temperature increases from 10 to 300 K, in agreement with the temperature shift of the PL peak for structure *A*. At the same time, as the nitrogen content in the layers increases (structures *B*–*D*), the temperature shift of the PL peaks decreases and the type of dependence itself is changed. This phenomenon can be attributed to the fact that, at low temperatures, the radiative processes in InGaAsN QWs are governed by the recombination of excitons localized on fluctuations of the QW potential [11]. As the temperature increases, carriers are excited from localized states to a level corresponding to the carrier localization in the QW plane (the portion of the temperature dependences where the PL peaks are shifted to higher photon energies). For structures *B*–*D*, the inflection point in the temperature dependence, which corresponds to the transition to the recombination of carriers localized in the QW plane, lies at 140 K. Studies of the PL excitation (PLE) spectra at low temperatures allow a quantitative determination of the energy of the carrier localization during fluctuations of the QW potential. Figure 4 shows the PL and PLE spectra of structures *A*–*D*, which were recorded at 7 K. The PLE spectra of the QWs demonstrate features related to absorption on the levels corresponding to the optical transitions *e*1–*hh*1, *e*1–*lh*1, and *e*2–*hh*2 [12]. The Stokes shift between the energies of the PL peak and PLE peak corresponding to the *e*1–*hh*1 transition is not observed in structure *A*; however, owing to the existence of localized states responsible for luminescence at low temperatures, this shift can be observed in structure *B* (Δ_S = 45 meV), and, as the nitrogen content in the QW grows, it increases to 80 meV in structure *D*.

In order to suppress the decomposition of the InGaAsN solid solution, related to a high N and In content, and to improve the optical properties of InGaAsN QWs emitting in the vicinity of 1.55 μm, it has been suggested that thin InAs insertions should be introduced into the center of an InGaAsN QW embedded in an InGaAsN/GaAsN superlattice. This method potentially makes it possible to considerably reduce the average N and In content in the active region of a structure.

In order to study the effects related to the introduction of these additional layers into the InGaAsN/GaAsN superlattice, we grew structure *E*. In this structure, the parameters of the $In_{0.38}Ga_{0.42}As_{0.974}N_{0.026}$ QW were the same as in structure *B*. Figure 2 shows the PL spectrum of structure *E*, which was recorded at room tempera-

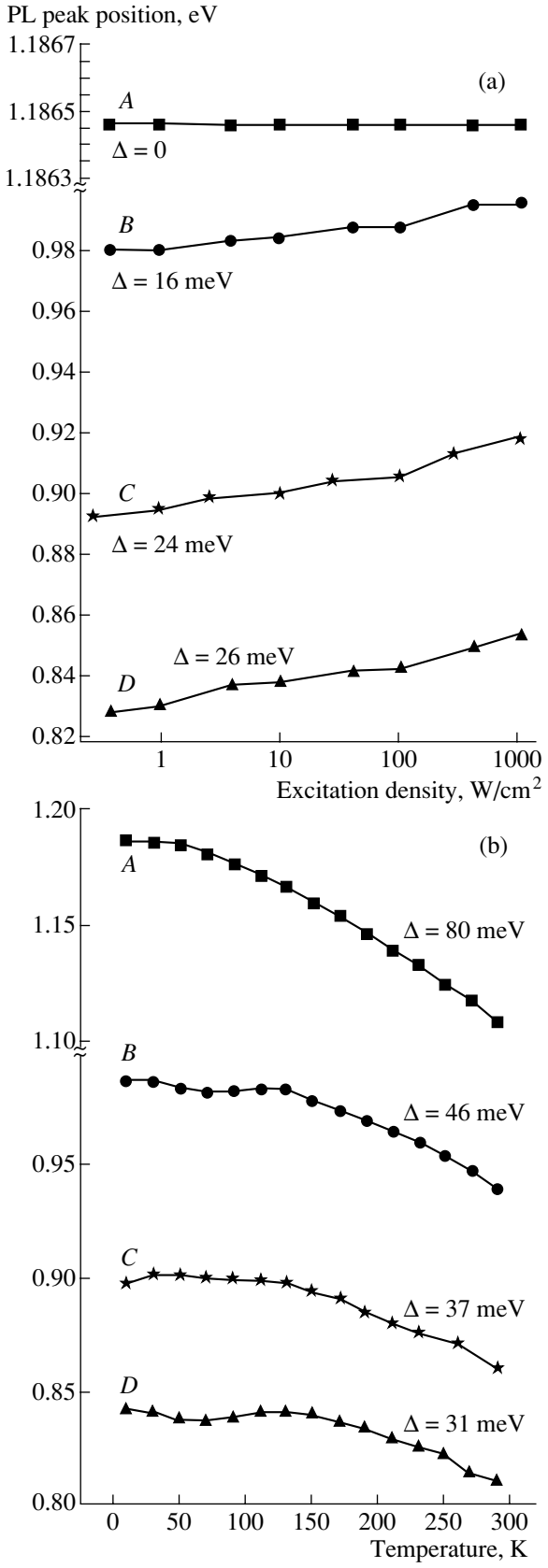

Fig. 3. PL peak positions for structures *A*–*D* vs. (a) the excitation density at T = 10 K and (b) the temperature at an excitation density of 10 W cm^{-2}.

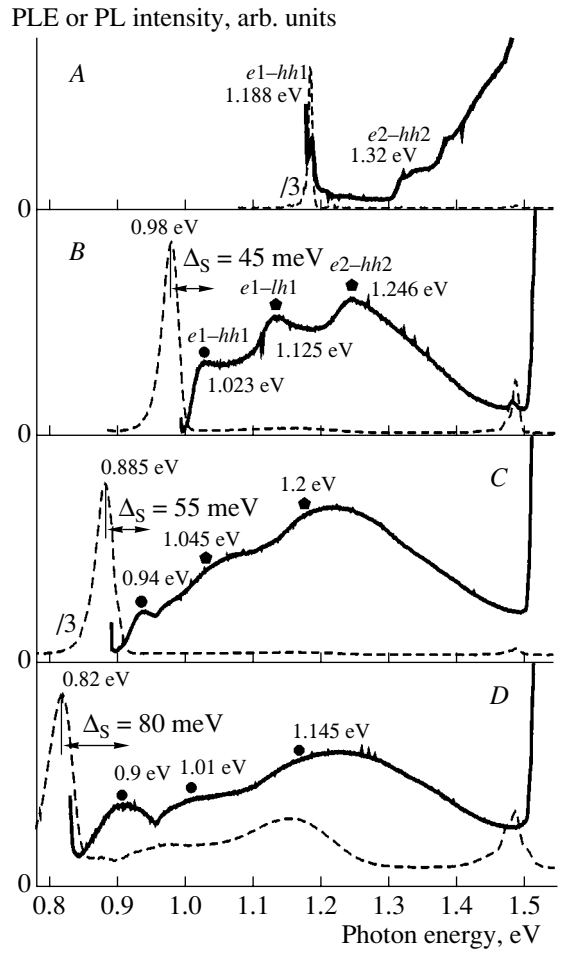

Fig. 4. PL (dashed lines) and PLE spectra (solid lines) for structures *A*–*D*, recorded at $T = 7$ K.

ture. As can be seen, the use of the InGaAsN/GaAsN superlattice leads, owing to a significant reduction of the barrier height, to a shift in the wavelength of the QW emission to 1.4 μm without any major deterioration in the structure quality. The average lattice constant of the used InGaAsN/GaAsN superlattices is close to the lattice constant of GaAs; therefore, it does not induce additional elastic stress in a heterostructure during its formation. Thus, we chose structure *E* as the basis for these InGaAsN QWs with thin InAs insertions (structures *F* and *G*, respectively). Figure 2 also shows the room-temperature PL spectra of structures *F* and *G*. It can be seen that the insertion of a thin InAs layer (structure *F*) into the center of the QW shifts the emission line to 1.49 μm without any reduction of the PL intensity, and the addition of one period to the embedding InGaAsN/GaAsN superlattice results in a further red shift. This heterostructure (structure *G*) emits at 1.505 μm. For all the heterostructure types under study, the advance toward longer emission wavelengths is followed by a decrease in the integral PL intensity. However, for structure *G*, this decrease is significantly

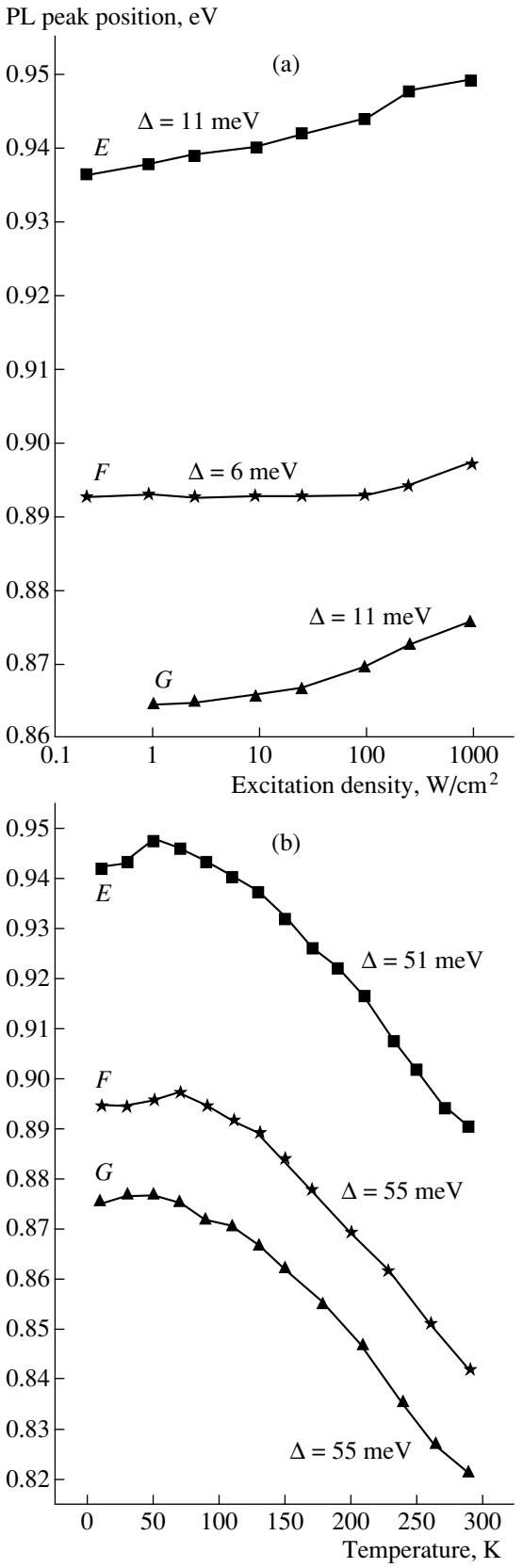

Fig. 5. PL peak positions for structures *E*, *F*, and *G* vs. (a) the excitation density at $T = 10$ K and (b) the temperature at an excitation density of 10 W cm^{-2}.

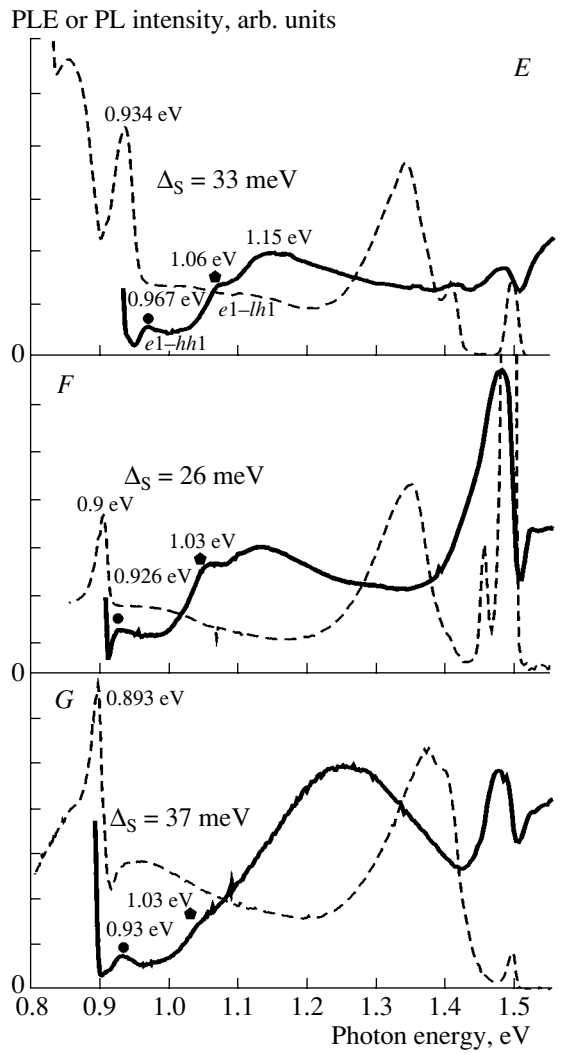

Fig. 6. PL (dashed lines) and PLE spectra (solid lines) for structures *E*, *F*, and *G*, recorded at T = 7 K.

smaller than for structure *D* (by factors of 24 and 330, respectively, as compared to structure *A*).

Figure 5a shows the positions of the PL peaks for structures *E*–*G* as functions of the excitation density at 10 K. It can be seen that the blue shift of the PL peak for structure *E*, Δ = 11 meV, is less than that for structure *B* (16 meV) for the same QW composition; moreover, it is further reduced to 6 meV for structure *G*, in which an InAs layer is added in the center of the InGaAsN QW. The temperature dependences of the PL peak position (Fig. 5b) also demonstrate that the contribution of localized states in the radiative processes in structures *F* and *G* is reduced. For these structures, the PL peaks are shifted by 55 meV in the temperature range 10–300 K, and the inflection point in the temperature dependence of the PL peak position is at about 70 K, whereas, in conventional QW structures, it lies at about 140 K.

The PLE spectra for structures *E*–*G* are shown in Fig. 6. In the PL spectra, the peak in the 1.3–1.4 eV range corresponds to emission from a GaAsN layer, which is formed in the buffer GaAs layer at the instant when the plasma discharge is initiated in the source of atomic nitrogen. The PLE spectra for structures *E*–*G* demonstrate peaks at energies ~1.1–1.15 eV, which correspond to absorption in the InGaAsN/GaAsN layers, and peaks corresponding to the *e*1–*hh*1 and *e*1–*lh*1 optical transitions in the QW (marked by points). The Stokes shift Δ_S in structures *E*–*G* is much smaller than in structures with simple QWs (samples *B*–*D*, Fig. 4). This suggests a significant depression of the effects related to the formation of localized states produced as the nitrogen content in the structures under study is raised; this rise is necessary for a further advance of the emission wavelength to 1.55 μm.

4. CONCLUSION

The optical properties of two types of heterostructures have been studied: (1) those with conventional InGaAsN QWs in GaAs, and (2) structures in which the active region consists of a short-period GaAsN/InGaAsN superlattice and a InGaAsN QW with a submonolayer-thick InAs insertion placed in its center. At room temperature, the studied heterostructures emit light in a wavelength range from ~1.3 to ~1.55 μm. A significant difference in the optical properties of the first and second types of heterostructures is found. The second type of heterostructure is characterized by a smaller half-width of the PL line and smaller Stokes shift between the energies of the PL peak and the PLE peak corresponding to the *e*1–*hh*1 transition. These facts indicate that, owing to the lower average content of nitrogen and indium used in these heterostructures as compared to in a conventional QW, local fluctuations in the elemental composition are depressed. Thus, the growing technique used and the design of structures allowed us to reach the necessary wavelength while improving the luminescence efficiency. The second type of heterostructures offer good prospects for the improvement of laser characteristics and can be regarded as potential candidates for the fabrication of 1.55-μm injection lasers.

ACKNOWLEDGMENTS

The study was supported by the Ministry of Science and Technology of the Russian Federation under the program “Physics of Solid-State Nanostructures,” the Russian Foundation for Basic Research (project no. 02-02-17677), the project “Self-Assembled Semiconductor Nanostructures for New Devices in Photonics and Electronics” (SANDiE), and INTAS under the program “Young Scientist Fellowships” (project no. 03-55-882).

REFERENCES

1. J. S. Harris, Semicond. Sci. Technol. **17**, 880 (2002).
2. B. Borchert, A. Yu. Egorov, S. Illek, *et al.*, Electron. Lett. **35**, 2204 (1999).
3. M. C. Larson, M. Kondow, T. Kitatani, *et al.*, IEEE Photonics Technol. Lett. **10**, 188 (1998).
4. M. Kondow, K. Uomi, K. Hosomi, and T. Mozume, Jpn. J. Appl. Phys. **33**, L1056 (1994).
5. M. Fischer, M. Reinhardt, and A. Forhel, Electron. Lett. **36**, 1208 (2000).
6. L. F. Bian, D. S. Jiang, S. L. Lu, *et al.*, J. Cryst. Crowth **250**, 339 (2003).
7. A. Yu. Egorov, V. V. Mamutin, and V. M. Ustinov, RF Patent No. 2 004 113 171 (28 April 2004).
8. B. V. Volovik, A. R. Kovsh, W. Passenberg, *et al.*, Semicond. Sci. Technol. **16**, 186 (2001).
9. X. P. Xin and C. W. Tu, Appl. Phys. Lett. **72**, 2442 (1998).
10. V. A. Odnoblyudov, A. Yu. Egorov, N. V. Kryzhanovskaya, *et al.*, Tech. Phys. Lett. **28**, 964 (2002).
11. A. Polimeni, M. Capizzi, M. Geddo, *et al.*, Appl. Phys. Lett. **77**, 2870 (2000).
12. H. D. Sun, M. D. Dawson, M. Othman, *et al.*, Appl. Phys. Lett. **82**, 376 (2003).

Translated by D. Mashovets

Semiconductors, Vol. 39, No. 6, 2005, pp. 709–711. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 741–743.

AMORPHOUS, VITREOUS, AND POROUS SEMICONDUCTORS

Structural Defects and Electrical Conductivity in Nanocrystalline SiC:H Films Doped with Boron and Grown by Photostimulated Chemical-Vapor Deposition

O. I. Shevaleevskiy*^, S. Y. Myong, K. S. Lim**, S. Miyajima***, and M. Konagai*****

**Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, ul. Kosygina 4, Moscow, 119991 Russia*

^e-mail: O_Chevale@Yahoo.com

***Korea Advanced Institute of Science and Technology, Daejeon 305-701, South Korea*

****Tokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, Tokyo 152-8552, Japan*

Submitted July 21, 2004; accepted for publication October 6, 2004

Abstract—The paramagnetic *DB* defects and dark conductivity σ_d in films of nanocrystalline hydrogenated silicon doped with boron and carbon (*nc*-SiC:H) and grown by photostimulated chemical vapor deposition are studied. It is shown that an increase in the doping level leads to a phase transition from the crystalline structure to an amorphous structure. The electrical conductivity increases as the doping level increases and attains the value of $\sigma_d = 5.5 \times 10^{-2}\ \Omega^{-1}\ cm^{-1}$; however, the conductivity decreases once the phase transition has occurred. The concentration of *DB* defects decreases steadily as the doping level increases and varies from $10^{19}\ cm^{-3}$ (in the crystalline structure) to $9 \times 10^{17}\ cm^{-3}$ (in the amorphous structure). © *2005 Pleiades Publishing, Inc.*

Nanocrystalline hydrogenated silicon doped with carbon (*nc*-SiC:H) is a promising semiconductor material for the components of nanoelectronics and solar cells and also for the fabrication of optically transparent thin-film conducting coatings [1, 2]. Heterogeneous *nc*-SiC:H films grown by photostimulated chemical vapor deposition (photo-CVD) have a wide optical-transmission band (~2.2 eV) and low dark electrical conductivity $\sigma_d \sim 10^{-6}\ \Omega^{-1}\ cm^{-1}$ [3, 4]. However, the value of σ_d needs to be increased if these materials are to be used effectively in electronic devices. In the cases of hydrogenated microcrystalline silicon (μ*c*-Si:H) and nanocrystalline silicon carbide (*nc*-SiC:H), this increase is attained by doping with boron; the conductivity increases in proportion to the doping level at boron concentrations of 10^{18}–$10^{20}\ cm^{-3}$ [5–7]. However, the film-growth conditions and the methods of introducing the impurity can profoundly affect the dependence of σ_d on the doping level [8]. In heterogeneous structures, the doping level and the electrical conductivity can have a complex interrelation, since the introduction of doping atoms during growth may affect the crystalline-phase formation. This issue requires additional study because high electrical conductivity in μ*c*-Si:H and *nc*-SiC:H is typically related to their crystalline phase and to the presence of a high concentration of *DB* defects (~$10^{18}\ cm^{-3}$) in this phase [3, 7]. At the same time, there are no available experimental data on the interrelation between the paramagnetic *DB* defects, electrical conductivity, and structural phase composition of *nc*-SiC:H doped with boron.

In this paper, we report the results of studying the effect of doping with boron on the structural and electrical parameters of thin *nc*-SiC:H films formed using photo-CVD. We used electron spin resonance (ESR) to measure the concentrations of paramagnetic defects and study their interrelation with the electronic transport.

Boron-doped *nc*-SiC:H films with a thickness of ~200 nm were deposited onto glass substrates (Corning 7059) using the dissociation of monosilane (SiH_4), hydrogen (H_2), and ethylene (C_2H_4) under the effect of ultraviolet radiation (the photo-CVD method). Doping with boron was accomplished by introducing diborane (B_2H_4) into the reaction chamber. The doping level was determined as the volume ratio between the gas fluxes $C_B = [B_2H_4]/[SiH_4]$, which was varied from 10^{-3} to 10^{-2} in the course of growth. During the deposition of the films, the ratios $[H_2]/[SiH_4] = 20$ and $[C_2H_4]/[SiH_4] = 0.07$ were kept constant. The residual-gas pressure in the growth chamber was 0.46 Torr, and the substrate temperature was 250°C. The experimental aspects of the fabrication of such samples were described in detail by Myong *et al.* [9]. The density of paramagnetic defects (N_s) was determined from an analysis of the ESR spectra using the method described in [4]. The dark electrical conductivity was measured at room temperature using the four-probe method.

The structure of the films was studied using Raman scattering spectroscopy. In Fig. 1, we show the characteristic Raman spectra of the *nc*-SiC:H samples at different boron concentrations. The spectra were analyzed by decomposing the spectral curve into Gaussian components that corresponded to the following structural phases: crystalline silicon *c*-Si (the peak at ~520 cm^{-1}), *a*-SiC:H with a low carbon content (~480 cm^{-1}), and an

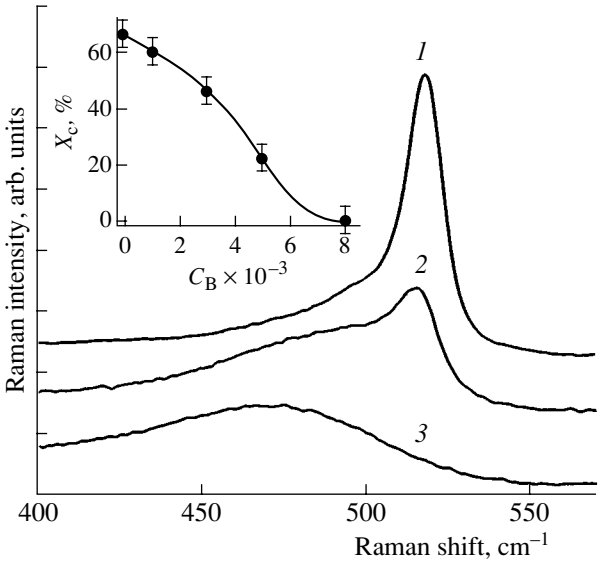

Fig. 1. Experimental Raman spectra of the *nc*-SiC:H samples that were doped with boron at levels of C_B = (*1*) 10^{-3}, (*2*) 3×10^{-3}, and (*3*) 8×10^{-3}. The degree of sample crystallinity X_c as a function of the degree of boron doping is shown in the inset.

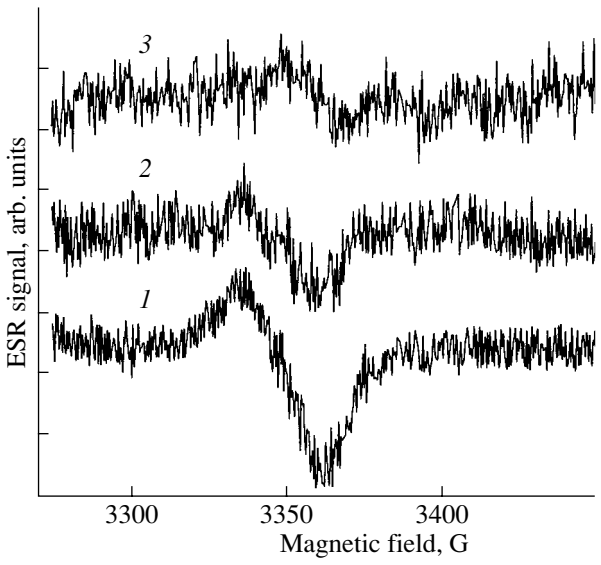

Fig. 2. Evolution of the ESR spectrum in *nc*-SiC:H as a result of doping with boron and variation in the degree of crystallinity: (*1*) $C_B = 10^{-3}$, $X_c = 65\%$; (*2*) $C_B = 3 \times 10^{-3}$, $X_c = 48\%$; and (*3*) $C_B = 8 \times 10^{-3}$, $X_c = 0$.

intermediate phase in the vicinity of the nanocrystal boundaries (~510 cm^{-1}). The volume content of the nanocrystalline fraction (X_c) was defined as the ratio between the intensities of the corresponding Raman peaks: X_c (%) = $[(I_{520} + I_{510})/(I_{520} + I_{510} + I_{480})] \times 100$ [3]. We believe that carbon is mainly localized in the region of grain boundaries, where its content is estimated to be ~6 at % [3]. In the inset in Fig. 1, we show data on variations in the quantity X_c in relation to the level of doping with boron for all the samples under study. As can be seen from Fig. 1, an increase in the boron content during film growth stimulates a structural phase transition from the crystalline structure ($X_c \approx 70\%$) to a completely amorphous structure. This transition occurs at the doping level $C_B = 8 \times 10^{-3}$. An analysis of the curves shows that the heterogeneous samples have a mixed structure and consist of *c*-Si nanocrystals in an amorphous matrix if the doping level is relatively low ($C_B < 4 \times 10^{-3}$). A further increase in the boron content leads to the transformation of the structure into a predominantly amorphous state.

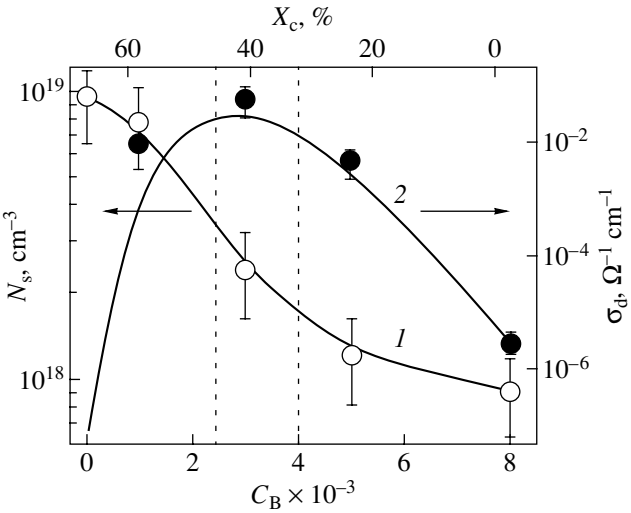

Fig. 3. Dependences of the (*1*) concentration of *DB* defects N_s and (*2*) dark conductivity σ_d in *nc*-SiC:H on the level of doping with boron C_B and on the degree of crystallinity X_c. The region of the structural phase transition is delimited by dashed lines.

Figure 2 illustrates the evolution of the ESR spectra in the *nc*-SiC:H films as the level of doping with boron increases. All the spectra feature a broad asymmetric ESR line with a half-width (ΔH_{pp}) of about 20–30 G and *g* factor equal to 2.006 ± 0.0005, which corresponds to the neutral *DB* defects in silicon [10]. The shape and position of the ESR spectra are controlled by paramagnetic *DB* defects located at the surface of nanocrystals and in the region of grain boundaries; at the same time, the broadening and asymmetry of the resonance line are typically related to a large spread of paramagnetic defects in the crystalline fraction [4, 11].

Figure 3 illustrates the quantitative interrelation between the doping level (C_B), concentration of paramagnetic defects (N_s), and value of the dark electrical conductivity (σ_d). In the films with a high degree of crystallinity ($X_c > 50\%$), an increase in the level of doping with boron to $C_B \approx (3–4) \times 10^{-3}$ is accompanied by a

steady increase in σ_d from 10^{-7} Ω^{-1} cm^{-1} (in an undoped sample) to the largest value of 5.5×10^{-2} Ω^{-1} cm^{-1}. A further increase in the doping level ($C_B > 4 \times 10^{-3}$) initiates the phase transition from the crystalline state to the amorphous state. The dashed lines in Fig. 3 indicate the approximate boundaries of the structural phase transition. In the region beyond the transition, an increase in the doping level ($4 \times 10^{-3} < C_B < 8 \times 10^{-3}$) leads to a sharp (by two orders of magnitude) decrease in σ_d. We assume that this decrease in σ_d is accounted for by destruction of the residual crystalline phase in a sample, occurring as the boron content increases, when the amorphous phase is prevalent in it.

Let us consider the behavior of σ_d and N_s from the standpoint of variation in the phase composition of the films in the course of doping. We believe that the paramagnetic defects, whose number is excessive with respect to the amorphous phase, are caused by the nanocrystalline phase and that their number is proportional to the sample crystallinity X_c. Indeed, as follows from Fig. 3, the value of N_s decreases steadily as a result of doping, from $N_s = 10^{19}$ cm^{-3} in an undoped sample, where $X_c = 70\%$, to $N_s = 9 \times 10^{17}$ cm^{-3} in a completely amorphous structure obtained at a doping level $C_B = 8 \times 10^{-3}$. In turn, the behavior of the conductivity curve is representative of changes in the mechanisms of charge transport resulting from the transition from the crystalline structure to the amorphous state. At low concentrations of doping impurity, the large value of σ_d in the samples with a high level of crystallinity is caused by efficient charge transport via surface defects in the crystallites [11]. In this case, an introduction of boron acceptor impurity increases σ_d even further. Once the structural phase transition has occurred and the amorphous phase becomes prevalent in the sample, an expected increase in σ_d due to an increase in the boron concentration does not compensate for the decrease in σ_d caused by the destruction of crystalline phase in the sample. As a result, the conductivity curve exhibits a maximum in the transition region, and a further increase in the boron concentration is accompanied by a decrease in the value of σ_d. As was mentioned above, the decrease in conductivity as the doping level is increased is caused, in this case, by a decrease in the value of X_c. We can state that an increase in the conductivity as a result of doping heterogeneous *nc*-SiC:H films possessing mixed crystalline and amorphous phases with boron is effective only in samples with a high degree of crystallinity.

Thus, in this study, we used ESR measurements to study, for the first time, the paramagnetic defects in thin *nc*-SiC:H films doped with boron and grown by the photo-CVD method. We found that there is an interrelation between the degree of crystallinity of the samples, the concentration of *DB* defects, and the value of conductivity. It is shown that an increase in the level of doping to $C_B \approx (3\text{–}4) \times 10^{-3}$ during film growth initiates a structural phase transition from the crystalline structure to an amorphous state.

This study was supported in part by the Russian Foundation for Basic Research, project no. 04-02-17050.

REFERENCES

1. S. Y. Myong, S. S. Kim, and K. S. Lim, J. Appl. Phys. **95**, 1525 (2004).
2. Y. Hamakawa, H. Okamoto, and Y. Tawada, Int. J. Sol. Energy **1**, 125 (1982).
3. O. I. Shevaleevskiĭ, A. A. Tsvetkov, L. L. Larina, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **38**, 547 (2004) [Semiconductors **38**, 528 (2004)].
4. O. Chevaleevski, S. Y. Myong, and K. S. Lim, Solid State Commun. **128**, 355 (2003).
5. A. G. Kazanskiĭ, H. Mell, E. I. Terukov, and P. A. Forsh, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 41 (2002) [Semiconductors **36**, 38 (2002)].
6. H. K. Lee, S. Y. Myong, K. S. Lim, and E. Yoon, J. Non-Cryst. Solids **316**, 297 (2003).
7. F. Finger, J. Muller, C. Malten, *et al.*, J. Non-Cryst. Solids **266–269**, 511 (2000).
8. M. M. Mezdrogina and A. V. Patsekin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 354 (2000) [Semiconductors **34**, 348 (2000)].
9. S. Y. Myong, T. H. Kim, K. S. Lim, *et al.*, Sol. Energy Mater. Sol. Cells **81**, 485 (2004).
10. C. Boehme and K. Lips, J. Non-Cryst. Solids **338–340**, 434 (2004).
11. M. Stutzmann, M. S. Brandt, and M. W. Bayerl, J. Non-Cryst. Solids **266–269**, 1 (2000).

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 712–715. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 744–747.

PHYSICS OF SEMICONDUCTOR DEVICES

The Physical Properties of CdTe Doped with V and Ge

S. Yu. Paranchych^, L. D. Paranchych, V. N. Makogonenko, Yu. V. Tanasyuk, M. D. Andriĭchuk, and V. R. Romanyuk

Chernovtsy National University, Chernovtsy, 58012 Ukraine

^e-mail: parsu@chnu.cv.ua

Submitted May 27, 2004; accepted for publication August 9, 2004

Abstract—CdTe:(V, Ge) single crystals are grown using the Bridgman–Stockbarger method. The impurity concentrations in the melt are $N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 5 \times 10^{18}$ and 1×10^{19} cm^{-3}. Electrical and galvanomagnetic characteristics are studied in the temperature range 300–400 K. It is found that the equilibrium characteristics are governed by deep levels (ΔE = 0.75–0.95 eV) located close to the midgap. Low-temperature optical absorption spectra indicate that the impurity levels of V and Ge ions in the low-energy region are in different charge states. In addition, the samples are annealed in Cd vapor and then rapidly cooled. This annealing causes the decomposition of various complexes formed during the crystal growth and an increase in both electrical conductivity and charge carrier concentration. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Recently, semiconductor CdTe single crystals doped with iron Group transition elements (3*d* elements) have been extensively studied in relation to their potential for use as photorefractive materials for the near-IR spectral region. A particular feature of these materials is that their electrooptical coefficients (z_{41}) exceed the corresponding values for practically important crystals such as GaAs and InP by a factor of 3 [1, 2].

Another advantage of CdTe in comparison with oxide $BaTiO_3$ and $LiNbO_3$ crystals, whose photorefractive properties have been thoroughly studied, is a much shorter response time, due to the high mobility of photoexcited electrons in this material.

It is known that only certain impurities, namely V, Ti, Ge, Sn, Fe, and Mn, can form donor–acceptor levels close to the midgap in CdTe [3, 4], and the greatest number of studies have been carried out on CdTe:V [5–8]. In these investigations, it was confirmed that V exists in two charge states (V^{2+} and V^{3+}), and the depth of the V^{2+} level was determined as 0.94 eV relative to the conduction-band bottom. Studies of the optical and photoelectrical inhomogeneity of CdTe:V crystals indicate that the impurity is distributed nonuniformly in the axial and radial directions when the crystals are grown by the Bridgman–Stockbarger method [9]. Thus, the problem of obtaining a uniform photorefractive material remains unsolved.

However, recent studies devoted to CdTe:Ge indicate that this material, along with CdTe:V and CdTe:Ti, is also a promising photorefractive material [10–12].

It has been shown [13–15] that, irrespective of the growth method, a semi-insulating state is attained in CdTe:Ge crystals under the condition that the concentration of the doping impurity N exceeds a certain critical value N_{cr}, which equals $(2–3) \times 10^{16}$ cm^{-3}. For CdTe:Ge, resistivity ρ abruptly increases by six to seven orders of magnitude in the region of the critical doping level, and the photosensitivity and thermal stability of the material increase [13, 16]. This behavior is attributed to a compensating effect and the amphoteric characteristics of Ge in the CdTe lattice.

Thus, taking into account the defect formation in both sublattices in Ge-doped CdTe (Ge_{Cd}, Ge_{Te}) and the nonuniform distribution of V in CdTe both in the radial and axial directions, codoping with V and Ge should significantly improve the material's homogeneity and electrical properties.

The aim of this study is to obtain CdTe:(V, Ge) crystals with $N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 5 \times 10^{18}$ and 1×10^{19} cm^{-3} using the Bridgman–Stockbarger method and to examine the homogeneity of their equilibrium electrical and optical properties.

2. EXPERIMENTAL

We grew single crystals of CdTe:(V, Ge) by the Bridgman–Stockbarger method from melts with the content $N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 5 \times 10^{18}$ and 1×10^{19} cm^{-3}. The resulting crystals were 12–18 mm in diameter and 6–8 cm in length. During their growth in a quartz cell, the free space above the melt was equal to the crystal volume.

The purity of the initial Cd and Te was provided for by the completion of additional purification processes. A zone purification of Te was carried out in Pyrex containers, which were evacuated, filled with hydrogen, and sealed. The concentration of residual impurities, which was determined from the Hall measurements, was $(1–2) \times 10^{14}$ cm^{-3} at 77 K per 2/3 of the ingot length. Cd-0000 grade Cadmium was additionally puri-

fied in graphite containers placed into quartz cells, which were evacuated, filled with hydrogen, and sealed. The purification was carried out by zone melting, with 10–12 passages of the melted zone along the ingot at a rate of 3.5 cm/h. The rate of the last passage, 1.5 cm/h, allowed us to obtain a single-crystal material, which is more suitable for further use.

The samples to be used in the electrical and optical measurements were prepared from various parts of the crystal via mechanical cutting, grinding, and optical polishing, followed by chemical treatment in a bromomethyl etchant. To study the kinetic effects, we used samples $10 \times 2.5 \times 1.5$ mm in size. The measurements were carried out using a specially designed setup for the measurement of semi-insulating crystals in the temperature range 300–400 K. The setup included two identical electrometric amplifiers mounted according to a standard scheme [17]. In order to increase the input resistance of the electrometric amplifiers, we designed the input differential stage based on a carefully selected pair of field-effect KPS-104A transistors. This design provided an increase in the input resistance of the amplifiers to 10^{13} Ω. The input circuits of the electrometric amplifiers were switched by RES-60 electromagnetic relays with hermetic contacts and high isolation resistivity. Optical properties were studied at 300 and 80 K using a standard procedure based on MDR-23 and IKS-21 monochromators.

The samples were annealed in quartz cells. They were placed in one part of the cell and a certain amount of Cd was placed in another part in order to impose vapor pressure. Diffusion annealing was carried out in a horizontal two-zone furnace. The temperature in the sample zone was 720°C, and the temperature in the zone of pure Cd was 690°C. The samples were annealed under these conditions for 72 h.

3. RESULTS AND DISCUSSION

Figure 1 shows the temperature dependences of the conductivity σ and Hall coefficient R for CdTe:(V, Ge) samples ($N_V = 1 \times 10^{19}$ cm^{-3}, $N_{Ge} = 5 \times 10^{18}$ cm^{-3}) cut from various parts of the crystal. The samples had p-type conductivity at room temperature, and the hole mobility was no higher than 85 cm^2/(V s). It is noteworthy that studies of the heating–cooling cycle confirmed the high reproducibility of the results, which is indicative of the high thermal stability of the material. The shown dependences $\sigma(T)$ and $R(T)$ are typical of compensated semiconductors, and the distinctions in σ and R for the samples cut from the medium and upper parts of the crystal were small, which is indicative of high the uniformity of the properties along the grown crystal.

The mobility varies with the temperature close to the dependence $T^{3/2}$, which indicates that scattering by optical phonons is prevalent [18]. Since the effective densities of states in the conduction band, $N_c = 2(m_n^* kT/2\pi\hbar)^{3/2}$, and in the valence band, $N_v = 2(m_p^* kT/2\pi\hbar)^{3/2}$, are proportional to $T^{3/2}$, the temperature dependence of the resistivity, $\rho = (en\mu_n + ep\mu_p)^{-1}$, is exclusively determined by the exponential function (m_n^* and m_p^* are the effective masses of an electron and hole, and μ_n and μ_p are the carrier mobilities). Therefore, the dependences $\rho(T)$ in the coordinates $\log(\rho)$–$10^3/T$ are represented by straight lines, whose slopes yield the activation energy ΔE. Thus, the estimated activation energy ΔE and the value of ΔE obtained from the dependences $\log(RT^{3/2}) = f(10^3/T)$ for various samples were 0.75–0.95 eV. This value indicates that a deep impurity level plays the main role in the charge transport.

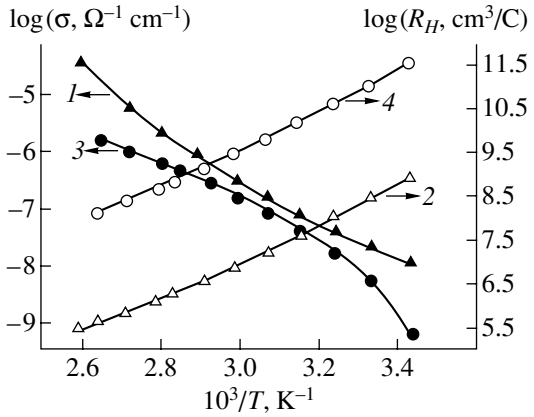

Fig. 1. Temperature dependences of the conductivity σ and Hall coefficient R for samples cut from various parts of the CdTe:(V, Ge) single crystal. $N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 5 \times 10^{18}$ cm^{-3}. Curves *1* and *2* correspond to the middle part of the crystal, and curves *3* and *4*, to the upper part.

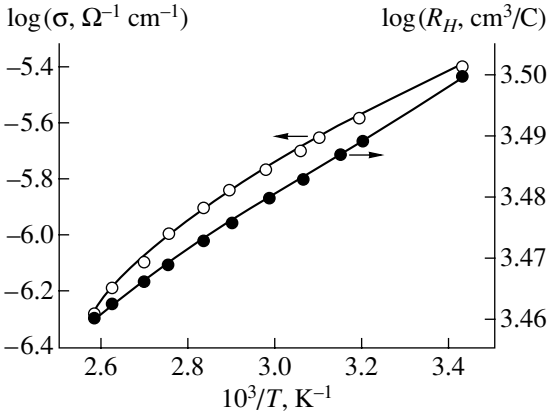

Fig. 2. Temperature dependences of the conductivity σ and Hall coefficient R for the sample cut from the top of the CdTe:(V, Ge) single crystal. $N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 5 \times 10^{18}$ cm^{-3}.

The top part of the crystal had electrical characteristics typical of a semimetal (Fig. 2). This finding can be

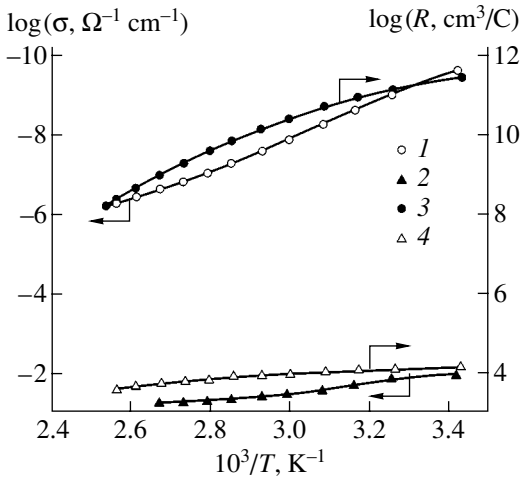

Fig. 3. Temperature dependences of the (*1*, *2*) conductivity and (*3*, *4*) Hall coefficient of the CdTe:(V, Ge) crystals: (*1*, *3*) prior to the heat treatment, and (*2*, *4*) after the heat treatment.

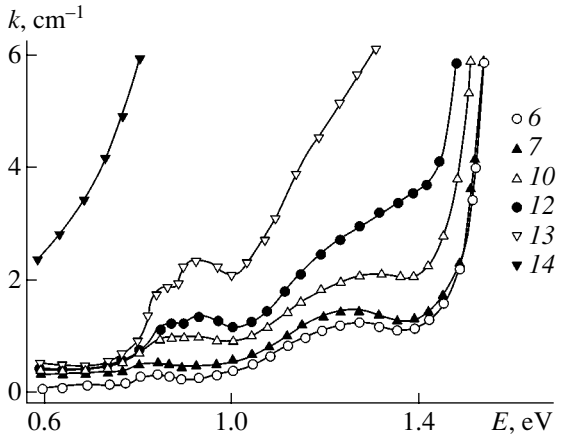

Fig. 4. Absorption spectra at T = 78 K of single-crystal wafers cut from various parts of the CdTe:(V, Ge) crystal. $N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 1 \times 10^{19}$ cm^{-3}. The numbers correspond to the location of the wafers in the crystal from the growth onset.

accounted for by the segregation of foreign impurities, for example, copper, in this part of the crystal during growth.

Figure 3 shows the temperature dependences $\sigma(T)$ and $R(T)$ for the CdTe:(V, Ge) samples ($N_V = 1 \times 10^{19}$ cm^{-3} and $N_{Ge} = 5 \times 10^{18}$ cm^{-3}) both prior to and after the heat treatment in Cd vapor. It can be seen from Fig. 3 that the conductivity σ and carrier concentration n prior to the thermal treatment were 10^{-10} Ω^{-1} cm^{-1} and 10^6–10^7 cm^{-3}, respectively. After the heat treatment and subsequent rapid cooling, the value of σ increased by four orders of magnitude and the concentration n increased to 10^{14}–10^{15} cm^{-3}. The effect of this heat treatment of CdTe:(V, Ge), in contrast to CdTe:Ge,

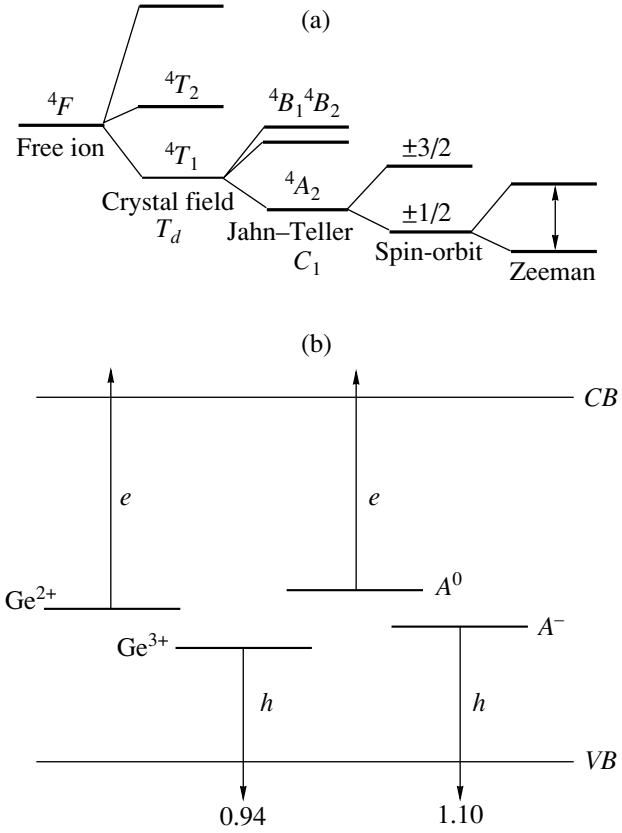

Fig. 5. (a) Splitting of energy levels of the V^{2+} ion in the CdTe lattice [19]. (b) A schematic energy diagram of the levels introduced by the Ge impurity into the CdTe band gap [10].

indicates that rapid cooling from the annealing temperature affects the defect system in such a way that various complexes, which were formed during growth, decompose.

In order to determine the energy of the optical ionization of a level located deep in the band gap, band-to-level or level-to-band transitions are conventionally used.

The results of optical studies carried out at 78 K for samples cut from various parts of the crystal are shown in Fig. 4. Several bands are observed in the absorption spectra. The intensity of these bands increases along the direction of the crystal growth. In order to identify these bands, let us consider a schematic diagram of the splitting of the energy levels of a V^{2+} ion by the crystal field and a schematic diagram of the levels in the band gap of CdTe:Ge [10, 19] (Figs. 5a, 5b).

Thus, it can be seen that an increase in absorption at 0.85 eV corresponds to an intracenter transition of V^{2+} ions between the states $4T_1(F)$ and $^4A_2(F)$ [20, 21]. The absorption band at 0.94 eV corresponds to the photoneutralization of the level of ionized Ge^+ donors (paramagnetic Ge^{3+} ion) by valence band electrons [11]. The broad band in the energy range 1.1–1.4 eV may be due to the mutual effect of the ionized states of V and Ge.

Two absorption bands, whose peak energies were equal to 1.1 and 1.22 eV, have been observed previously in CdTe:Ge crystals [11, 12]. It is assumed that the former band is related to the existence of an A center (a complex consisting of a donor and a cadmium vacancy) and its neutralization due to the capture of a free hole, which is formed during the neutralization of Ge^+. The second peak at 1.22 eV corresponds to the electron excitation from the level A^- into the conduction band. Due to this process, the density of states A^0 increases. The electrons generated under the effect of radiation are captured by the Ge^+ ions, which then transform into the neutral state Ge^0. In addition, a broad absorption band, which corresponds to the intracenter transition between the states $4T_1(F)$ and $^4T_1(P)$ of the V^{2+} ions and their subsequent self-ionization, was observed in the vicinity of 1.22 eV. A peak at 1.35 eV is observed for the CdTe:Ge crystals. This absorption band corresponds to electron excitation from the level corresponding to the neutral state Ge^0 (diamagnetic state Ge^{2+}) into the state of ionized donors Ge^+.

Thus, peaks associated with the impurity levels are revealed in the low-temperature absorption spectra in the lower-energy region. These levels are attributed to V and Ge ions in various charge states. The ionization of these ions contributes to the total photoconductivity.

4. CONCLUSIONS

Codoping of CdTe with V and Ge impurities yields compensated p-CdTe. The electrical and optical characteristics of the samples studied, which were cut from various parts of the crystal, are, excluding the uppermost part of the crystal, indicative of high material homogeneity.

Peaks attributed to the impurity levels are observed in the low-temperature absorption spectra in the lower-energy region. These levels are induced by V and Ge ions in various charge states.

The heat treatment of CdTe:(V, Ge) crystals (N_V = 1×10^{19} cm^{-3} and N_{Ge} = 5×10^{18}, 1×10^{19} cm^{-3}) in Cd vapor with a subsequent rapid cooling affects the defect subsystem of the crystals. Various complexes, which were formed during crystal growth, decompose, which leads to an increase in the conductivity and concentration of charge carriers.

REFERENCES

1. D. D. Nolte, *Photorefractive Effects and Materials* (Kluwer Academic, Boston, 1995).
2. C. Gu, Y. Xu, Y. Liu, *et al.*, Opt. Mater. **23**, 219 (2003).
3. J. Kreissl and H.-J. Schulz, J. Cryst. Growth **161**, 239 (1996).
4. H.-J. Schulz, Mater. Chem. Phys. **15** (5–6), 373 (1987).
5. M. Gauneau, R. Volle, G. Martel, and J. Y. Moisan, Opt. Mater. **7**, 21 (1997).
6. A. Partovi, J. Millerd, E. M. Garmire, *et al.*, Appl. Phys. Lett. **57**, 846 (1990).
7. Kh. Allachen, M. Tapiero, Z. Guellil, *et al.*, J. Cryst. Growth **184–185**, 1142 (1998).
8. A. Zerrai, G. Marrakachi, G. Bremond, *et al.*, J. Cryst. Growth **161**, 264 (1996).
9. S. Yu. Paranchich, L. D. Paranchich, V. N. Makogonenko, *et al.*, Zh. Prikl. Spektrosk. **70**, 381 (2003).
10. B. Briat, K. Shcherbin, B. Farid, and F. Ramaz, Opt. Commun. **156**, 337 (1998).
11. B. Briat, F. Ramaz, B. Farid, *et al.*, J. Cryst. Growth **197**, 724 (1999).
12. K. Shcherbin, S. Odulov, F. Ramaz, *et al.*, Opt. Mater. **18**, 151 (2001).
13. V. V. Matlak, K. D. Tovstyuk, A. V. Savitskiĭ, and E. S. Nikonyuk, Fiz. Tekh. Poluprovodn. (Leningrad) **6**, 2065 (1972) [Sov. Phys. Semicond. **6**, 1760 (1972)].
14. L. P. Shcherbak, E. S. Nikonyuk, O. É. Panchuk, *et al.*, Neorg. Mater. **13**, 415 (1977).
15. O. Panchuk, A. Savitsky, P. Fochuk, *et al.*, J. Cryst. Growth **197**, 607 (1999).
16. V. P. Zayachkivskiĭ, A. V. Savitskiĭ, E. S. Nikonyuk, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **8**, 1036 (1974) [Sov. Phys. Semicond. **8**, 675 (1974)].
17. A. G. Kuksov, Prib. Tekh. Éksp., No. 3, 132 (1987).
18. *Physics and Chemistry of II–VI Compounds*, Ed. by M. Aven and J. S. Prener (North-Holland, Amsterdam, 1967; Mir, Moscow, 1970).
19. P. Christman, J. Kreissl, D. M. Hofmann, *et al.*, J. Cryst. Growth **161**, 259 (1996).
20. Yu. P. Gnatenko, I. O. Faryna, P. M. Bakivskij, *et al.*, J. Phys.: Condens. Matter **14**, 7027 (2002).
21. Yu. P. Gnatenko, R. V. Gamernik, I. A. Farina, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 1975 (1996) [Semiconductors **30**, 1027 (1996)].

Translated by N. Korovin

Semiconductors, Vol. 39, No. 6, 2005, pp. 716–721. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 748–753.

PHYSICS OF SEMICONDUCTOR DEVICES

A New Memory Element Based on Silicon Nanoclusters in a ZrO_2 Insulator with a High Permittivity for Electrically Erasable Read-Only Memory

V. A. Gritsenko*^, K. A. Nasyrov, D. V. Gritsenko*, Yu. N. Novikov*, A. L. Aseev*, J. H. Lee***, J.-W. Lee***, and C. W. Kim*****

**Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
^e-mail: grits@isp.nsc.ru
***Institute of Automation and Electrometry, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
****Samsung Advanced Institute of Technology, P.O. Box 111, Suwon 440-600, Republic of Korea*
Submitted August 11, 2004; accepted for publication October 14, 2004

Abstract—The write and erase function and the data retention characteristics of a memory element designed to be used in electrically erasable read-only memory and based on a silicon–oxide–(silicon dot)–oxide–polysilicon structure, in which either a SiO_2 insulator or a ZrO_2 high-permittivity insulator are used as blocking oxides, are simulated. It is established that the use of the high-permittivity insulator gives rise to a number of effects: spurious injection from *poly*-Si is reduced; the electric field in the tunneling oxide increases; it becomes possible to increase the thickness of the tunneling insulator and, consequently, to increase the data retention time; and lower voltages for the write and erase functions can be used. Programming with a pulse of ±11 V possessing a width of 10 ms makes it possible to retain a memory window of ~3 V for 10 years. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

At present, the market for electrically alterable read-only memory (EAROM or flash memory) devices exceeds the markets for random-access memory devices and microprocessors taken together [1]. EAROM devices with a floating gate are prevalent in the nonmilitary market [2]. For custom applications, radiation-resistant silicon–oxide–nitride–oxide–silicon (SONOS) devices based on the memory effect in silicon nitride are used [3]. It is predicted that memory elements with a floating gate will be superseded by the SONOS structures in the development of terabit (10^{12} bits per crystal) EAROM devices [4].

One disadvantage of EAROM devices with a floating gate is that the bit transfer rate cannot be increased from the gigabit range to the terabit range. At present, a tunneling oxide with a thickness of ~7.0 nm and located between the silicon substrate and the floating gate is used in gigabit-capacity EAROM devices. A decrease in the channel length is accompanied by a decrease in the tunneling-oxide thickness. However, a further decrease in the thickness of the tunneling oxide leads to a rapid drain of electrons from polysilicon to the substrate because of the resulting degradation (increase) of the tunneling-oxide conductivity during reprogramming (alternation of pulses with different polarities at the gate). As a result of this reprogramming, traps are formed in the tunneling insulator and, as a consequence, the leakage current (stress-induced leakage current (SILC)) increases [5]. The leakage current in the tunneling oxide (SILC) gives rise to the problem of information retention (retention of the charge in the floating gate), which is required for ten years at 85°C for civil applications and at 125°C for special applications.

At present, EAROM devices based on the conducting nanoclusters in an insulator are actively being developed for terabit memory elements [6–12]. Semiconductor (Si, Ge, Si_xGe_y) or metal clusters with sizes of 1–10 nm act as the memory medium in these EAROM devices. A major advantage of EAROM devices based on nanocrystals in comparison with those based on a floating gate is that the nanoclusters are isolated from each other in the direction parallel to the silicon–insulator interface. As a result, defects (voids) in the tunneling insulator do not give rise to a drain of the complete charge accumulated in these nanoclusters.

Silicon dioxide SiO_2 [6–11], silicon nitride Si_3N_4 [10], and hafnium oxide HfO_2 [11, 12] are the typically used tunneling insulators in memory devices. The blocking insulator between the nanoclusters and the gate is typically fabricated of SiO_2 [6–11]. Since silicon dioxide exhibits a low relative permittivity ($\varepsilon \approx 3.9$), a high electric field (and, consequently, a large spurious voltage drop) appears in the blocking insulator in the course of writing or erasing information (reprogramming). A spurious voltage drop across the blocking SiO_2 layer leads to an undesirable increase in the voltage required for this reprogramming process. In addition, a high electric field in the blocking oxide can give rise to a spurious injection of electrons and holes from the poly-

silicon gate and, thus, to a decrease in the memory window (in the difference between the threshold voltages of the memory element in the logic states "0" and "1").

Recently, great efforts have been made to replace silicon dioxide (SiO_2) and oxynitride (SiO_xN_y) in complementary metal–oxide–semiconductor (MOS) transistors with alternative insulators known as high-*k* insulators, which have a high permittivity [13–15]. The most promising alternative insulators are believed to be ZrO_2, HfO_2 ($\varepsilon \approx 25$), Y_2O_3 ($\varepsilon \approx 15$), and Al_2O_3 ($\varepsilon \approx 10$). It has been shown both theoretically and experimentally that the replacement of the blocking SiO_2 insulator in a SONOS memory element with an alternative insulator (for example, ZrO_2 [16–18] or Al_2O_3 [19]) reduces the voltage drop across the blocking insulator and, consequently, lowers the reprogramming voltage in a SONOS EAROM device. The aim of this study was to gain insight into the data retention in EAROM devices based on a silicon–oxide–(silicon dot)–oxide–silicon (SODOS) structure by simulating the processes of reprogramming and retention (spreading) of the charge. Silicon dioxide and zirconium dioxide were used as the blocking insulators.

2. ENERGY DIAGRAMS OF SODOS STRUCTURES WITH DIFFERENT BLOCKING INSULATORS

In Fig. 1, we show energy diagrams of SODOS structures with SiO_2 (on the left) and ZrO_2 (on the right) oxides used as the blocking insulators without applied voltage and with negative or positive potentials applied to polysilicon. The barrier heights for electrons and holes at the Si/SiO_2 interface are 3.14 and 3.8 eV, respectively [20, 21]. The published data on the barrier height for electrons at the Si/ZrO_2 interface differ: 1.95 ± 0.08 eV [22], 2.0 eV [23], and 1.23 eV [24]. The values of 5.4 eV [23] and 5.5 eV [24] have been reported for the band gap of ZrO_2. In this study, we used the value of 5.5 eV in our calculations. Values of 2.0 and 2.4 eV were used for the barrier height for electrons and holes, respectively. Flat-band voltage U_{FB}, which can be measured in a memory element, was used as a control parameter. This voltage is applied to the gate of a memory structure to ensure that the electric field in silicon at the silicon/insulator interface vanishes.

3. THEORETICAL MODEL

In this study, we simulated the memory properties of SODOS structures with a blocking layer that consisted of either silicon dioxide or the alternative ZrO_2 insulator. An intrinsic semiconductor was used for the Si dots (Si nanoclusters). The acceptor concentration in both the silicon substrate and the polysilicon gate was equal to 2×10^{14} cm^{-3}. The thickness of the tunneling oxide was varied in the range from 1.5 to 5.0 nm. The size of the silicon nanoclusters was fixed in all cases and was equal to 5.0 nm. We disregarded both the quantization

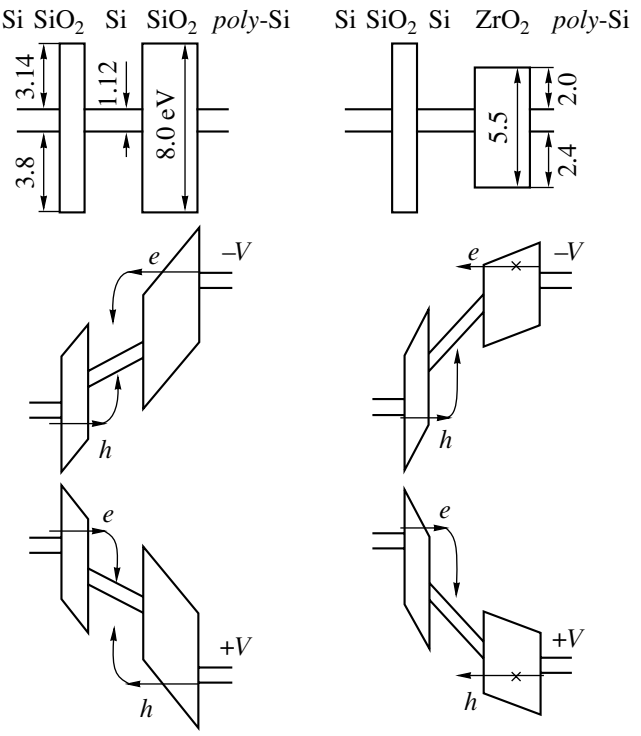

Fig. 1. An energy diagram of SODOS structures with a blocking insulator composed of SiO_2 (on the left) and ZrO_2 (on the right): (at the top) without an applied voltage, (in the middle) with a negative voltage applied to the gate, and (at the bottom) with a positive voltage applied to the gate. The energies are expressed in electronvolts. For simplicity, the voltage drops across the silicon substrate and the silicon gate are not shown.

of the electron and hole spectra, as well as the effect of the Coulomb blockade in the Si dots. The thickness of the blocking SiO_2 layer was 5.0 nm, and that of the blocking ZrO_2 layer was 8.0 nm. We used a one-dimensional two-band model in which electron injection from the negatively biased electrode and hole injection from the positively biased electrode were taken into account (Fig. 1).

We used a modified Fowler–Nordheim formula in calculations of the tunneling current. If the fields and insulator thickness are such that $F_{ox}d_{ox} > \Phi$ (here, d_{ox} is the insulator thickness, Φ is the height of the potential barrier at the Si–insulator interface, and F_{ox} is the electric field in the insulator), the tunneling occurs through a triangular barrier; as a result,

$$j = AF_{ox}^2 \exp\left\{-\frac{4}{3}\frac{\sqrt{2m^*}}{\hbar e}\frac{\Phi^{3/2}}{F_{ox}}\right\}. \quad (1)$$

Here, $A = 2.2 \times 10^{-6}$ A/V^2 and e is the elementary charge. We assumed that the tunneling effective masses m^* for electrons and holes in SiO_2 and ZrO_2 were equal to $0.5m_e$ [20].

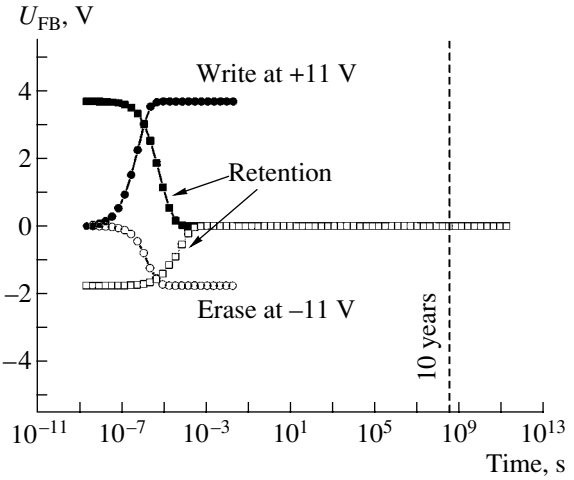

Fig. 2. Characteristics of the write and erase processes in a SODOS structure in which 5.0-nm-thick SiO_2 is used as the blocking insulator. The filled circles represent the accumulation of negative charge, and the open circles represent the accumulation of positive charge. The amplitudes of the write and erase pulses are +11 and –11 V, respectively, and the duration of the pulses is 10 ms. The filled and open squares illustrate the process of charge drain in the shorted state. The Si nanoclusters are 5.0 nm thick, and the tunneling-oxide thickness is 1.5 nm.

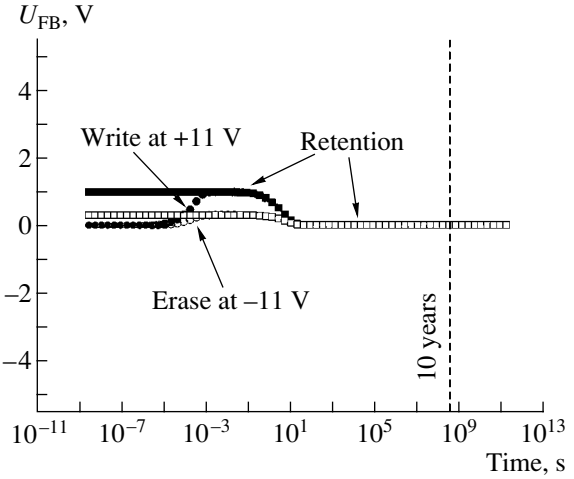

Fig. 3. Characteristics of the write and erase processes in a SODOS structure where 5.0-nm-thick SiO_2 is used as the blocking insulator. The filled and open circles represent the accumulation of negative and positive charges, respectively. The amplitudes of the write and erase pulses are +11 and –11 V, respectively; the duration of the pulses is 10 ms. The filled and open squares illustrate the process of charge drain in the shortened state. The thickness of the Si nanoclusters is 5.0 nm, and the tunneling-oxide thickness is 2.5 nm.

In our calculations, we used the following formula for the tunneling current in the case of a trapezoidal barrier ($F_{ox}d_{ox} < \Phi$):

$$j = AF_{ox}^2 \exp\left\{-\frac{4}{3}\frac{\sqrt{2m^*}}{\hbar e}\frac{(\Phi^{3/2} - (\Phi - F_{ox}d_{ox})^{3/2})}{F_{ox}}\right\}. \quad (2)$$

4. COMPARISON OF THE MEMORY CHARACTERISTICS OF SODOS STRUCTURES WITH SiO_2 AND ZrO_2 AS THE BLOCKING INSULATORS

We carried out the first simulation of the write and erase process and the retention of charge in SODOS structures with the SiO_2 blocking insulator at a tunneling-oxide thickness of 1.5 nm (Fig. 2). The duration of the write and erase pulses was 10 ms in all cases. The positive potential +11 V at the polysilicon gate gave rise to the injection of electrons from the silicon substrate through the tunneling oxide into the Si nanoclusters, the accumulation of negative charge, and a flat-band voltage shift to +4 V. In this case, the memory window amounted to +6 V. In the write and erase mode, we observed a spurious injection of charge through the blocking oxide in accordance with the Fowler–Nordheim mechanism. The drain of negative and positive charges from the Si nanoclusters in the information-retention mode (shortened state) occurred in 10^{-5} s. The drain of electrons and holes from a nanocluster occurs owing to tunneling through the tunneling oxide to the silicon substrate. The discharge due to the tunneling of charge carriers through the blocking oxide into the gate is negligible as a result of the large thickness of this oxide.

When attempting to increase the information-retention time to ten years ($\sim 3 \times 10^8$ s), the most obvious step is to increase the tunneling-oxide thickness so as to suppress the electron and hole tunneling from an Si dot to the substrate. However, an increase in the tunneling-oxide thickness to 2.5 nm at the reprogramming-pulse amplitude of ±11 V leads to a decrease in the memory window to ~1 V (Fig. 3). The 2.5-nm-thick tunneling oxide sharply reduces the electron injection and completely suppresses the hole injection from the substrate. An erase pulse with an amplitude of –11 V does not result in the accumulation of positive charge; however, an insignificant accumulation of negative charge is observed. This behavior is caused by spurious electron injection from the polysilicon gate. An increase in the tunneling-oxide thickness to 2.5 nm results in a slowing down of the charge drain and to an increase in the retention time to ~10 s. However, this value is more than eight orders of magnitude smaller compared to the information-retention time required of an EAROM (ten years). A further increase in the tunneling-oxide thickness is accompanied by a drastic decrease in the memory window, due to an exponential decrease in the injection current of electrons and holes into the Si nanoclusters.

Figure 4 illustrates the dependence of the memory window on the tunneling-oxide thickness in the write and erase mode for voltages of ±8 and ±11 V. It is worth noting that the memory window increases by 2–3 V as the write (erase) voltage decreases from 11 to 8 V. This effect is related to a decrease in the spurious injection through the blocking insulator. It can be seen from

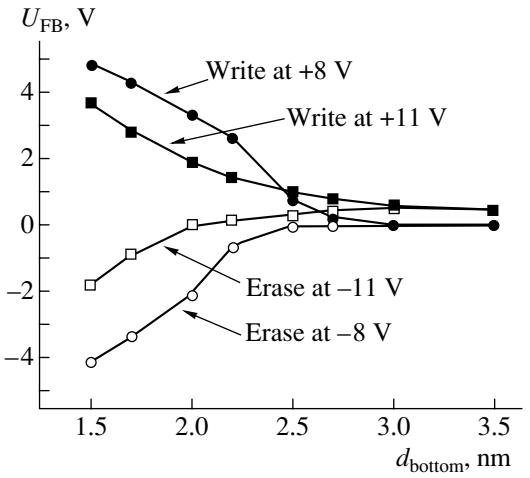

Fig. 4. The dependence of the memory window in a SODOS structure on the tunneling-oxide thickness in the write mode. A 5.0-nm-thick SiO_2 layer was used as the blocking insulator. Voltage pulses with amplitudes of +8 V (filled circles) and +11 V (filled squares) were used for writing and pulses with amplitudes of –8 V (open circles) and –11 V (close squares) were used for erasing. The duration of the pulses was 10 ms. The tunneling-oxide thickness was varied from 1.5 to 3.5 nm.

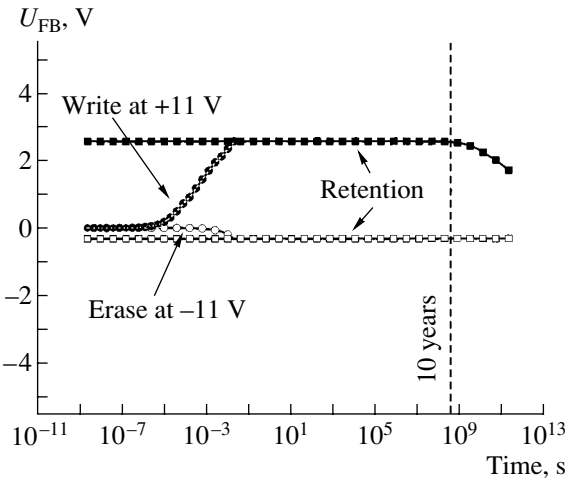

Fig. 5. Characteristics of the write and erase processes in a SODOS structure in which 8.0-nm-thick ZrO_2 is used as the blocking insulator. The filled and open circles represent the accumulation of negative and positive charge, respectively. Voltage pulses with a duration of 10 ms and amplitudes of +11 V (–11 V) were used for the write (erase) operations. The filled and open squares illustrate the process of charge drain in the shortened state. The Si nanocluster thickness was 5.0 nm, and the SiO_2 tunneling-oxide thickness was 5.0 nm.

Fig. 4 that the memory window decreases to less than 1 V if the tunneling-oxide thickness exceeds 2.5 nm.

Thus, the analysis shows that the memory effect (accumulation of electrons and/or holes in SODOS structures) manifests itself only in a situation where the thickness of the bottom oxide is smaller than that of the top oxide. In this case, a charge is accumulated in the silicon dots owing to the higher current of the direct tunneling through the bottom oxide. If the values of the thickness of the bottom and blocking oxide layers are equal to each other and exceed the tunneling length in the insulator, the SODOS structure does not exhibit a memory effect. This circumstance is related to the fact that the current of the electron injection from the substrate into the silicon dots is equal to the current of the electron injection from the silicon dots into the gate.

Replacement of the blocking layer made of SiO_2 by one made of ZrO_2 makes it possible to obtain a flat-band shift of +2.5 V at a tunneling-oxide thickness of 5.0 nm and programming-pulse amplitude of +11 V (Fig. 5). An erase pulse with an amplitude of –11 V leads to only an insignificant accumulation of positive charge (–0.4 V). However, in contrast to SiO_2, a spurious injection of electrons is not observed when ZrO_2 is used as the blocking insulator. The use of the ZrO_2 insulator and a relatively thick tunneling oxide makes it possible to greatly slow down the drain of electrons and obtain a memory window of 3 V even after ten years (Fig. 5).

In Fig. 6, we show the dependence of the memory window on the tunneling-insulator thickness for a

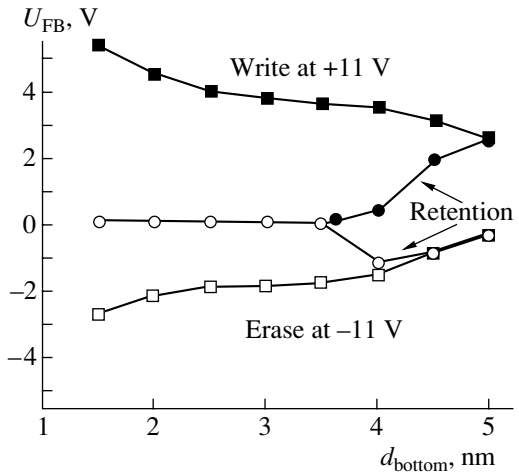

Fig. 6. The dependence of the memory window in a SODOS structure on the thickness of the tunneling (bottom) insulator in the writing and retention modes. A 8.0-nm-thick ZrO_2 layer was used as the blocking insulator. Voltage pulses with a duration of 10 ms and amplitudes of +11 V (filled squares) and –11 V (open squares) were used for the write and erase operations. The tunneling-oxide thickness was varied from 1.5 to 5.0 nm. The memory window (the filled and open circles) is shown after ten years of storage.

SODOS structure with the ZrO_2 blocking insulator. The amplitudes of the write and erase pulses were +11 V and –11 V, respectively. An increase in the tunneling-oxide thickness from 1.5 to 5.0 nm leads to a decrease in the memory window from 8 to 3 V. The memory window for a SODOS structure charged preliminarily with electrons or holes after ten years is also shown in Fig. 6.

The measured memory window is observed for ten years if the tunneling-oxide thickness exceeds 3.5 nm (Fig. 6).

5. DISCUSSION

The replacement of a blocking insulator composed of SiO_2 by one composed of ZrO_2 with a higher permittivity in memory devices based on nanoclusters leads to two effects: (i) the electric field in the tunneling insulator increases and (ii) the field in the blocking insulator decreases. An increase in the field in the tunneling insulator gives rise to an exponential increase in the injection current of electrons and holes from silicon into the nanoclusters, which leads to an increase in the memory window. In addition, it becomes possible to reduce the amplitude and/or the duration of the programming pulse for a fixed memory window. The effect of an increase in the field in the tunneling oxide makes it possible to increase the thickness of this oxide and, thus, suppress the charge drain into the silicon substrate in the mode of information retention.

At present, memory devices based on nanoclusters and insulators with a high permittivity are being studied extensively. For example, Lee *et al.* [12] studied devices with silicon nanoclusters and with HfO_2 used as both the tunneling insulator and the blocking insulator. Similar devices with Si_xGe_{1-x} nanoclusters were studied by Kim *et al.* [11]. However, the effect of an increase in the electric field in the tunneling oxide in the aforementioned structures was not observed. It is noteworthy that the first memory device based on silicon nanoclusters deposited on silicon oxide and coated with silicon nitride was apparently fabricated by Belyĭ *et al.* [25]. In this device [25], the effect of an increase in the field in the tunneling oxide was observed due to the relatively high permittivity of Si_3N_4 ($\varepsilon = 7$) (to be compared with $\varepsilon = 3.9$ in SiO_2).

In the model used in this study, it is assumed that the charge-carrier injection through an insulator with a high permittivity is limited by the Fowler–Nordheim tunneling mechanism. However, in actual insulators with a high permittivity (for example, in Al_2O_3 or HfO_2), there always is a high concentration of traps; therefore, the electrical conductivity is limited by the ionization of defects according to either the Frenkel model or the multiphonon mechanism. Thus, in order to design the devices we have suggested properly, it is necessary to develop the technology of alternative insulators with a low concentration of traps and low leakage currents.

6. CONCLUSIONS

Using the example of ZrO_2, we showed that the application of alternative insulators in the blocking layer in silicon–oxide–(silicon dot)–oxide–silicon (SODOS) structures leads to a number of advantages. These consist in the following:

(i) The electric field in the tunneling oxide increases and, as a result, there is increase in the injection current. This effect makes it possible to use a thicker tunneling insulator at a fixed memory window, which enables an increase in the charge-retention time.

(ii) Lower write and erase voltages can be used for an unchanged SODOS configuration.

(iii) The speed of the response can be increased without changing either the SODOS configuration or the write-and-erase voltages.

(iv) A decrease in the voltage drop across the blocking insulator leads to a reduction in the spurious injection of electrons and holes from the polysilicon gate into the write or erase modes.

The simulation of the charge-retention mode showed that the retention time of the charge is no longer than 10 s if SiO_2 is used as the blocking insulator (for any thickness of the blocking and tunneling oxides). If 8.0-nm-thick ZrO_2 is used as the blocking insulator and 5-nm-thick SiO_2 is used as the tunneling insulator, it becomes possible to attain a memory window of ~3 V, even after ten years at room temperature.

REFERENCES

1. S. M. Sze, in *Future Trends in Microelectronics*, Ed. by S. Luryi, J. Xu, and A. Zaslavsky (Wiley, New York, 1999), p. 135.
2. P. Pavan, R. Bez, P. Olivo, and E. Zanony, Proc. IEEE **85**, 1248 (1997).
3. J. Bu and M. H. White, Solid-State Electron. **45**, 113 (2001).
4. *International Technology Roadmap*, http://public.itrs.net/ (2003).
5. K. Komiya and Y. Omura, J. Appl. Phys. **92**, 2953 (2002).
6. Y. I. Hanafi, S. Tiwari, and I. Khan, IEEE Trans. Electron Devices **43**, 1553 (1996).
7. Y.-C. King, T.-J. King, and C. Hu, IEEE Electron Device Lett. **20**, 409 (1999).
8. B. De Salvo, G. Gibaudo, G. Pananakakis, *et al.*, IEEE Trans. Electron Devices **48**, 1789 (2001).
9. Z. Liu, C. Lee, V. Narayanan, *et al.*, IEEE Trans. Electron Devices **49**, 1614 (2002).
10. M. She and T.-J. King, IEEE Trans. Electron Devices **50**, 1934 (2003).
11. D.-W. Kim, T. Kim, and S. K. Banerjee, IEEE Trans. Electron Devices **50**, 1823 (2003).
12. J. J. Lee, X. Wang, W. Bai, *et al.*, IEEE Trans. Electron Devices **50**, 2067 (2003).
13. G. D. Wilk, R. M. Wallace, and J. M. Anthony, J. Appl. Phys. **89**, 5243 (2001).
14. E. P. Gusev, E. Cartier, D. A. Buchanan, *et al.*, Microelectron. Eng. **59**, 341 (2001).
15. G. Ya. Krasnikov, *Structural and Technological Features of Submicron MOS Transistors* (Tekhnosfera, Moscow, 2002) [in Russian].

16. V. A. Gritsenko, K. A. Nasyrov, and Yu. N. Novikov, in *Proceedings of 12th Workshop on Dielectrics in Microelectronics (WODIM)* (Grenoble, France, 2002), p. 179.
17. V. A. Gritsenko, K. A. Nasyrov, Yu. N. Novikov, and A. L. Aseev, Mikroélektronika **32** (2), 69 (2003) [Russ. Microelectronics **32**, 69 (2003)].
18. V. A. Gritsenko, K. A. Nasyrov, Yu. N. Novikov, *et al.*, Solid-State Electron. **47**, 1651 (2003).
19. C. Lee, S. Hur, Y. Shin, *et al.*, in *Abstracts of 2002 International Conference on Solid State Devices and Materials* (Nagoya, Japan, 2002), p. 162.
20. V. A. Gritsenko, E. E. Meerson, and Yu. N. Morokov, Phys. Rev. B **57**, R2081 (1997).
21. V. A. Gritsenko, *Design and Electronic Properties of Amorphous Insulators in Silicon MIS Structures* (Nauka, Novosibirsk, 1993) [in Russian].
22. V. V. Afanas'ev, M. Houssa, A. Stesmans, *et al.*, Microelectron. Eng. **59**, 335 (2001).
23. M. Houssa, M. Tuominen, M. Naili, *et al.*, J. Appl. Phys. **87**, 8615 (2000).
24. S. Miyazaki, M. Narasaki, M. Ogasawara, and M. Hirose, Microelectron. Eng. **59**, 373 (2001).
25. V. I. Belyĭ, V. V. Voskoboĭnikov, A. S. Ginovker, *et al.*, Mikroélektronika **2**, 182 (1971).

Translated by A. Spitsyn

Semiconductors, Vol. 39, No. 6, 2005, pp. 722–729. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 754–761.

PHYSICS OF SEMICONDUCTOR DEVICES

Special Features of Charge Transport in Schottky Diodes Based on Semi-insulating CdTe

L. A. Kosyachenko^, O. L. Maslyanchuk, and V. M. Sklyarchuk

Chernovtsy National University, Chernovtsy, 58012 Ukraine

^e-mail: lakos@chv.ukrpack.net

Submitted October 4, 2004; accepted for publication November 10, 2004

Abstract—The electrical characteristics of CdTe-based Schottky-diode detectors of X-ray and γ radiation are studied. Experimental data are obtained for Al/*p*-CdTe diodes with a substrate resistivity from 10^2 to 10^9 Ω cm (300 K). The obtained results are interpreted in the context of the Sah–Noyce–Shockley theory of generation–recombination, taking into account the special features of the Schottky diode. It is shown that, when semi-insulating CdTe is used, the considerable forward currents observed are caused by electron injection into the substrate. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

For several decades, CdTe has been an important material in semiconductor detectors, which are widely used in science and engineering, medicine, and other fields. Due to the larger atomic numbers of CdTe as compared with Si (48/52 for CdTe and 14 for Si), the spectral range of CdTe-based detectors is extended to a higher-energy region of detected photons, and its wider band gap (1.5 eV) provides for operation without cryogenic cooling. However, designers of CdTe-based detectors come up against numerous problems associated with obtaining homogeneous semi-insulating single crystals with conductivity close to the intrinsic value and a prolonged carrier lifetime.

As early as in the 1960s, it was shown that it was possible to design surface-barrier CdTe-based spectrometry-type detectors of γ-ray radiation operating at room temperature [1]. These findings were repeatedly confirmed in the following years [2–4]. However, these studies lost their importance after start of the commercial fabrication of detectors based on homogeneous CdTe single crystals and, then, $Cd_{1-x}Zn_xTe$ alloys ($x \approx 0.1$) with characteristics acceptable for practical usage. For a single crystal 1–5 mm thick, the spectral characteristic of such a detector extended into the region of γ-ray photons with energies of 1 MeV or higher, while the active region (depletion layer) in detectors with a surface–barrier structure (the Schottky barrier) was no thicker than 50–100 μm, which substantially restricted their potential [2–4]. However, at the end of the 1990s, Takahashi *et al.* [5–9] showed that this limitation could be overcome due to the extremely low dark currents and favorable conditions of charge collection in a CdTe-based Schottky diode. This circumstance made it possible to fabricate CdTe-based detectors of high-energy photons (for example, 662 keV and higher) possessing a limiting energy resolution without a special circuit for processing the electrical signal in the detector circuit [10].

Despite this considerable technical progress, a series of physical issues remained unclear with respect to the properties of a Schottky diode based on a semi-insulating material. Specifically, the mechanism of the reverse current is still not satisfactorily understood. A low reverse current admits high operational voltages and, consequently, high efficiency of the charge collection. The nature of the observed considerable forward diode currents, amongst other factors, is also unclear for such a high-resistivity base. The aim of this study is to offer some clarification of these issues.

2. EXPERIMENT AND RESULTS

In order to fabricate the diodes, we used single-crystal *p*-CdTe wafers with a resistivity ρ from 10^2 to 10^9 Ω cm (300 K). The wafers were ground and polished, and then treated with $K_2Cr_2O_7 + HNO_3 + H_2O$ in several stages until the surface became uniform, which was judged with the use of a microscope. A Schottky contact of 1–3 mm^2 in area was fabricated by the vacuum (10^{-6} Torr) deposition of a 1-μm-thick Al film at a substrate temperature of 150–200°C. Nonrectifying contacts were obtained by vacuum deposition of Ni.

Figure 1 shows the current–voltage (*I–V*) characteristic of one of the diodes that were fabricated on a substrate with a resistivity of 2×10^4 Ω cm at room temperature. It is important that, at both the temperatures used (295 and 343 K), an extended portion following the dependence close to $I \propto \exp(eV/2kT) - 1$ is observed in a forward *I–V* characteristic in semilog coordinates. However, at voltages exceeding ~0.5 V, the experimental points sharply deviate towards higher voltages.

It is natural to assume that, in the region of large forward currents, a part of the applied voltage drops across

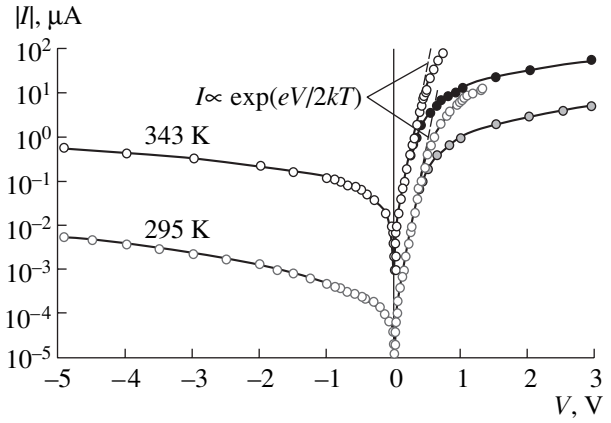

Fig. 1. Current–voltage characteristics of an Al/*p*-CdTe diode at 295 and 343 K. The substrate resistivity is 2×10^4 Ω cm at 295 K.

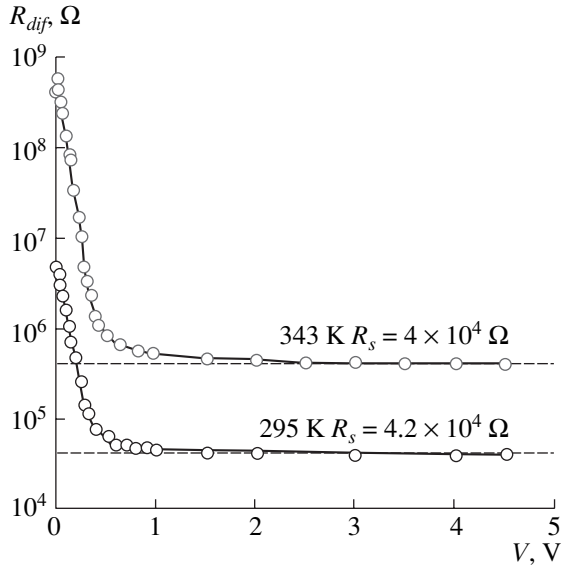

Fig. 2. Differential resistance of a diode at 295 and 343 K. The value of R_{dif} in the saturation region corresponds to the substrate resistance R_s.

the substrate resistance R_s, whose magnitude can be found from the voltage dependence of the differential resistance of a diode R_{dif} (Fig. 2). Figure 2 shows that R_{dif} rapidly decreases at low biases, which corresponds to an exponential increase in the current. However, at $V = 0.5$–2 V, the curve $R_{dif}(V)$ levels off (saturates).

The presence of saturation in the curve $R_{dif}(V)$ means that the voltage at the barrier region of a diode virtually ceases to increase, and a further increase in the voltage applied to a sample leads to an increase in the voltage drop only across the series linear substrate resistance. In other words, variation in the current ΔI in this region is associated with variation in the voltage applied to the substrate ΔV; i.e., the magnitude $\Delta V/\Delta I$, or the differential resistance at fairly low values of ΔV and ΔI, equals the substrate resistance R_s. Thus, the determined resistance R_s for the diode shown in Fig. 1 is 0.42 and 0.04 MΩ at 295 and 343 K, respectively. If we take into account the voltage drop on R_s, i.e., plot the dependence of I versus $V–IR_s$, the forward portions of the $I–V$ characteristic in the region of high currents vary (open circles in Fig. 1). Nevertheless, in terms of the voltage drop on R_s, the points in the region of a high current do not correspond to the exponential dependences $I \propto \exp(eV/2kT)$ (dashed straight lines). An explanation of this observation should be sought in the special features of the transport mechanism in the diodes under study.

Another characteristic feature of the diodes under study is the fact that the reverse current at low biases is sublinear relative to the voltage, and its magnitude at a fixed voltage is proportional to the concentration of intrinsic carriers in the semiconductor: $n_i = (N_c/N_v)^{1/2}\exp(-E_g/2kT)$ [15, 16]. Figure 3 shows such a dependence in the coordinates $I/T^{3/2}$–$1000/T$. The slope of the obtained straight line yields the energy 0.80 eV,

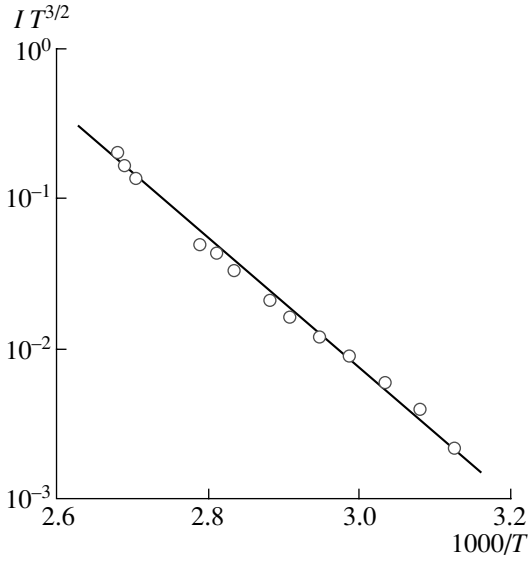

Fig. 3. Temperature dependence of the reverse current at a fixed bias voltage of 0.3 V (circles). The straight line has a slope corresponding to $E_g/2 = 0.8$ eV at 0 K.

which corresponds to half of the CdTe band gap at 0 K: $E_{g0} = E_g(T) + \gamma T = 1.46 + 4 \times 10^{-4} \times 300 = 1.58$ eV.

The above features of the $I–V$ curves are characteristic of Al/*p*-CdTe diodes, which are fabricated on crystals with a resistivity varying over the entire range under consideration ($\rho = 10^2$–10^9 Ω cm). However, as the substrate resistivity increases, the variance in the range of voltage, in which the forward current follows the dependence $I \propto \exp(eV/2kT) - 1$, narrows so that, as ρ approaches 10^9 Ω cm, the linear portion in the voltage

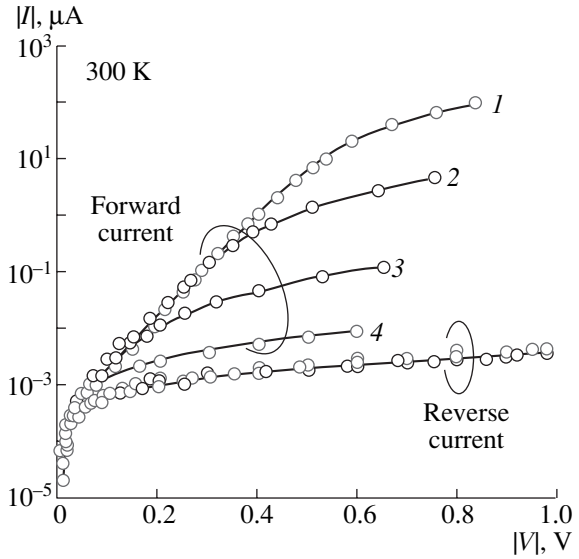

Fig. 4. Current–voltage characteristics of Al/*p*-CdTe diodes normalized for a reverse bias of 0.3 V. The diodes are fabricated on substrates with the resistivities (*1*) 3×10^2, (*2*) 2×10^4, (*3*) 2×10^7, and (*4*) 2×10^9 Ω cm.

dependence of log *I* is completely unobservable and the rectifying properties of a diode virtually vanish. The shape of the reverse portion of the *I–V* characteristic is retained in this case, and the magnitude of the reverse current varies only within the limits of an order of magnitude when ρ varies from 10^2 to 10^9 Ω cm. We can assume that this circumstance is related to the other characteristics of the material rather than to its resistivity. If we normalize the reverse current at a fixed voltage (for example, 0.3 V), the generalized *I–V* characteristic of an Al/*p*-CdTe diode corresponds to that shown in Fig. 4.

The existence of the dependence $I \propto \exp(eV/2kT)$ for the forward current and the sublinear voltage dependence of the reverse current are characteristic indications that the generation–recombination mechanism of charge transport is operating in the diodes under study. The Sah–Noyce–Shockley theory does not imply a deviation from the exponential dependence at high forward currents [11]. However, if an allowance is made for the specific features of the Schottky diode, especially the band bending in the barrier region, which is smaller than in the *p–n* junction, it is possible to explain this special feature of the *I–V* characteristics of the diodes based on low-resistivity $Cd_{1-x}Zn_xTe$ crystals ($x = 0.05$–0.1) of both *n*-type and *p*-type conductivity [12, 13]. Let us consider a description of the evolution of the *I–V* characteristic of a Schottky diode according to the Sah–Noyce–Shockley model as the semiconductor resistivity increases from 10^2–10^3 to 10^9–10^{10} Ω cm.

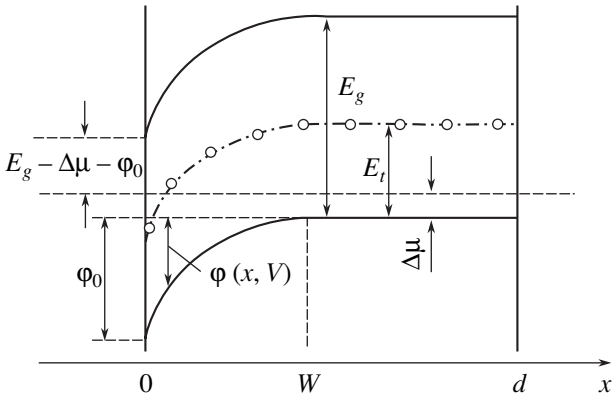

Fig. 5. Energy diagram of the diode structure under consideration.

3. THEORETICAL CONSIDERATION

According to the Sah–Noyce–Shockley theory, the current through a diode can be calculated by integration of the generation–recombination rate over the entire space-charge region (SCR) [11]:

$$I_{g-r} = Ae\int_0^W \frac{n(x,V)p(x,V) - n_i^2}{\tau_{po}[n(x,V)+n_1] + \tau_{no}[p(x,V)+p_1]}dx, \quad (1)$$

where *A* is the diode area, *e* is the elementary charge, $n(x, V)$ and $p(x, V)$ are the carrier concentrations in the bands, and τ_{no} and τ_{po} are the electron and hole lifetimes in the SCR. The values of n_1 and p_1 are equal to the equilibrium electron and hole concentrations n_0 and p_0 with the proviso that the Fermi level coincides with the level under consideration; i.e., $n_1 = N_c\exp(-E_t/kT)$ and $p_1 = N_v\exp[-(E_g - E_t)/kT]$, where $N_c = 2(m_nkT/2\pi\hbar^2)^{3/2}$ and $N_v = 2(m_pkT/2\pi\hbar^2)^{3/2}$ are the effective densities of states in the conduction and valence bands, respectively; m_n and m_p are effective masses of electrons and holes; and E_t is the energy measured from the valence band top (Fig. 5). In the chosen reference system, it is possible to write the following expressions for the electron and hole concentrations at the point *x* at a voltage *V*:

$$n(x,V) = N_c\exp\left[-\frac{\Delta\mu + \varphi(x,V)}{kT}\right], \quad (2)$$

$$p(x,V) = N_v\exp\left[-\frac{E_g - \Delta\mu - \varphi(x,V) - qV}{kT}\right]. \quad (3)$$

Here, Δμ is the energy spacing between the Fermi level and the valence-band top in the bulk of a semiconductor. The potential distribution in a Schottky diode is parabolic:

$$\varphi(x,V) = (\varphi_0 - eV)\left(1 - \frac{x}{W}\right)^2. \quad (4)$$

Here, φ_0 is the equilibrium barrier height from the semiconductor side ($\varphi_0/e = V_{bi}$ is the contact potential), the coordinate is measured from the semiconductor surface, and the energy $\varphi(x, V)$ is measured from the valence-band top downwards in the neutral part of the semiconductor.

The width of the SCR W, which enters formula (1), requires special consideration. For a Schottky diode, this quantity can be usually written as

$$W = \sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_0 - eV)}{e^2(N_a - N_d)}}, \quad (5)$$

where ε is the relative permittivity, ε_0 is the permittivity of free space, and $N_a - N_d$ is the concentration of uncompensated acceptors (we are considering a semiconductor with p-type conduction).

For a semiconductor doped with shallow-level (completely ionized) impurities, the concentration $N_a - N_d$ of uncompensated acceptors is virtually equal to the hole concentration in the valence band. However, it is known that even the purest and most perfect CdTe single crystals contain impurities (defects) at a concentration of as high as 10^{15}–10^{17} cm^{-3}. Among these defects, there are the deep-level acceptors with various ionization energies [14, 15]. In this case, it is implied that $N_a - N_d$ in formula (5) (giving W) stands for the total acceptor concentration. Actual CdTe crystals always contain donor impurities of several different types [14, 15], and the quantity $N_a - N_d$ is equal to the total acceptor concentration minus the total donor concentration. For lightly compensated CdTe at $N_a = 10^{15}$–10^{17} cm^{-3}, $\varepsilon = 10.6$, $\varphi_0 = 1$ eV, and $V = 0$, we have $W = 0.3$–3 μm. For a material compensated to a greater degree, $N_a - N_d$ becomes smaller and, correspondingly, W becomes larger. As the degree of compensation increases, the SCR can become much larger than the mentioned magnitude and can theoretically extend over an entire sample; however, this does not represent a realistic case. The solution of the electroneutrality equation shows that it is virtually impossible to obtain a semi-insulating state in a semiconductor, i.e., to keep the Fermi level close to the midgap, by compensating its shallow-level impurities [16, 17]. In order to achieve this aim, a very accurate compensation is necessary. Even with the Use of modern growth technology for CdTe single crystals, it is unattainable for any considerable volume of the crystal. If there is the smallest deviation from the necessary value for the degree of compensation, the Fermi level abruptly shifts upwards or downwards from the midgap, and the semiconductor develops low-resistivity with n-type or p-type conductivity, respectively. The problem can be solved via compensation of the impurities or the defects, whose level is located close to the midgap. In this case, it is this level that can pin the Fermi level. This pinning requires a partial rather than complete compensation of the impurity level. For a 50% compensation, the Fermi level exactly coincides with the level under consideration and slightly shifts from this position if the degree of compensation varies in a range of about 20–80% [17]. This implies that $N_a - N_d = (0.2$–$0.8)N_a$; i.e., according to formula (5), the width of the SCR increases no more than severalfold in contrast to an uncompensated semiconductor.

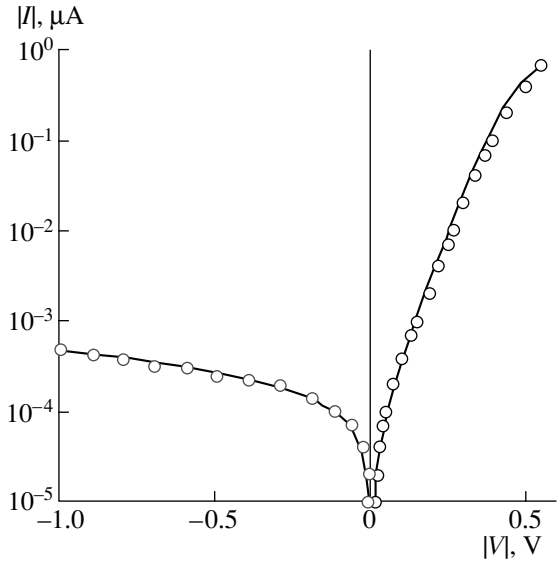

Fig. 6. Current–voltage characteristic of an Al/p-CdTe diode ($\rho = 2 \times 10^4$ Ω cm). The circles represent the measured characteristic and the solid lines correspond to the results of calculations using formula (1).

Making allowance for the above, in our further calculations, we assume that the concentration of the uncompensated acceptors is equal to the average value of the above-mentioned range, i.e., 10^{16} cm^{-3}, which corresponds to $W = 1$ μm. This assumption introduces no substantial error into the calculations. Note that, as the selected value of $N_a - N_d$ decreases or increases by an order of magnitude, the SCR is broadened or narrowed by a factor $\sqrt{10} \approx 3$.

4. RESULTS OF CALCULATIONS

Figure 6 shows a comparison of the I–V characteristic calculated using formula (1) and taking into account relations (2)–(5) with that measured for an Al/p-CdTe diode based on a crystal with the resistivity 2×10^4 Ω cm (295 K). We used the following values of the parameters: $m_n = 0.11m_0$, $m_p = 0.35m_0$ (m_0 is the free-electron mass), $E_t = 0.73$ eV, $p = (e\rho\mu_p)^{-1} = 2 \times 10^{12}$ cm^{-3}, $\mu_p = 80$ cm^2/(V s), and $\Delta\mu = kT\ln(N_v/p) = 0.374$ eV. The adjustable parameters were τ_{no} and τ_{po}, which were assumed to be equal 9×10^{-11} s to obtain the best agreement between the calculation and the experiment. The deviation from the dependence $I \propto \exp(eV/2kT)$ with an increase in V depends heavily on the barrier height φ_0. In order to obtain the required deviation from the

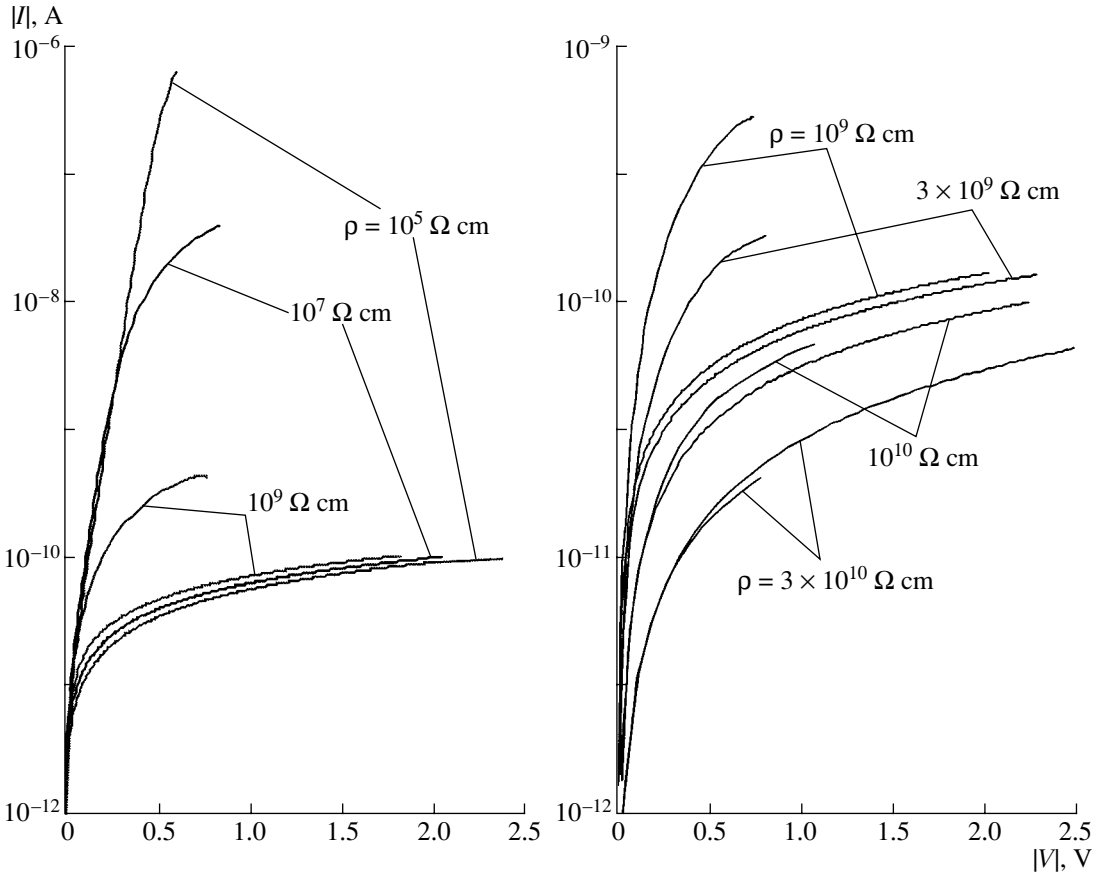

Fig. 7. Current–voltage characteristics of a CdTe-based Schottky diode calculated for various substrate resistivities ($N_a - N_d$ = 10^{16} cm^{-3}, T = 300 K).

straight line in Fig. 6 at $V > 0.4$ V, we have to assume that the value of φ_0 equals 0.52 eV. Figure 6 shows that excellent agreement between the results of the calculation and experiment is attained both for the forward and reverse biases at the selected parameters. Thus, there is no doubt that the charge is transported in the diodes under study by the generation–recombination mechanism.

Let us now consider the behavior of the *I–V* characteristic of the Schottky diode as the substrate resistivity increases, especially at high resistivity values (to the extent of conductivity close to the intrinsic value).

Figure 6 shows the *I–V* characteristic calculated using formula (1) and taking into account relations (2)–(5). In the calculation, it was assumed that, in all cases, the Fermi level at the metal–semiconductor interface is 1-eV above the valence band top; i.e., the barrier height at the contact φ_0 is 1 eV – $\Delta\mu$. A similar pinning of the Fermi level by high-density surface states is characteristic of many semiconductors [18]. The effective lifetime in the SCR was assumed to be equal $\tau = 3 \times 10^{-9}$ s for the diode area A = 10 mm^2. Since it is necessary to take into account the electron component of the current for the semi-insulating material, the Fermi level energy was determined from the specified value of the material resistivity ρ using the formula [17]

$$\Delta\mu = kT\ln\left(\frac{1-\sqrt{1-4e^2\rho^2\mu_n\mu_p n_i^2}}{2e\rho\mu_n n_i^2/N_v}\right). \qquad (6)$$

The electron mobility μ_n was assumed to be equal 1000 cm^2/(V s). The left-hand side of Fig. 7 shows the results of the calculation for values of ρ in the range 10^5–10^9 Ω cm, and the right-hand side shows more detailed results from $\rho = 10^9$ Ω cm to $\rho = 3 \times 10^{10}$ Ω cm, which practically corresponds to the intrinsic conductivity of CdTe.

Figure 7 shows that the rectifying properties of the diode worsen as the material resistivity increases, and they virtually vanish at $\rho = 3 \times 10^9$–3×10^{10} Ω cm. Another important result is that the reverse portion of the *I–V* characteristic varies only slightly over the entire range of ρ. The forward current through the diode is also independent of the resistivity at low voltages, where $I \propto \exp(eV/2kT)$. This result is in good agreement with the experimental data, which are shown, in a generalized form, in Fig. 4. It follows from the above

that the main systematic features in the behavior of the *I–V* characteristic of the Al/*p*-CdTe diodes can now be explained. (i) The reverse current, which has a generation origin, cannot vary in a wide range as the material resistivity ρ does since this current is governed by the carrier lifetime and by the thickness of the SCR, which have no direct relation with a value of ρ. (ii) In the region of low forward biases, where the dependence $I \propto \exp(eV/2kT) - 1$ holds, the current is governed by the same parameters and, therefore, is also only slightly ρ-dependent. (iii) As ρ increases, the Fermi level recedes from the valence band; i.e., $\Delta\mu$ increases at the same time as φ_0 decreases. In this case, the segment of the forward portion, where the forward current is proportional to $\exp(eV/2kT)$, is increasingly restricted from above, as is observed in the experimental curves.

5. RESISTANCE MODULATION FOR A HIGH-RESISTIVITY SUBSTRATE

It follows from the experimental data reported by Takahashi *et al.* [3–7] that the forward current through a Schottky-barrier CdTe detector fabricated based on a crystal with a resistivity of (1–3) × 10^9 Ω cm increased rapidly at a voltage higher than ~10 V and considerably exceeded the reverse current through the diode. A 0.5-mm-thick substrate with the above resistivity of its 4 × 4-mm^2 diode area has a resistance of (0.3–1) × 10^9 Ω. Therefore, the current limited by this resistance should not exceed ~10^{-8} A at a voltage of ~10 V. However, the experimental value of the forward current at 10 V is ~10^{-5}–10^{-4} A; i.e., this value is three–four orders of magnitude larger. We observed similar properties in the Schottky-barrier diodes fabricated on the basis of semi-insulating CdTe single crystals. As an example, Fig. 8a shows the *I–V* characteristic of a CdTe-based diode fabricated on a substrate with a resistivity of 2 × 10^9 Ω cm. The portion of linear increase in the current (the dashed straight line in Fig. 8a) is observed only for forward voltages lower than ~2 V. At higher voltages, the experimental points deviate upwards.

We can explain an increase in the current at increased forward biases by the decrease in the series substrate resistance R_s. The voltage dependence of the differential resistance of a diode $R_{dif} = dV/dI$ for a forward bias provides the quantitative characteristic for such a decrease in the resistance. Figure 8b shows these types of curves for three diodes fabricated on substrates with various resistivities: 2 × 10^4, 2 × 10^7, and 2 × 10^9 Ω cm. It is clear that, for each diode, as the voltage increases, the differential resistance initially decreases; then, the portion of constant resistance is observed to follow a pattern similar to Fig. 2. However, as the voltage further increases, the differential resistance decreases again. The higher the material resistivity, the larger the decrease in the differential resistance. It decreases for the lower curve by a factor of approximately two, for the middle curve by an order of magnitude, and for the upper curve more than by two orders of magnitude.

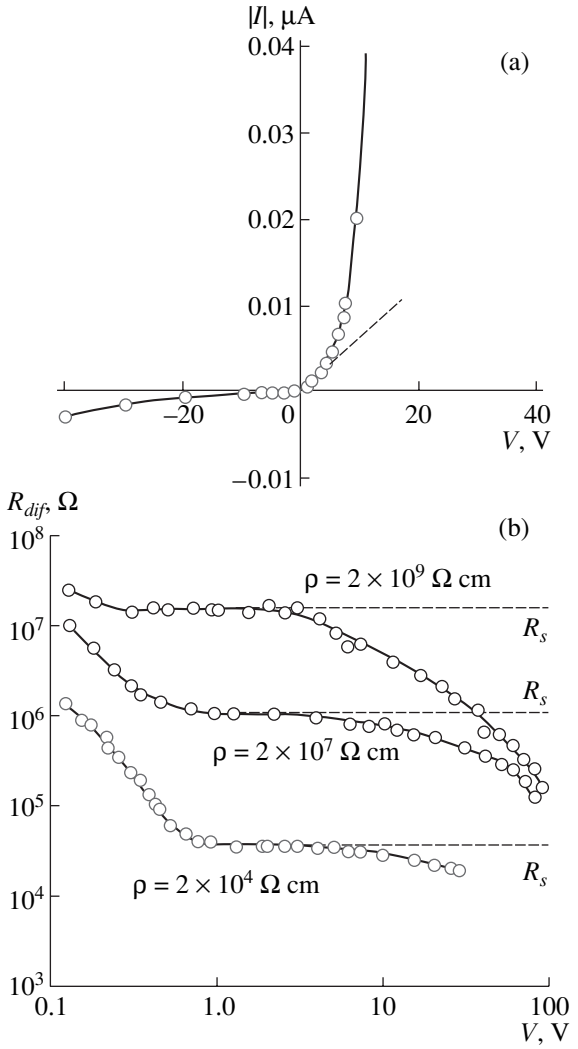

Fig. 8. (a) Current–voltage characteristic of an Al/*p*-CdTe diode on a substrate with a resistivity of 2 × 10^9 Ω cm, and (b) the differential resistance of diodes on substrates with various resistivities.

Let us consider the modulation of the bulk region of a diode (base) possessing *p*-type conductivity due to electron injection. The electron concentration at the interface between the SCR and the neutral part of the crystal (we assume that $x = 0$ for this point) is given by

$$\Delta n(0) = N_c \exp\left(-\frac{E_g - \Delta\mu}{kT}\right)\exp\left(\frac{eV}{kT}\right), \qquad (7)$$

where the voltage V of the forward bias approaches the contact potential difference V_{bi}; i.e., for the purpose of estimation, we can assume that $V = \varphi_0/e$ (Fig. 3).

The excess electron concentration Δn decreases as we move the variable point x away from the point $x = 0$ and deep into the crystal, both due to the diffusion and

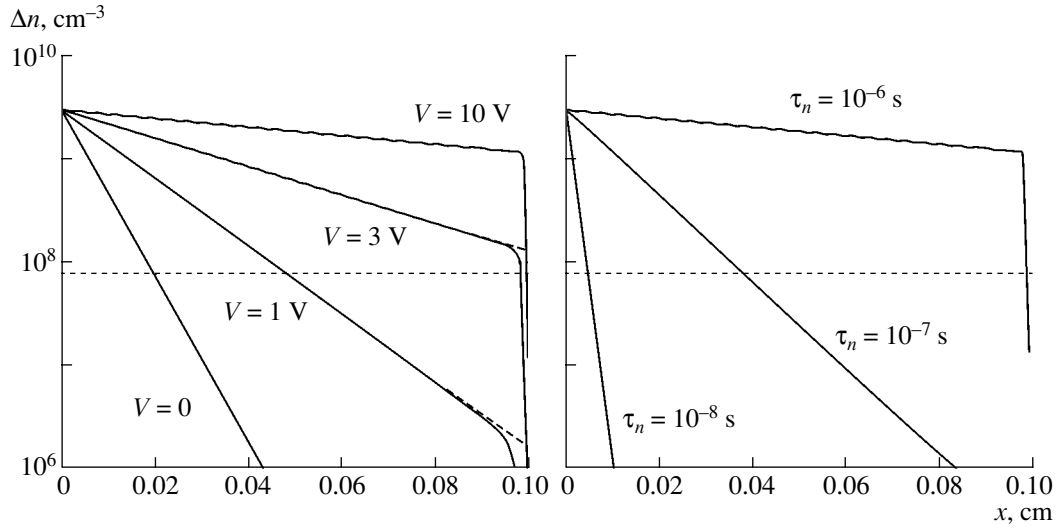

Fig. 9. Distribution of electrons injected into the *p*-type substrate: at various applied voltages and $\tau_n = 10^{-7}$ s on the left-hand side, and at various lifetimes and $V = -10$ V on the right-hand side. The substrate resistivity was 10^9 Ω cm.

the drift from the voltage drop across the substrate. For the situation of interest, i.e., high voltages, this drop equals $V - V_{bi}$. Taking into account the diffusion and drift, the distribution of Δn by the coordinate is described by the equation

$$D_n \frac{d^2\Delta n}{dx^2} + \mu_n F \frac{d\Delta n}{dx} - \frac{\Delta n}{\tau_n} = 0, \tag{8}$$

where $F = (V - V_{bi})/(d - W) \approx V/d$, since V_{bi} is an insignificant part of the voltage applied and W is much smaller than the crystal thickness d. Let us seek the solution of this equation in the form $\Delta n = C\exp(kx)$, whose substitution into formula (7) leads to the characteristic equation

$$k^2 + \frac{L_{dr}}{L_n^2}k - \frac{1}{L_n^2} = 0, \tag{9}$$

whose roots are

$$\frac{1}{L_{1,2}} = -\frac{L_{dr}}{2L_n^2} \pm \sqrt{\left(\frac{L_{dr}}{2L_n^2}\right)^2 + \frac{1}{L_n^2}}, \tag{10}$$

where $L_n = (D_n\tau_n)^{1/2}$ denotes the diffusion length of the electrons and $L_{dr} = \mu_n F\tau_n$ is the drift length (at $F \longrightarrow 0$, the roots $k_{1,2}$ are equal to $\pm 1/L_n$). Thus,

$$\Delta n(x) = A\exp\left(\frac{x}{L_1}\right) + B\exp\left(\frac{x}{L_2}\right). \tag{11}$$

In order to determine coefficients A and B, we use the following boundary conditions:

(i) at $x = 0$, $\Delta n = \Delta n(0)$, from which we have

$$A + B = \Delta n(0); \tag{12}$$

(ii) at $x = d$, the surface recombination is balanced by the electron inflow, therefore,

$$S\Delta n(d) = -D_n \left.\frac{d\Delta n}{dx}\right|_{x=d} - \mu_n F\Delta n(d). \tag{13}$$

Substituting Δn and $\Delta n/dx$ into Eq. (8), we derive an equation supplementary to Eq. (12) for A and B, whose solution relative to A is

$$A = \left[1 - \left(S + \frac{D_n}{L_1} + \mu_n F\right)\left(S + \frac{D_n}{L_2} + \mu_n F\right)^{-1} \times \exp\left(\frac{d}{L_1} - \frac{d}{L_2}\right)\right]^{-1}. \tag{14}$$

Finding A, we obtain

$$B = \Delta n(0) - A. \tag{15}$$

Figure 9 shows the results of a calculation of $\Delta n(x)$ using formula (11) and taking into account relations (14) and (15). The distribution of excess electrons in a *p*-type substrate with a resistivity of 10^9 Ω cm and $\tau_n = 10^{-6}$ s is shown at various voltages on the left-hand side. We took the typical values of ρ and τ_n for the material, which is used in the CdTe detectors. The variation in the electron distribution with a decrease in τ_n is shown on the right-hand side.

It is evident that the electric field, which emerges due to the voltage drop across the substrate, profoundly affects the distribution of injected electrons. At $V = 0$, all the electrons are accumulated in a thin layer (thickness L_n) that borders the SCR. In contrast, as the voltage increases, the region with the increased electron concentration is extended so that, at $V = 10$ V, the electron concentration already noticeably exceeds the hole con-

centration over the entire substrate thickness (the dashed horizontal line). Close to the nonrectifying contact, the surface recombination manifests itself, although insignificantly ($S = 10^6$ cm/s), as an abrupt drop at $x \longrightarrow d$.

The effect of the conductivity modulation of the substrate due to the electron injection is rapidly weakened as the electron lifetime decreases; thus, at $\tau_n < 10^{-7}$ s, the injection, even at a considerable voltage drop, becomes noticeable only in the thin layer adjoining the barrier region of a diode (Fig. 9, on the right). It is evident that the effect of the injected electrons is weakened as the substrate resistivity decreases, since the horizontal line in Fig. 9, which corresponds to the hole concentration, shifts upwards along the carrier concentration scale and, finally, exceeds the electron concentration over the entire substrate thickness.

Thus, the observed superlinear increase in the current when considerable forward voltages are applied to the diode structures, which are fabricated based on semi-insulating single crystals, is explained by the modulation of the substrate resistance due to minority carriers, specifically, in the case under study, due to electrons. Knowing the distribution of excess electrons and taking into account the injection, we calculated the substrate resistance, the *I–V* characteristic, and the differential resistance of a diode. This procedure seems to be unjustified since the Schottky diode has a reverse bias in detectors of X- and γ-ray radiation.

6. CONCLUSIONS

Al/*p*-CdTe diodes based on single crystals with a resistivity of 10^2–10^9 Ω cm (300 K) have been fabricated. The current–voltage characteristic of diodes fabricated on the basis of single crystals with both low and high resistivity is described by generation–recombination processes in the space charge region. This conclusion is in accordance with the Sah–Noyce–Shockley model after its adaptation to a Schottky diode. The observed considerable forward currents and rectifying properties of Schottky diodes based on high-resistivity CdTe are caused by the modulation of the substrate conductivity due to the injected minority carriers (electrons).

REFERENCES

1. E. N. Arkad'eva, O. A. Matveev, S. M. Ryvkin, and Yu. V. Rud', Zh. Tekh. Fiz. **36**, 1146 (1966) [Sov. Phys. Tech. Phys. **11**, 846 (1966)].
2. E. N. Arkad'eva, O. A. Matveev, S. M. Ryvkin, and Yu. V. Rud', Fiz. Tekh. Poluprovodn. (Leningrad) **1**, 805 (1967) [Sov. Phys. Semicond. **1**, 669 (1967)].
3. P. Siffert, B. Rabin, H. Y. Tabatabai, and R. Stuck, Nucl. Instrum. Methods **150**, 31 (1978).
4. A. J. Dabrovski, J. Iwanczyk, W. M. Shymczyk, *et al.*, Nucl. Instrum. Methods **150**, 25 (1978).
5. T. Takahashi, K. Hirose, C. Matsumoto, *et al.*, Proc. SPIE **3446**, 29 (1998).
6. C. Matsumoto, T. Takahashi, K. Takizawa, *et al.*, IEEE Trans. Nucl. Sci. **45**, 428 (1998).
7. T. Takahashi, B. Paul, K. Hirose, *et al.*, Nucl. Instrum. Methods Phys. Res. A **436**, 111 (1999).
8. T. Takahasi, S. Watanabe, G. Sato, *et al.*, IEEE Trans. Nucl. Sci. **48**, 287 (2001).
9. T. Takahasi and S. Watanabe, IEEE Trans. Nucl. Sci. **48**, 950 (2001).
10. Home page of Amptek Inc., www.amptek.com.
11. C. Sah, R. Noyce, and W. Shockley, Proc. IRE **45**, 1228 (1957).
12. L. A. Kosyachenko, I. M. Rarenko, Z. I. Zakharuk, *et al.*, Semiconductors **37**, 238 (2003).
13. L. A. Kosyachenko, O. L. Maskyanchuk, V. V. Motushchuk, and V. M. Sklyarchuk, Sol. Energy Meter. Sol. Cells **82** (1/2), 65 (2004).
14. D. M. Hofmann, W. Stadler, P. Chrismann, and B. K. Meyer, Nucl. Instrum. Methods Phys. Res. A **380**, 117 (1996).
15. M. Zha, E. Gombia, F. Bissoli, *et al.*, Phys. Status Solidi B **229**, 15 (2002).
16. M. Fiederle, C. Eiche, and M. Salk, J. Cryst. Growth **146**, 142 (1995).
17. L. A. Kosyachenko, O. L. Maslyanchuk, I. M. Rarenko, and V. M. Sklyarchuk, Phys. Status Solidi C **1**, 925 (2004).
18. S. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981; Mir, Moscow, 1984).

Translated by N. Korovin

Semiconductors, Vol. 39, No. 6, 2005, pp. 730–733. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 39, No. 6, 2005, pp. 762–766.

PHYSICS OF SEMICONDUCTOR DEVICES

A 4*H*-SiC *p–i–n* Diode Fabricated by a Combination of Sublimation Epitaxy and CVD

E. V. Bogdanova*^, A. A. Volkova*, A. E. Cherenkov*, A. A. Lebedev*, R. D. Kakanakov, L. P. Kolaklieva**, G. A. Sarov**, T. M. Cholakova**, A. V. Kirillov***, and L. P. Romanov*****

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

^e-mail: LeBog@mail.ioffe.ru

***Institute of Applied Physics, Plovdiv, 4000 Bulgaria*

****Svetlana-Elektropribor, St. Petersburg, 194156 Russia*

Submitted October 25, 2004; accepted for publication November 10, 2004

Abstract—The possibility of fabricating heavily doped ($N_a - N_d \geq 1 \times 10^{19}$ cm^{-3}) p^+-4*H*-SiC layers on CVD-grown lightly doped *n*-4*H*-SiC layers by sublimation epitaxy has been demonstrated. It is shown that a Au/Pd/Ti/Pd contact, which combines a low specific contact resistance (~2×10^{-5} Ω cm^2) with high thermal stability (up to 700°C), is the optimal contact to *p*-4*H*-SiC. The *p–n* structures obtained are used to fabricate packaged diodes with a breakdown voltage of up to 1400 V. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

It is known that, owing to its electrical parameters, silicon carbide (SiC) is a promising material for microwave devices [1]. However, for a long time, it has been impossible to demonstrate the advantage of SiC in the fabrication of microwave *p–i–n* diodes for a number of technological and design reasons. The reason for this difficulty was primarily the lack of thick (6–10 μm) lightly doped SiC epitaxial films with a high structural perfection. Recent advances in the development of the CVD technology mean that it is now possible to obtain *n*-4*H*-SiC layers with the required doping level and thickness. However, the problems related to the fabrication of low-resistance and thermally stable contacts to *p*-type 4*H*-SiC and the protection of the periphery of the *p–n* junction from surface breakdowns remain to be solved.

These difficulties in the formation of low-resistance contacts to *p*-type 4*H*-SiC are due to the wide band gap of SiC (3.2 eV for 4*H*-SiC). It is known that the specific contact resistance depends on the doping level of the layer in which the contact is fabricated. Theoretical calculations and experimental studies have shown that a dopant concentration exceeding 10^{19} cm^{-3} is necessary for obtaining a resistance on the order of 10^{-5} Ω cm^2 [2]. It is known that the maximum concentration of electrically active acceptors is obtained in *p*-type SiC layers grown by sublimation epitaxy [3]. Thus, the goal of the present study was to grow a p^+-type 4*H*-SiC emitter on a lightly doped *n*-type 4*H*-SiC layer by sublimation epitaxy in order to obtain a *p–n* structure suitable for the fabrication of a microwave *p–i–n* diode.

An additional goal was to optimize the post-growth technology involved in the formation of a 4*H*-SiC *p–i–n* diode to obtain low-resistance and thermally stable contacts and to lower the probability of a surface breakdown.

2. EPITAXIAL GROWTH OF p^+-TYPE 4*H*-SIC EMITTER LAYERS

The layers were grown by sublimation epitaxy in vacuum (SEV) [4] at a temperature $T = 1900$°C from an aluminum-doped source. The residual pressure in the chamber during the growth process was 3×10^{-6}–5×10^{-6} Torr, and the growth rate was ~1 μm/min.

Commercial n–n^+ structures (CREE Co.) served as the initial samples. A lightly doped ($N_d - N_a = 3.5 \times 10^{15}$ cm^{-3}) 6-μm-thick *n*-layer was grown on the "silicon" face of a 4*H*-SiC substrate with *n*-type conduction ($N_d - N_a = 5 \times 10^{18}$ cm^{-3}).

In order to fabricate high-quality *p–i–n* diodes, a p^+-type layer should be doped to $N_a - N_d \geq 1 \times 10^{19}$ cm^{-3} and have a thickness of ~1–1.2 μm (to lower the series resistance of the diode). In addition, it is necessary to obtain a high-quality metallurgical interface between the *n*- and *p*-type regions in order to eliminate leakage currents.

Earlier, we solved a similar problem in relation to 6*H*-SiC [5]. However, in the case of 4*H*-SiC, the situation was complicated by the large misorientation angle of the surface of the initial CVD-grown epitaxial layer (8°) (this angle is only 3° for 6*H*-SiC). This large misorientation angle led to an increase in the growth rate of SEV-produced epitaxial layers under the same technological conditions, which can result in a decrease in the level doping of the growing layer with aluminum atoms.

The optimization of the SEV process with respect to the temperature gradient within the growth cell and to the absolute value of the growth temperature yielded

the technological parameters at which we were able solve the problems mentioned above. The quality of the metallurgical boundary of the *p–n* junction was improved by in situ pregrowth polishing sublimation etching of the initial layer, which was effected by changing the sign of the temperature gradient within the growth cell [4].

The thicknesses of the layers grown (1.1–1.2 μm) were determined from transversal SEM images of the samples studied [6]. The atomic concentration of aluminum in the layers was measured by SIMS and found to be $\sim 1 \times 10^{20}$ cm^{-3} (Fig. 1).

The concentration and distribution of electrically active impurities in the layers were determined from the capacitance–voltage (*C–V*) characteristics. Measurements were performed with a mercury probe at a sinusoidal signal frequency of 1 kHz. It was found that a thin lightly doped film is formed on the surface of the layers (Fig. 2, curves *1*, *3*). Raising the concentration of Al in the growth source did not eliminate this film and only adversely affected the morphology of the layer.

Additional investigation demonstrated that this film was formed in the final stage of the growth process, after the crucible heater was switched off. Because of the thermal inertia of the apparatus, the temperature within the growth cell did not fall immediately, and this led to continued growth at a varying Si/C ratio in the vapor phase. As has been shown previously [7], a change in the Si/C ratio leads, if all the other conditions remain the same, to a significant change in the concentration of the main doping impurity in the growing layer.

The presence of such a film can result in an increased specific resistance in the ohmic contacts (ρ_c) to the given layer. To avoid such a situation, after growth, the samples were subjected to thermal oxidation in dry oxygen at 1150°C for 3 h, with a subsequent chemical removal of the oxide. As can be seen in Fig. 2 (curves *2*, *4*), the additional treatment made it possible to eliminate the lightly doped film and to raise the surface concentration of the acceptor impurity by a factor of 2–3.

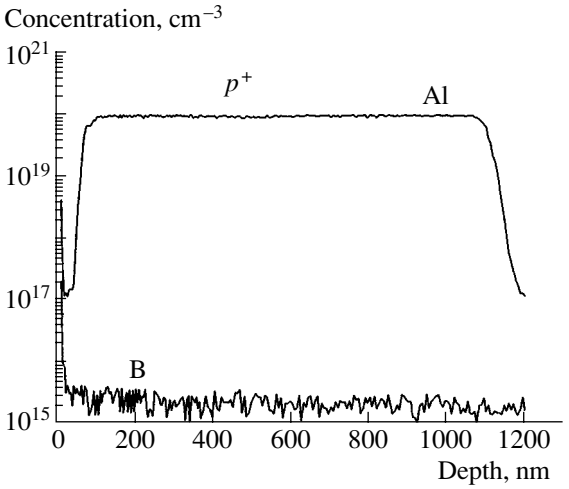

Fig. 1. SIMS profile of the p^+-type layer.

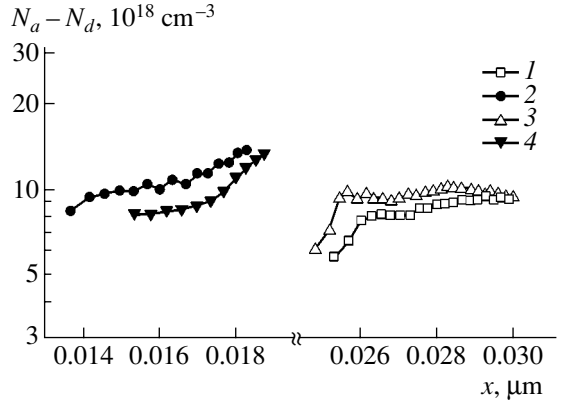

Fig. 2. ($N_a - N_d$) concentration distribution profiles for samples *S*-480 and *S*-481 (*1*, *3*) before oxidation and (*2*, *4*) after oxidation and chemical removal of the oxide.

3. FABRICATION OF *p–i–n* DIODES

3.1. Formation of Mesa Structures

It is known that the breakdown voltage of real *p–n* junctions may be lower than that expected theoretically because of the development of a surface breakdown. The probability of a surface breakdown is diminished by using various configurations of mesa structures, which make the electric field strength on the surface lower than that in the semiconductor bulk. In this study, the *p–n* junction was protected from a surface breakdown using a steplike mesa structure configuration [8, 9] (Fig. 3). The mesa structure was formed by plasmochemical etching in a vacuum installation with a saddlelike electric field generated with a dc power source in the reaction chamber. An SF_6 gas with

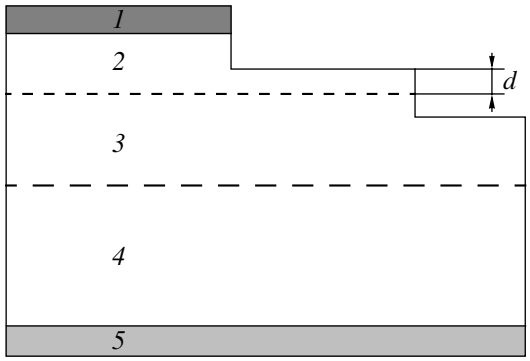

Fig. 3. Cross-section of a diode mesa structure. The ohmic contact to the (*1*) 4*H*-SiC p^+-type layer, (*2*) p^+-type 4*H*-SiC epitaxial layer, (*3*) *n*-type 4*H*-SiC epitaxial layer, (*4*) n^+-type 4*H*-SiC substrate, and (*5*) n^+-type 4*H*-SiC substrate.

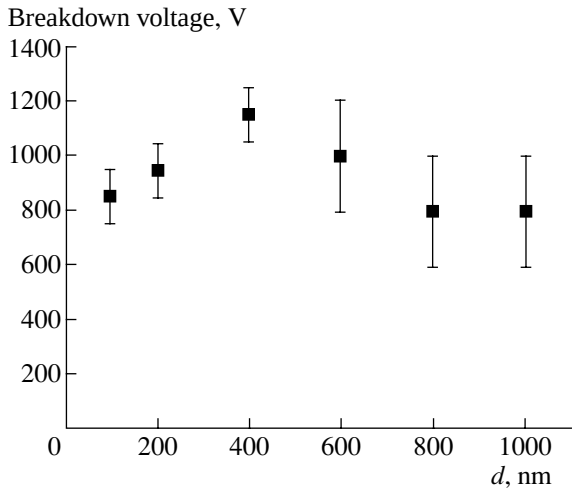

Fig. 4. Breakdown voltage of the *p–i–n* structure vs. the thickness of the expanded p^+-type layer.

an addition of O_2 (20%) was introduced into the source chamber. The working pressure in the vacuum chamber was 2×10^{-3} Torr. The etching rate of the p^+-type layer was 40 Å/min, whereas that of the *n*-type layers was higher (50 Å/min). The p^+-type layer of the samples was etched in stages and the breakdown voltage of the structures was measured after each stage.

Figure 4 shows the experimental dependence of the breakdown voltage U_{br} on the thickness *d* of the two-step p^+-type region of the *p–i–n* structure. It can be seen that, in the course of etching, the breakdown voltage of the diode structures increases from 700–900 to 1100–1200 V at an etched-off p^+-type layer thickness equal to 400 nm and to 1400 V for the separate samples. It is important to note that these results were obtained in measurements performed without any dielectric coating and in air.

Specific resistance of a contact in relation to its annealing and aging temperatures

Annealing temperature, °C	Specific resistance, Ω cm^2		
	Au/Pd	Au/Pd/Al	Au/Pd/Ti/Pd
600	7×10^{-4}	–	–
650	5×10^{-4}	–	–
700	3×10^{-4}	9×10^{-4}	3×10^{-3}
750	2×10^{-4}	4×10^{-4}	9×10^{-4}
800	1×10^{-4}	2×10^{-4}	2×10^{-4}
850	4×10^{-5}	7×10^{-5}	9×10^{-5}
900	5×10^{-5}	4×10^{-5}	2×10^{-5}
950	–	5×10^{-5}	6×10^{-5}
Aging temperature, °C			
20	4×10^{-5}	4×10^{-5}	3×10^{-5}
500	5×10^{-5}	3.8×10^{-5}	2.7×10^{-5}
600	1×10^{-4}	4×10^{-5}	3.2×10^{-5}
700	2×10^{-4}	4.2×10^{-5}	3.2×10^{-5}

3.2. Formation of Ohmic Contacts to 4H-SiC

The ohmic contact to the n^+-type substrate was composed of two metals: Ni(100 nm)/Au(200 nm). In order to form the ohmic contact, after the deposition of the metals, the samples were annealed in a reactor with resistive heating at 900°C for 5 min. The specific contact resistance was 1.8×10^{-6} Ω cm^2.

It has previously been shown that good results can be obtained in the fabrication of low-resistance contacts to p^+-type 4*H*-SiC if Pd is contained in the contact metallization [10, 11].

Three types of Pd-based contact systems were studied in order to fabricate low-resistance contacts to the p^+-type layer: Au(100 nm)/Pd(80 nm), Au(100 nm)/Pd(80 nm)/Ti(20 nm)/Pd(10 nm), and Au(100 nm)/Pd(67 nm)/Al(33 nm).

Pd was deposited by electron-beam evaporation in vacuum with a residual pressure of 1×10^{-6} Torr. Ti and Al were deposited by magnetron sputtering in an atmosphere of Ar at a pressure of 3×10^{-3} Torr. Au was deposited by thermal evaporation in vacuum at a residual pressure of 1×10^{-6} Torr. The contact pads were formed by explosive photolithography. In order to optimize the mode in which the ohmic contacts were formed, the dependence of the specific contact resistance on the annealing temperature was studied in the range 600–950°C. The annealing time at each temperature was 5 min. The specific contact resistance was measured by the TLM (transmission lime model) method on structures with five contact pads. The table presents the specific contact resistance in relation to the annealing temperature.

For the successful operation of SiC devices, it is also necessary to solve the problem related to the stability of ohmic contacts at high temperatures and high switched power. The thermal stability of the contacts obtained was studied in a temperature range from 500 to 700°C (see table). As can be seen from the table, the optimal contact to 4*H*-SiC is the Au/Pd/Ti/Pd contact, which combines a high specific contact resistance and good thermal stability.

3.3. Packaging of 4H-SiC p–i–n Structures

After the diode structures were formed, the wafer was cut into separate crystals (chips) of 0.6×0.6 mm in size. The preliminarily tested diode structures were soldered to an *M*15 case with an Au(88%)–Ge(12%) solder at 360°C in the atmosphere of pure nitrogen. The contact between the diode case and the p^+-type region was formed with an Au wire of 25 μm in diameter by thermal compression. Control measurements of the

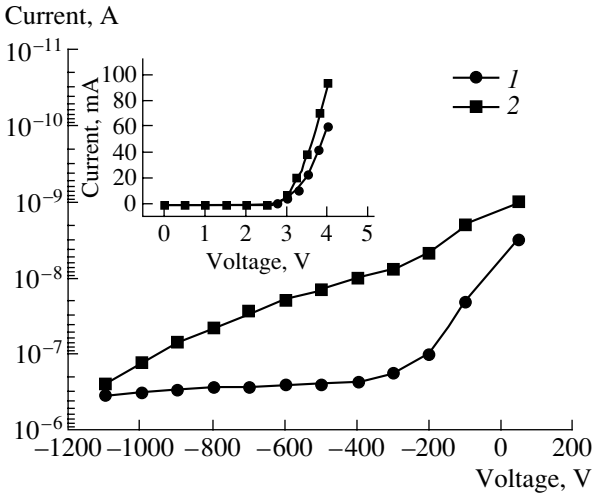

Fig. 5. *I–V* characteristics of 4*H*-SiC *p–i–n* diodes with different mesa diameters: (*1*) 130 and (*2*) 80 μm.

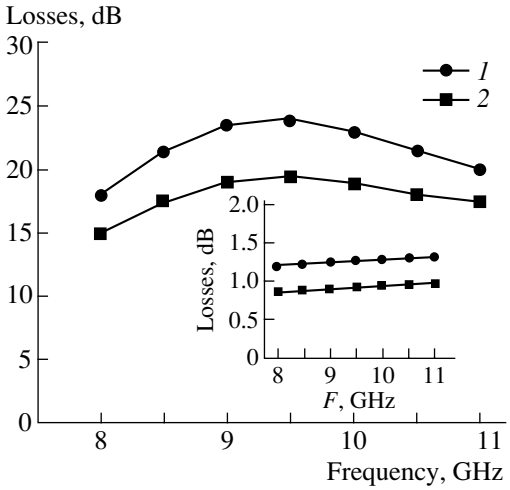

Fig. 6. Frequency characteristics of the attenuation and insertion loss of SiC *p–i–n* diode switches with *d* equal to (*1*) 130 and (*2*) 80 μm.

current–voltage (*I–V*) characteristics of the mounted diodes did not reveal any changes in the parameters of the devices.

The room-temperature reverse *I–V* characteristics of the packaged *p–i–n* diodes obtained are shown in Fig. 5. The reverse currents of most of the diodes studied fall within the region confined between the two curves shown in the figure. The difference in the slope of the curves has no effect on the breakdown voltage of the diodes, and the reverse currents do not exceed 1 μA at 1100 V. In the range 1100–1200 V, a surface electric breakdown is observed at voltages somewhat lower than the calculated bulk breakdown voltage.

The inset in Fig. 5 shows the typical room-temperature forward *I–V* characteristics of the *p–i–n* diodes under study for the mesa diameters of 130 and 80 μm.

The packaged SiC *p–i–n* diodes were studied in a microwave waveguide switch that operated in the *X* band. An isolation of 24 dB (Fig. 6) at an insertion loss of 1.2 dB (Fig. 6, inset) was obtained for a diode with a mesa diameter of 130 μm. The respective values for diodes with a mesa diameter of 80 μm were 20 and 0.9 dB. High-power tests of the serviceability of the switch samples demonstrated their stability at an input power of up to 1.0–1.5 kW in a pulsed mode.

4. CONCLUSION

In this study, it is demonstrated that 4*H*-SiC *p–i–n* structures can be fabricated by means of the SEV growth technique for p^+-type emitters on a preliminarily CVD-grown lightly doped *n*-type layer. As a result of the optimization of the post-growth treatment, 4*H*-SiC *p–i–n* diodes with a breakdown voltage of up to 1400 V were fabricated on the basis of these structures and experimental samples of switches operating in the *X* band at a pulsed power of 1–1.5 kW were created using these diodes. The parameters of the devices can be further improved by developing effective protection of the periphery of the mesa structure using dielectric films.

ACKNOWLEDGMENTS

The study was supported in part by INTAS (grant no. 01-603).

REFERENCES

1. A. S. Tager, Izv. Vyssh. Uchebn. Zaved., Fiz. **22** (10), 6 (1979).
2. J. Crofton, P. A. Barnes, and J. R. Williams, Appl. Phys. Lett. **62**, 384 (1993).
3. Yu. A. Vodakov and E. N. Mokhov, Fiz. Tverd. Tela (St. Petersburg) **41**, 822 (1999) [Phys. Solid State **41**, 742 (1999)].
4. N. S. Savkina, A. A. Lebedev, D. V. Davidov, *et al.*, Mater. Sci. Eng. B **61–62**, 165 (1999).
5. A. A. Lebedev, A. A. Mal'tsev, N. K. Poletaev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 1805 (1996) [Semiconductors **30**, 944 (1996)].
6. V. A. Solov'ev, M. P. Mikhaĭlova, K. D. Moiseev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 1300 (1998) [Semiconductors **32**, 1157 (1998)].
7. D. J. Larkin, P. G. Neudeck, J. A. Powell, and L. G. Matus, Appl. Phys. Lett. **65**, 1659 (1994).
8. X. Li, K. Tone, L. Hui, *et al.*, Mater. Sci. Forum **338–342**, 1375 (2000).
9. G. Sarov, T. Cholakova, and R. Kakanakov, Mater. Sci. Forum **457–460**, 1005 (2004).
10. L. Kassamakova, R. Kakanakov, I. Kassamakov, *et al.*, IEEE Trans. Electron Devices **46**, 605 (1999).
11. E. V. Kalinina, G. F. Kholuyanov, A. V. Shchukarev, *et al.*, Diamond Relat. Mater. **8**, 1114 (1999).

Translated by M. Tagirdzhanov