Semiconductors, Vol. 34, No. 8, 2000, pp. 861–866. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 897–902.

ATOMIC STRUCTURE AND NONELECTRONIC PROPERTIES OF SEMICONDUCTORS

Radiation Defects in *n*-6*H*-SiC Irradiated with 8 MeV Protons

A. A. Lebedev*, A. I. Veĭnger*, D. V. Davydov*, V. V. Kozlovskiĭ, N. S. Savkina*, and A. M. Strel'chuk***

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

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

Submitted January 31, 2000; accepted for publication February 1, 2000

Abstract—Capacitance methods and electron spin resonance (ESR) were applied to study deep centers in *n*-6*H*-SiC irradiated with 8 MeV protons. Schottky diodes and *p*–*n* structures grown by sublimation epitaxy or commercially produced by CREE Inc. (United States) were used. The type of the irradiation-induced centers is independent of the material fabrication technology and the kind of charged particles used. Irradiation results in an increase in the total concentration of donor centers. The possible structure of the centers is suggested on the basis of data on defect annealing and ESR. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Proton irradiation is used to passivate the periphery of semiconductor *p*–*n* structures [1]. SiC layers semi-insulating at room temperature have been obtained using proton irradiation [2–4]; however, the parameters and concentrations of deep centers formed in the course of irradiation have not been determined. Deep-center parameters have been studied elsewhere [5, 6]; however, compensation appearing under irradiation has not been considered.

The aim of this work was to obtain semi-insulating 6*H*-SiC layers by means of proton irradiation and to determine the parameters of deep centers responsible for the compensation.

2. SAMPLES

As samples, we used SiC *p*–*n* structures and epilayers either commercially produced by CREE [7] or grown by sublimation epitaxy [8]. The thicknesses of the *n*- and *p*-type layers and the substrate were 5, ~1, and ~400 μm, respectively. The Schottky barriers fabricated on epilayers and diode mesa structures were 300–1200 μm in diameter. The concentrations of uncompensated donors in the *n*-type layer, $N_{d-a} = N_d - N_a$, and the substrate were $(0.8–4) \times 10^{16}$ and $(3–5) \times 10^{18}$ cm^{-3}, respectively, and the concentration of acceptors in the *p*-type layer was $\sim 5 \times 10^{18}$ cm^{-3}. Thus, the doping level of the substrate and the *p*-type emitter exceeded by no less than two orders of magnitude that of the *n*-type base layer. Since the generation rate of radiation defects in a semiconductor is virtually independent of its doping level [9], we assume that the forward resistance of *p*–*n* structures (R_b) is determined by the carrier concentration in the *n*-type layer (by donor compensation as a result of radiation defects). To study deep centers by the method of electron spin resonance (ESR), we used ~0.5-mm-thick Lely substrates with an initial nitrogen impurity concentration of 2×10^{17} cm^{-3}.

The samples were irradiated using an MGTs-20 cyclotron with 8 MeV protons in the irradiation dose range $D = 1 \times 10^{14}–2 \times 10^{16}$ cm^{-2}.

3. EXPERIMENTAL

3.1. Capacitance–Voltage Characteristics

Capacitance–voltage (*C*–*V*) characteristics were measured using a conventional setup with a parallel equivalent circuit at a sinusoidal signal frequency of 10 kHz. A study of irradiated samples revealed a decrease in the measured value of N_{d-a} at room temperature, compared with the as-grown samples. At the same time, N_{d-a} at 650 K was higher than that in the initial structures prior to irradiation (Fig. 1). This difference increased with the irradiation dose (Fig. 2). Irradiation also caused the value of R_b to increase, and, as a result, the measured capacitance ceased to depend on applied voltage at high irradiation doses [10]. On heating, the R_b value decreased exponentially with activation energy ε_A (Fig. 3). With an increase in the irradiation dose, ε_A increased, approaching asymptotically the value of ~1.1 eV (Fig. 4). Such behavior of the resistance under irradiation (Fermi level "pinning") is also characteristic of other semiconducting materials (see, e.g., [11] for GaAs). Thus, with an increasing irradiation dose, the Fermi level is pinned in the vicinity of the level associated with the radiation defects introduced into the material at the highest rate.

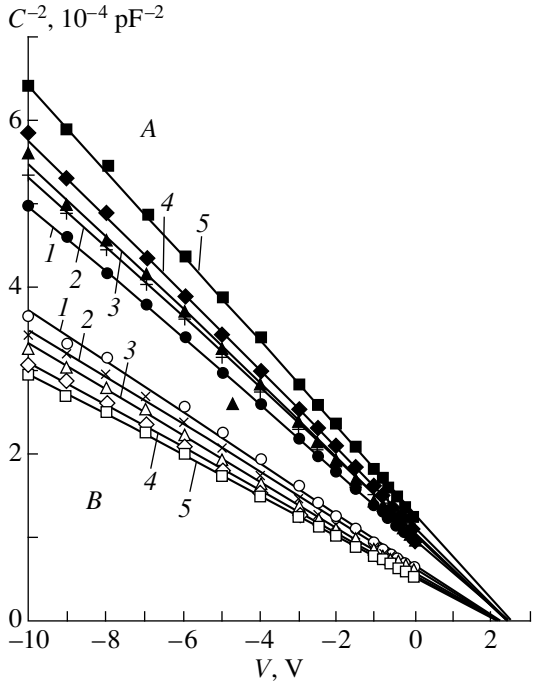


Fig. 1. *C–V* characteristics of a 6*H*-SiC *p–n* structure, measured at 300 (series *A*) and 650 K (series *B*) after irradiation with doses of (*1*) 1.1, (*2*) 2, (*3*) 3, (*4*) 4, and (*5*) 6×10^{14} cm^{-2}.

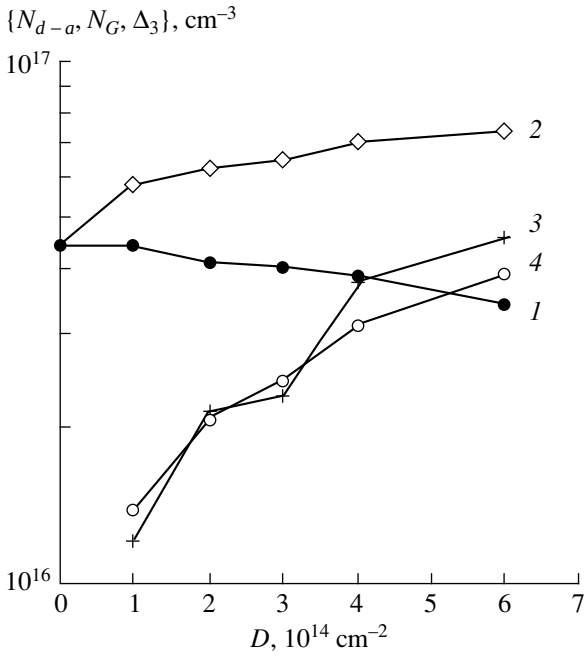


Fig. 2. N_{d-a} at (*1*) 300 and (*2*) 650 K, (*3*) the difference between these Δ_3, and (*4*) concentration of E_c – (1.1–1.22) eV centers, N_G, vs. irradiation dose.

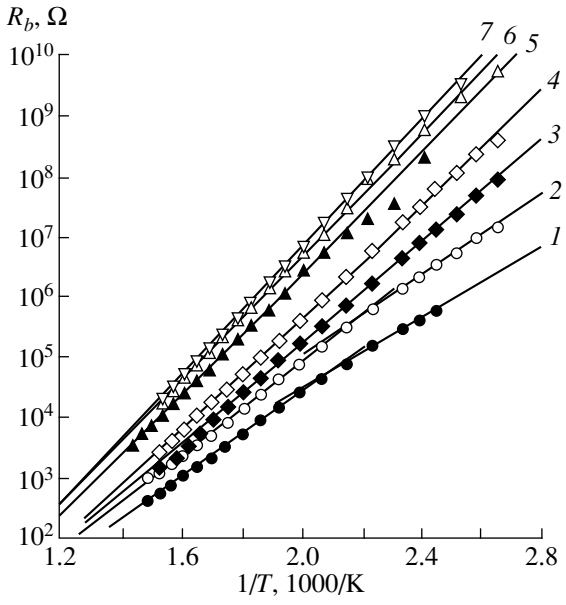


Fig. 3. Temperature dependences of the forward resistance of a Schottky diode for irradiation doses of (*1*) 3, (*2*) 4, (*3*) 6, (*4*) 10, (*5*) 20, (*6*) 50, and (*7*) 100×10^{14} cm^{-2}.

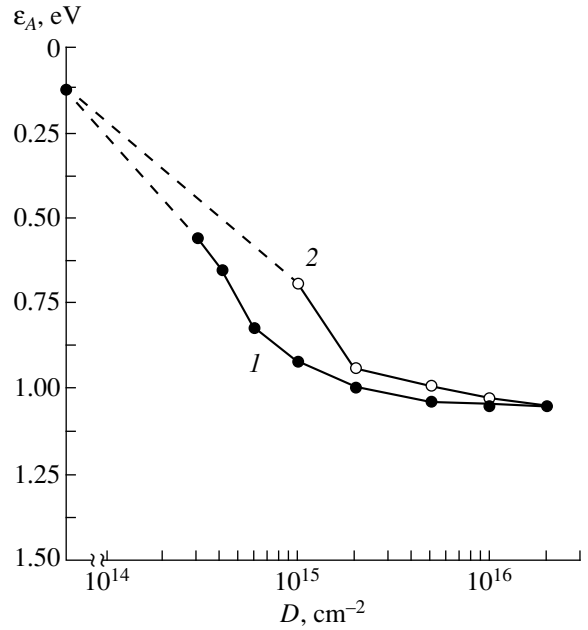


Fig. 4. Dependence of ε_A on irradiation dose for (*1*) a CREE *p–n* structure and (*2*) a Schottky diode based on an epilayer grown by sublimational epitaxy.

3.2. DLTS Measurements

A study of deep centers in the upper half of the forbidden band revealed six types of deep centers. Most of these either have been observed previously in electron-irradiated *n*-6*H*-SiC [12] or have parameters close to those of the structural defects. Table 1 lists the ionization energies of the centers ($E_c - E_0$) and their electron-capture cross sections (σ_n) and concentrations (N_g) for an irradiation dose of 2×10^{14} cm^{-2}. Two of the observed centers, at E_c – (0.16–0.2) eV and E_c – 0.5 eV,

were completely annealed out at temperatures T_a = 500–650 K (Fig. 5). In some samples, the concentration of E_c – 0.7 eV centers increased upon annealing at T_a = 800 K. No noticeable difference was observed between the spectra of deep centers formed in epilayers produced by CREE and grown by sublimation epitaxy. The introduced radiation defects were annealed out completely at $T_a \approx 1500$ K.

Deep levels lying in the lower half of the forbidden band were studied in *p*–*n* structures. Center recharging was achieved by injecting holes by switching to forward bias for a short time. In some samples, centers close in parameters to the *D* centers (E_v + 0.58 eV) [15] appeared upon irradiation with a dose of $\sim 2 \times 10^{15}$ cm^{-2} and subsequent annealing at 800 K. According to published data, the *D* center is a complex composed of a carbon vacancy and a boron atom [16]. At high irradiation doses, the signal from the centers in the lower half of the band gap disappeared. This is possibly associated with an abrupt decrease in the hole diffusion length upon irradiation [17]. Therefore, the results of DLTS measurements with injection will be incorrect under such conditions [18].

3.3. ESR Study of Deep-Level Centers Produced by Proton Irradiation

Bulk samples were irradiated under the same experimental conditions (irradiation dose 4×10^{15} cm^{-2}). The measurements were performed using a Varian E-112 ESR spectrometer at 77 K. The observed ESR signal was related to impurity nitrogen atoms whose concentration decreased only slightly as compared with that in unirradiated samples. In addition, two new spectra were recorded, one being a single line nearly coinciding with the central line of the nitrogen spectrum and the other comprising 12 lines at arbitrary sample orientation. In addition, much weaker lines that did not belong to these two spectra were observed.

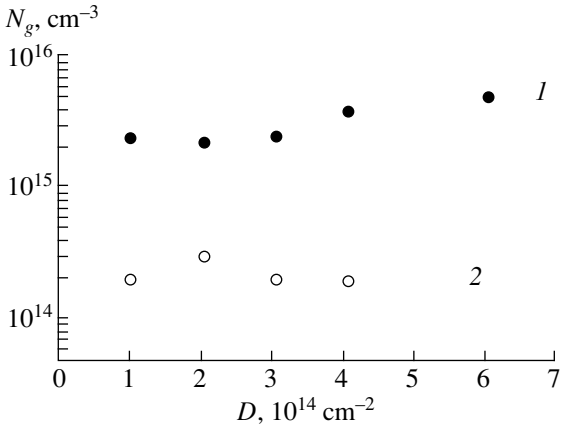


Fig. 5. Concentration of (*1*) E_c – 0.5 eV and (*2*) E_c – (0.16–0.2) eV centers in relation to irradiation dose.

4. DISCUSSION

4.1. Compensation

As shown by this investigation, proton irradiation of *n*-6*H*-SiC at the highest doses leads to the formation of deep centers (E_c – 1.22 eV) with parameters close to those of the known structural defect, the *R* center. For high irradiation doses, the sample conductivity starts to be determined by thermal excitation of electrons from this center and the activation energy of resistance approaches the ionization energy of the center. A question arises as to whether this center is an acceptor or a donor. Let us discuss this in more detail.

We first consider changes in the charge states of all the possible types of radiation defects and compare the results with experimental data on N_{d-a} variation at 300 and 650 K. As follows from the parameters of the center at E_c – (1.1–1.22) eV, the time constant (τ) of its recharging is about two weeks at 300 K. Thus, the charge state of the given center (and those lying deeper) will remain unchanged during *C*–*V* measurements at room temperature. At the same time, τ is 3.3 s for the closest level at E_c – 0.8 eV. Consequently, the E_c – 0.8 eV centers (and all those shallower) can be considered

Table 1. Parameters and properties of observed deep centers

Parameters of observed deep centers				Consistency with published data		Possible structure
$E_c - E_0$, eV	σ_n, cm^2	N_g, cm^{-3}	T_a, K	electron irradiation [12]	structural defects	
0.16–0.2	6×10^{-17}	3×10^{14}	500–650	L_1		Primary defect [*]
0.36/0.4	2×10^{-15}	3.3×10^{15}	800 < … < 1500	L_3/L_4	*E*1/*E*2 [5], *S* [13]	
0.5	5×10^{-15}	2.2×10^{15}	500–650	L_6		V_C [12, *]
0.7	4×10^{-15}	1.3×10^{15}	800 < … < 1500	L_7/L_8	Z_1/Z_2 [5]	$V_C + V_S$ [5, *]
0.8	4×10^{-15}	6×10^{14}	800 < … < 1500	L_9		
1.1–1.22	2×10^{-15}	2×10^{16}	800 < … < 1500	L_{10}	*R* [13]	$V_C + V_{Si}$ [14, *]

* This work.

Table 2. Temperature dependence of the measured N_{d-a} value for the two types of deep centers formed

Deep center	N_{d-a} (300 K)	$\Delta_1 = N_{d-a}$ (300 K) $- N_0$	N_{d-a} (650 K)	$\Delta_2 = N_{d-a}$ (650 K) $- N_0$	$\Delta_3 = N_{d-a}$ (650 K) $- N_{d-a}$ (300 K)
$E_c - 1.22$ eV, acceptor	$N_0 + N_D - N_A - N_G$	$N_D - N_A - N_G$	$N_0 + N_D - N_A$	$N_D - N_A$	N_G
$E_c - 1.22$ eV, donor	$N_0 + N_D - N_A$	$N_D - N_A$	$N_0 + N_D - N_A + N_G$	$N_D - N_A + N_G$	N_G

completely depopulated in room-temperature *C–V* measurements. When these centers are acceptors, they make no contribution to the initial (prior to irradiation) N_{d-a} concentration (N_0) at 300 K, because acceptors are neutral without bound electrons. The depopulated donor levels (N_D) are positively charged and contribute to N_0. Upon heating the structure to 650 K, the charge state of both types of centers remains unchanged.

In the case where the center at E_c – (1.1–1.22) eV with concentration N_G is a donor, it remains neutral at 300 K and does not contribute to the measured capacitance. At 650 K, this center is depopulated, which makes the initial concentration higher by N_G. If the *R* center is an acceptor, then, since the acceptors filled with electrons are charged negatively, N_0 decreases by N_R in *C–V* measurements at 300 K. As the temperature increases, the *R* center becomes depopulated (neutral) and makes no contribution to the measured capacitance (concentration).

Also, account should be taken of the possibility that radiation defects are produced in the lower half of the forbidden band. If these centers are donors, they are filled (neutral) and make no contribution to the measured concentration. Formation of acceptors, also filled with electrons (negatively charged, with concentration N_A), reduces the N_0 value by N_A. In the temperature range in question, the charge state of both types of centers remains unchanged.

For more clarity, all the aforesaid is summarized in Table 2.

It can be seen from Table 2 that, in both possible cases, the difference of the N_{d-a} values measured at 300 and 650 K (Δ_3) must be equal to the concentration of E_c – (1.1–1.22) eV centers, determined by DLTS. As can be seen from Fig. 2, this equality agrees well with the experiment for doses $\leq 3 \times 10^{14}$ cm^{-2}. With an increasing irradiation dose, the concentration of deep-level radiation defects becomes comparable with that of the shallow-level impurity; i.e., conditions are violated under which the expressions used in determining the concentration of deep-level centers from DLTS spectra are derived. Thus, the experimentally observed (for doses $>3 \times 10^{14}$ cm^{-2}) relationship $\Delta_3 < N_R$ may be due to the limited applicability of the DLTS technique.

As can be seen from Table 2, the available experimental data do not permit an unambiguous conclusion as to whether the E_c – (1.1–1.22) eV center is a donor or an acceptor. Since, at doses $<1.8 \times 10^{15}$ cm^{-2}, no observable radiation defects are formed in the lower half of the forbidden band, it may be assumed that $N_A \approx 0$. Then, in the case where this center is a donor (Table 2), $\Delta_1 > 0$, which contradicts the experiment. Thus, the only remaining possibility is that this center is an acceptor. This is in agreement with the results obtained in studying the Pool–Frenkel effect for the given center [5]. On the other hand, taking $N_A \approx 0$, we assume that no centers are formed lying closer to the valence band than E_v + (0.1–0.15) eV, i.e., the range in which DLTS measurements are impossible in our case. Therefore, the conclusion that the center at E_c – 1.22 eV is an acceptor can be considered tentative.

In the case where the center in question is a donor, a higher irradiation dose will not obviously "depress" the Fermi level any further. However, if the given deep center is an acceptor (which is possible at $N_A \approx 0$), the Fermi level will also be pinned near the level corresponding to the given defect, the deepest of all those induced by radiation. This center will be partly filled with electrons coming from the shallower donor levels, and only the degree of the sample compensation will become larger with increasing irradiation dose. Thus, irrespective of the nature of the given deep center, the maximum activation energy of the base resistance will be close to the ionization energy of the center, i.e., ~1.22 eV. This means that proton irradiation can be used to obtain *n*-6*H*-SiC layers, semi-insulating only near room temperature. Upon heating, the resistance of the irradiated structures decreases exponentially and the leakage currents increases, which has been observed experimentally [2].

4.2. Identification and Possible Structure of the Centers

It can be seen from Table 1 that there is good agreement between the spectra of deep centers produced in 6*H*-SiC by proton or electron irradiation. There also exists a correlation between the temperatures at which these centers are annealed out. Such a correlation is also preserved for centers produced in SiC by irradiation with alpha particles [4]. Thus, a conclusion can be made that, in SiC, as also in other semiconducting materials [11], various kinds of radiation treatment produce the same set of deep centers, but with different concentration ratios.

Despite the coincidence of the radiation defect parameters with those of some native defects (*R*, *E*1/*E*2, *Z*1/*Z*2), it seems premature to state the full identity of the deep centers and the native defects. It should be recalled that the radiation defects are annealed out completely at ~1500 K, while the native defects with nearly the same parameters exist in sublimation-grown *n*-6*H*-SiC epilayers up to ~2100 K. Their concentration remains unchanged on sublimational overgrowth of an additional *p*-type epilayer [13] at the same growth temperature.

In [6, 14], the mechanism of formation and annealing of radiation defects in SiC was studied on the basis of positron lifetime experiments. It was found that proton irradiation produces carbon and silicon vacancies and other primary defects, recombining upon annealing to 400–600 K to form thermally stable complexes. Thus, our DLTS investigations suggest that the centers at E_c – (0.16–0.2) eV and E_c – 0.5 eV are related to primary radiation defects and that the E_c – 0.7 eV center is, similarly to the *D* center, a vacancy complex.

Previously, for some structural defects appearing in SiC samples upon quenching, the structure of the defects has been related to their electrical properties, taking advantage of the photosensitivity of their ESR spectra [14]. A study of the effect of light on the amplitude of ESR absorption spectra taken for the samples studied in this work demonstrated their insensitivity to light. In this connection, to correlate the structure and electrical properties of deep centers, we compared the parameters of the ESR spectra taken in this work with the parameters obtained in studying thermal defects in quenched samples [14] and radiation defects in samples irradiated with neutrons and alpha particles [19]. As shown in [18, 19], defects in SiC give rise to two kinds of ESR spectra: with spins $S = 1/2$ and $S = 1$. The latter exhibit a fine structure.

In our case, we also observed two types of spectra. A weakly anisotropic line with a *g*-factor nearly coinciding with that of nitrogen corresponds to the center with $S = 1/2$. Since the nitrogen-related spectrum did not disappear in the samples studied, it masked the possible weak lines of the hyperfine structure, related to the interaction of the electron spin with the nuclear spins of the isotopes ^{29}Si and ^{13}C. In this connection, we have thus far been unable to identify the exact structure of this center. However, comparison with the spectrum associated with quenched-in defects suggests that the ESR line with spin $S = 1/2$ originates from an isolated vacancy in the carbon sublattice with a thermal activation energy of 0.5 eV (in [14], the measured optical activation energy was 0.65 eV).

A much more precise identification could be obtained for the center with spin $S = 1$. Taken in arbitrary orientation, the spectrum contains 12 lines, which means that the center responsible for this spectrum possesses a twofold symmetry, with the symmetry axis lying in the plane containing the **c**-axis and one of the twofold axes. The maximum observable fine splitting is 580 Oe (510×10^{-4} cm^{-1}) at an angle of 30° between the magnetic field direction and the **c**-axis. A spectrum of this kind is characteristic of the center *G*2, observed in [19] and interpreted as a pair of nearest vacancies in the carbon and silicon sublattices. The photosensitive thermal defect *P*7 with optical activation energy of 1.17 eV, observed in [14], possesses a similar spectrum. It should also be noted that variously oriented divacancies have close activation energies. Hence, it follows that the spectrum of the center with $S = 1$ can be identified in all probability with the deepest of the centers revealed by DLTS, with an activation energy of 1.1–1.2 eV.

Thus, ESR data made it possible to correlate the structure and electrical properties of two deep centers, namely, a carbon vacancy with an activation energy of 0.5 eV and a pair of vacancies in the carbon and silicon sublattices with an activation energy of 1.1 eV. The different energies of thermal ionization of the centers E_c – 1.22 eV and E_c – 0.7 eV at the same presumed structure ($V_C + V_{Si}$) can apparently be explained by different distances between pair components (vacancies) characteristic of each of these radiation defects.

5. CONCLUSION

(i) The properties and parameters of radiation defects formed in *n*-6*H*-SiC under irradiation with different types of charged particles are the same.

(ii) The total concentration of uncompensated donors in proton-irradiated *n*-6*H*-SiC samples increases. At the same time, the irradiation leads to the formation of deep acceptor centers to which electrons are transferred from shallower donor levels. This depresses the Fermi level and gives rise to *n*-6*H*-SiC layers that are semi-insulating at room temperature. Such layers can be used in fabrication of devices that are not intended for operation at high temperatures, e.g., photodetectors and detectors of various kinds of radiation.

REFERENCES

1. V. V. Kozlovskiĭ, S. I. Ponomarev, and I. A. Kozlovskaya, Pis'ma Zh. Tekh. Fiz. **20** (2), 1 (1994) [Tech. Phys. Lett. **20**, 46 (1994)].
2. A. O. Konstantinov, V. N. Kuz'min, L. S. Lebedev, *et al.*, Zh. Tekh. Fiz. **54**, 1622 (1984) [Sov. Phys. Tech. Phys. **29**, 949 (1984)].
3. G. C. Rybicki, J. Appl. Phys. **78**, 2996 (1995).
4. R. K. Nadela and M. A. Capano, Appl. Phys. Lett. **70**, 886 (1997).
5. T. Dalibor, G. Pensl, H. Matsunami, *et al.*, Phys. Status Solidi A **162**, 199 (1997).
6. W. Puff, P. Mascher, A. G. Balogh, and H. Baumann, Mater. Sci. Forum **258**–**263**, 733 (1997).
7. J. W. Palmor, J. A. Edmond, H. S. Kong, and C. H. Carter, Jr., Physica B (Amsterdam) **185**, 461 (1993).

8. N. S. Savkina, A. A. Lebedev, D. V. Davydov, *et al.*, Mater. Sci. Eng. B **61/62**, 165 (1999).
9. *Radiation Processing in the Technology of Electronic Materials and Devices*, Ed. by V. S. Ivanov and V. V. Kozlovskiĭ (Énergoatomizdat, Moscow, 1997).
10. A. A. Lebedev and N. A. Sobolev, Fiz. Tekh. Poluprovodn. (Leningrad) **16**, 1874 (1982) [Sov. Phys. Semicond. **16**, 1207 (1982)].
11. V. N. Brudnyĭ and A. I. Potapov, in *Proceedings of the 7th Russia Conference "Gallium Arsenide," Tomsk, 1999*, p. 68.
12. V. S. Ballandovich, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 1314 (1999) [Semiconductors **33**, 1188 (1999)].
13. M. M. Anikin, A. S. Zubrilov, A. A. Lebedev, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **25**, 479 (1991) [Sov. Phys. Semicond. **25**, 289 (1991)].
14. A. I. Veĭner, V. A. Il'in, Yu. M. Tairov, and V. F. Tsvetkov, Fiz. Tekh. Poluprovodn. (Leningrad) **15**, 1557 (1981) [Sov. Phys. Semicond. **15**, 902 (1981)].
15. M. M. Anikin, A. A. Lebedev, A. L. Syrkin, and A. V. Suvorov, Fiz. Tekh. Poluprovodn. (Leningrad) **19**, 114 (1985) [Sov. Phys. Semicond. **19**, 69 (1985)].
16. P. G. Baranov, I. V. Il'in, and E. N. Mokhov, Fiz. Tverd. Tela (St. Petersburg) **40**, 36 (1998) [Phys. Solid State **40**, 31 (1998)].
17. A. A. Lebedev, A. M. Strel'chuk, V. V. Kozlovskiĭ, *et al.*, Mater. Sci. Eng. B **61/62**, 450 (1999).
18. A. A. Lebedev, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 999 (1996) [Semiconductors **30**, 531 (1996)].
19. N. M. Pavlov, M. I. Iglitsyn, M. G. Kosaganova, and V. N. Solomatin, Fiz. Tekh. Poluprovodn. (Leningrad) **9**, 1279 (1975) [Sov. Phys. Semicond. **9**, 845 (1975)].

Translated by M. Tagirdzhanov

Semiconductors, Vol. 34, No. 8, 2000, pp. 867–871. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 903–908.

ATOMIC STRUCTURE AND NONELECTRONIC PROPERTIES OF SEMICONDUCTORS

TEM Structural Studies of Undoped and Si-doped GaN Grown on Al_2O_3 Substrate

N. A. Cherkashin*, N. A. Bert*, Yu. G. Musikhin*, S. V. Novikov*, T. S. Cheng, and C. T. Foxon****

** Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia; e-mail: Cherkashin@pop.ioffe.rssi.ru*

*** School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*

Submitted February 9, 2000; accepted for publication February 17, 2000

Abstract—Transmission electron microscopy was used to study the microstructure of GaN films undoped or Si-doped to 10^{17} or 10^{18} cm^{-3} and grown by molecular-beam epitaxy on (0001) Al_2O_3 substrate without nitridation or a buffer layer. Defect structures including inversion domains, nanopipes, and (0001) stacking faults were studied. The influence of Si doping on the threading dislocation density and the dimensions of GaN grains bounded by inversion domains was assessed. Smoothing of the steplike morphology of the GaN film surface occurs at a Si concentration of 10^{17} cm^{-3}. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The use of thin GaN films as a material for bright blue light-emitting diodes and lasers [1] has aroused considerable interest in the microstructure of GaN films [2–8]. A specific feature of structures with a GaN film grown, in particular, on the Al_2O_3 substrate is their high defect density $\geq 10^9$ cm^{-2} [2–4], exceeding by several orders of magnitude that admissible for GaAs device structures. Previous studies of GaN/Al_2O_3 structures have shown that the main defects are dislocations, stacking faults in the basal plane (0001), and planar defects [2–8]. Most of the previously revealed threading dislocations were partial edge dislocations with the Burgers vectors $\mathbf{b} = \pm 1/3[11\bar{2}0]$, $\mathbf{b} = \pm 1/3[1\bar{2}10]$, and $\mathbf{b} = \pm 1/3[2\bar{1}\bar{1}0]$. Special attention has been given to the so-called inversion domains and nanopipes. Unfortunately, no unambiguous explanation of the structure and origin of these defects has been provided. Most of the studies have been performed with GaN structures grown on a buffer layer [2, 4–8]. The substrate nitridation and the introduction of a buffer layer (e.g., AlN, AlGaN, or GaN) allowed the strains associated with the significant mismatch of the Al_2O_3 and GaN lattices in the basal plane of growth to relax chiefly in the buffer layer. Further growth of GaN film has already proceeded without formation of a large number of defects. The GaN film growth was strongly affected by the buffer layer polarity and also by the presence of impurities in the material [7, 8]. As shown in [3], GaN films grow on Al_2O_3 substrates with and without a buffer layer in significantly different ways. For this reason, determining the influence of various growth conditions, in particular, those of silicon doping, on the structure of GaN films grown without a buffer layer on the Al_2O_3 substrate is of particular interest.

2. SAMPLES AND EXPERIMENTAL PROCEDURES

GaN (0001) films, undoped or Si-doped to 10^{17} cm^{-3} and 10^{18} cm^{-3}, were MBE-grown on (0001) Al_2O_3 substrates without additional treatment (e.g., nitridation) or a buffer layer.

An EM-420 Philips microscope with 100 kV acceleration voltage was used for transmission electron microscopy (TEM). The samples to be studied were prepared by a conventional method including mechanical treatment (grinding and polishing) and sputtering of the material with 4-keV Ar^+ ions at an angle of 14° to the sample surface.

3. RESULTS

All the samples were prepared in the cross-sectional geometry $(2\bar{1}\bar{1}0)$. Samples with GaN films, undoped and Si-doped to 10^{18} cm^{-3}, were also prepared in the planar geometry (0001).

The relative orientation of GaN and Al_2O_3 was determined by analyzing the diffraction patterns for samples of cross-sectional $(2\bar{1}\bar{1}0)$ geometry with the electron beam parallel to the interface. It was found that the GaN film is rotated by 30° about the direction normal to the interface (which corresponds to the **c**-axis). This rotation is caused by the fact that the system tends to relieve the stress due to a strong mismatch between the GaN and Al_2O_3 lattices in the basal plane of growth [2].

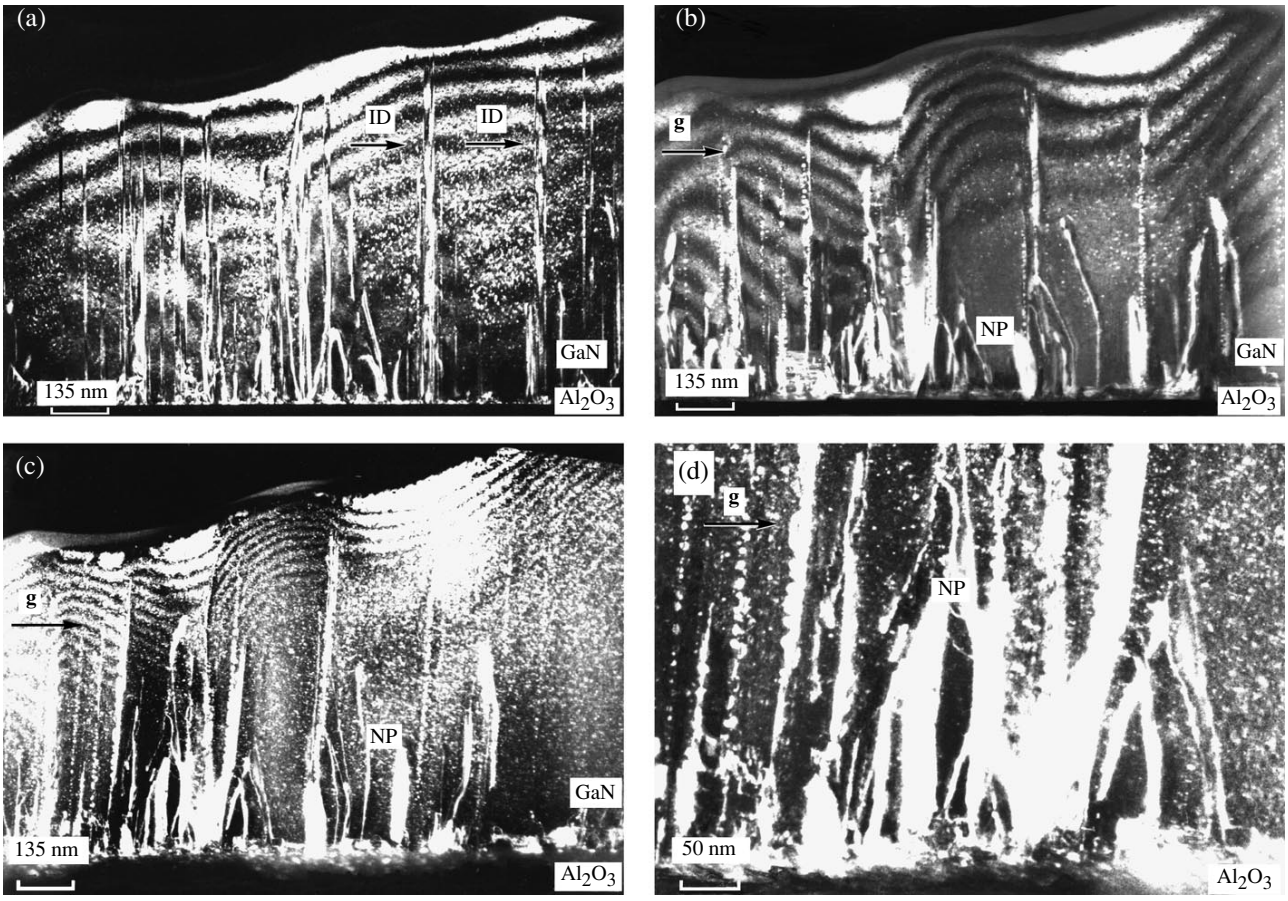


Fig. 1. Weak-beam (**g**, 3**g**) cross-sectional TEM images of GaN film doped with silicon to 10^{18} cm^{-3}: (a) for the $(2\bar{1}\bar{1}0)$ cross section with (0002) reflection excited; (b) for the $(2\bar{1}\bar{1}0)$ cross section with $(01\bar{1}0)$ reflection excited; (c) for the $(11\bar{2}0)$ cross section with $(1\bar{1}00)$ reflection excited; and (d) for the $(11\bar{2}0)$ cross section with $(1\bar{1}00)$ reflection excited. ID stands for inversion domains and NP stands for nanopipes; the same in Fig. 3.

Analysis of the contrast of TEM images shows the following basic types of defects in undoped and Si-doped GaN: three-dimensional (3D) inversion domains (IDs) and nanopipes, threading dislocations (of screw, edge, and mixed types), and stacking faults in the (0001) basal plane.

To identify the dislocation types, we used the method of extinction of the contrast from a dislocation at a direction of the excited reflection perpendicular to the Burgers vector **b**. Images were obtained in a weak-beam mode (**g**, 3**g**), when the strong contrast from the strain fields generated by dislocations is suppressed. A strong contrast from threading dislocations is observed under weak-beam conditions (**g**, 3**g**) with $(01\bar{1}0)$ reflection excited for the $(2\bar{1}\bar{1}0)$ cross section (Fig. 1b) and $(11\bar{2}0)$ reflection for the $(1\bar{1}00)$ cross section (Figs. 1c, 1d) and zero contrast is observed with (0002) reflection excited for the $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ cross sections (Fig. 1a). This suggests that these dislocations are of the edge type with $\mathbf{b} = \langle hkl0 \rangle$. It was shown in [2] that dislocations of this type are partial edge dislocations with $\mathbf{b} = 1/3\langle 11\bar{2}0 \rangle$ and $\mathbf{b} = 1/3\langle 2\bar{1}\bar{1}0 \rangle$. In the Si-doped GaN films, no dislocations were found that would show contrast with the $(11\bar{2}0)$ reflection excited for the $(1\bar{1}00)$ cross section and would show no contrast for any of the other above-described conditions. Therefore, partial edge dislocations with $\mathbf{b} = 1/3\langle 2\bar{1}\bar{1}0 \rangle$ are absent in these samples (or their density is negligible as compared with the densities of other types of dislocations).

Screw dislocations with **b** = [0002] were detected by the presence of contrast with the reflection (0002) excited for the $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ cross sections (Fig. 1a) and zero contrast with $(01\bar{1}0)$ and $(11\bar{2}0)$ reflections for the $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ cross sections, respectively (Figs. 1b, 1c).

Dislocations exhibiting contrast (or partially suppressed contrast) under any of the above-listed conditions have $\mathbf{b} = \langle hklm \rangle$. As shown in [2], this contrast behavior corresponds to mixed-type dislocations with $\mathbf{b} = 1/3[11\bar{2}3]$.

Thus, this method of dislocation-type identification was applied to analyze TEM images of the $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ cross sections for three types of GaN samples: undoped and Si-doped to 10^{17} and 10^{18} cm^{-3}. The results are listed in the table.

Inversion domains and nanopipes are present in all the samples along with dislocations and (0001) stacking faults. Cross-sectional $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ images with the (0002) reflection excited show well-defined stripes in light and dark fields; these stripes originate from the interface, terminate at the surface of the GaN layer, exhibit a contrast different from that of the matrix, and have sharp boundaries (Fig. 2). These defect structures are interpreted as inversion domains. An analysis of the contrast for these formations in relation to excited reflections shows that the stripe boundaries exhibit no contrast with the $(01\bar{1}0)$ reflection excited for the $(2\bar{1}\bar{1}0)$ cross section or with the $(11\bar{2}0)$ reflection for the $(1\bar{1}00)$ cross section (Figs. 1a, 1b). In this case, a contrast associated with small-sized inclusions stacked in the [0001] direction appears at the place corresponding to a quenched stripe; this suggests that screw dislocations with $\mathbf{b} = [0002]$ are present in these inversion domains. This conclusion was confirmed by images of the undoped and Si-doped (10^{18} cm^{-3}) GaN in the (0001) planar geometry with the $(01\bar{1}0)$ and $(11\bar{2}0)$ reflections excited (Fig. 3). Clearly visible are the GaN grains bounded by sharp-boundary stripes with a contrast different from that of the matrix. Also, screw dislocations can be seen both at the outer boundaries of stripes and at the stripe–matrix-grain interface. GaN matrix grains are hexagonal in shape and have boundaries whose orientation will be determined later. Their lateral dimensions for undoped and Si-doped (10^{17} and 10^{18} cm^{-3}) GaN are ~400–500, 300–400, and 200–300 nm, respectively. Superposition of the diffraction pattern onto images obtained in diffraction contrast with excited reflections $(11\bar{2}0)$, $(01\bar{1}0)$ allowed the orientation of grain boundaries to be identified as $\{01\bar{1}0\}$ and $\{11\bar{2}0\}$. A comparison of the images obtained from samples having cross-sectional geometry $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ with those obtained from samples of (0001) planar geometry demonstrated that the contrast from the stripes bounding the GaN grains corresponds to inversion domains with $\{01\bar{1}0\}$ and $\{11\bar{2}0\}$ boundaries. The inversion-domain width varies from 5 to 30 nm for Si-doped GaN, and from 15 to 50 nm for undoped GaN.

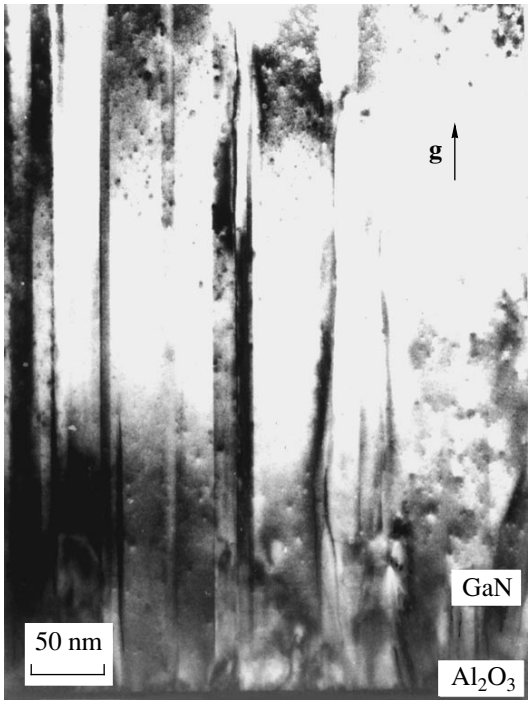


Fig. 2. Bright-field $(2\bar{1}\bar{1}0)$ cross-sectional TEM image of GaN film doped with silicon to 10^{18} cm^{-3} with (0002) reflection excited.

Densities of threading dislocations and nanopipes

GaN sample	Defect density, cm^{-2}				
	threading dislocations				nanopipes
	screw type with **b**	partial edge type with **b**		mixed type with **b**	
	[0002]	$1/3[11\bar{2}0]$	$1/3[2\bar{1}\bar{1}0]$	$1/3[11\bar{2}3]$	
Undoped	1×10^9	2×10^9	3×10^9	4×10^9	2×10^9
Doped with $n_{Si} = 10^{17}$ cm^{-3}	7×10^9	5×10^9	–	4×10^9	2×10^9
Doped with $n_{Si} = 10^{18}$ cm^{-3}	1×10^{10}	2×10^9	–	7×10^9	2×10^9

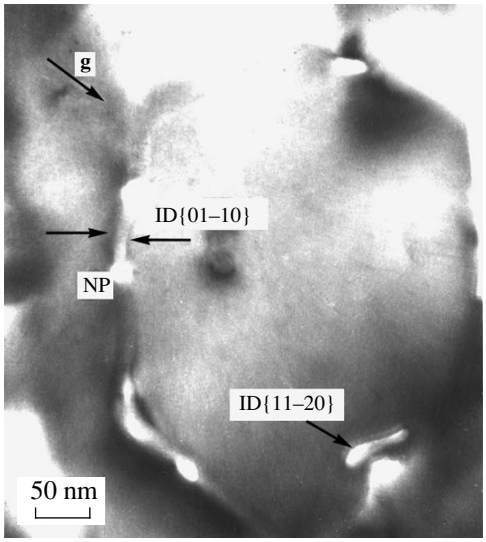


Fig. 3. Bright-field (0001) plan-view TEM image of GaN film doped with silicon to 10^{18} cm^{-3} with $(11\bar{2}0)$ reflection excited.

Figure 2 shows $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ cross-sectional images with the (0002) reflections excited. The contrast from stripes bounded by (0001) stacking faults either from above or from both above and below is superimposed onto the above-described stripes of inversion domains. The dislocations may originate at the interface and thread as far as these stacking faults, or these defect structures themselves may be a source of dislocations. This type of formation most clearly manifests itself in the weak-beam (**g**, 3**g**) $(1\bar{1}00)$ cross-sectional images with excited reflections $(11\bar{2}0)$ (Figs. 1c, 1d), in which case it gives rise to bright contrast typical of dislocations. These formations are nanopipes consisting of aggregated screw dislocations with an empty core [2–4, 7, 8]. The nanopipe width varies between 5 and 25 nm. It should be noted that, for images taken under these conditions, the presence of contrast from the dislocations originating at the interface and threading as far as nanopipes and those originating at the stacking faults bounding the nanopipes from above indicates that these dislocations are of the edge or mixed type. The images of a sample prepared in the (0001) planar geometry (Fig. 3) show bright-contrast hexagonal nanopipes located in the inversion-domain region for all types of reflections. Evaluation of the lateral dimensions of the nanopipes yields values in the range from 5 to 20 nm, correlating with their width obtained for the sample with cross-sectional $(2\bar{1}\bar{1}0)$ geometry. The nanopipe densities were evaluated for all the samples, both those with undoped and Si-doped (10^{17} and 10^{18} cm^{-3}) GaN films (see table).

An analysis of the diffraction patterns and diffraction images yielded the following results:

(1) The predominant types of defects in undoped and Si-doped GaN are 3D inversion domains, nanopipes, threading dislocations (of screw, edge, and mixed types), and stacking faults in the (0001) basal plane.

(2) The inversion domains bound hexagonal GaN grains with boundaries of $\{01\bar{1}0\}$ and $\{11\bar{2}0\}$ types.

(3) The nanopipes observed in all the samples are mainly located in the regions of inversion domains. They have a hexagonal shape and lateral dimensions from 5 to 25 nm.

(4) GaN doping with silicon leads to (a) a considerable increase in the screw dislocation density (see table); (b) suppression of the partial edge dislocations with **b** = $1/3[2\bar{1}\bar{1}0]$ and a considerable decrease in the density of edge dislocations with an increasing doping level (see table); (c) a nearly twofold increase in the density of mixed-type dislocations with increasing doping; (d) diminution of the lateral size of GaN matrix grains from ~400–500 nm for undoped GaN to 300–400 and 200–300 nm for Si concentrations of 10^{17} and 10^{18} cm^{-3}, respectively; and (e) diminution of the inversion domain width from 15–50 nm for undoped GaN to 5–30 nm for doped GaN films.

4. DISCUSSION

The initial growth of GaN films occurs in a strong strain field induced by considerable lattice mismatch between GaN and Al_2O_3 in the basal plane. This results in the formation of numerous defects at the interface containing dislocations, dislocation loops, stacking faults, and impurity clusters that can act as nuclei of threading defect boundaries in the film. According to [3, 8], without nitridation of the Al_2O_3 substrate and without the buffer layer, the GaN matrix has N-polarity and the inversion domains have Ga-polarity. As is known, the polarity inversion in GaN changes the growth rate of a material being deposited [8]. If the polarity direction changes from "Ga to N" to "N to Ga," the growth rate nearly doubles [8]. The growth rate in the nonpolar directions $[2\bar{1}\bar{1}0]$ and $[01\bar{1}0]$ is 50–100 times higher than that in the polar **C** direction, thus stimulating the formation of a homogeneous film. Nevertheless, the initial presence of nuclei and an increase in the growth rate lead to threading of the inversion domains as far as the surface. Under these growth conditions, the formation of defects only near the substrate/layer interface does not ensure total relief of stress. The inversion domains thread only where the GaN layer accumulates sufficient stress to be relieved via formation of a lower-energy boundary. Owing to the equivalence of the nonpolar $[2\bar{1}\bar{1}0]$ and $[01\bar{1}0]$ directions, the inversion domains emerge both at the boundaries of hexagonal grains in the GaN matrix and in the directions of the $\{01\bar{1}0\}$ type. In terms of this

approach, it becomes clear why the grain size in the GaN matrix decreases upon doping with silicon. The presence of additional impurities merely enhances the stress in the layer; this stress is relieved at smaller distances via formation of inversion domains. The general increase in the density of practically all types of dislocations confirms the fact that the stress is enhanced in the Si-doped GaN film and relaxes via defect formation. However, no correlation is observed between the densities of threading dislocations and nanopipes. It is noteworthy that a Si-doped GaN film with a dopant concentration of 10^{17} cm^{-3} has a smoother surface as compared with that of undoped and Si-doped (10^{18} cm^{-3}) GaN. This can be explained by the effect of "impurity poisoning" [8] of the $\{01\bar{1}0\}$ and $\{2\bar{1}\bar{1}0\}$ planes, the facets of the inversion domains. This stabilizes the growth of inversion domains and equalizes the growth rates of the N-polar GaN matrix and Ga-polar inversion domains. Additionally, the impurity deposition via formation of clusters stacked in the growth direction at the inversion domain boundaries is confirmed by the $(2\bar{1}\bar{1}0)$ and $(1\bar{1}00)$ cross-sectional images with, respectively, the $(01\bar{1}0)$ and $(11\bar{2}0)$ reflections excited (Figs. 1b, 1c).

5. CONCLUSION

In summary, we note that a large number of defects are present at the interface, including stacking faults, dislocations, and dislocation loops. GaN films are characterized by a high density of threading dislocations of the screw type with **b** = [0002], of the partial edge type with **b** = 1/3[$11\bar{2}0$] and **b** = 1/3[$2\bar{1}\bar{1}0$] (for undoped GaN), and of mixed type with **b** = 1/3[$11\bar{2}3$]. For undoped GaN film, the total dislocation density is about 1×10^{10} cm^{-2}. Doping the GaN film with silicon results in an increase in the densities of practically all types of dislocations. However, dislocations with **b** = 1/3[$2\bar{1}\bar{1}0$] are suppressed. For Si-doped GaN films (10^{17} and 10^{18} cm^{-3}), the total dislocation density is about 2×10^{10} cm^{-2}. All types of GaN films contain the following 3D defects: inversion domains with $\{01\bar{1}0\}$ and $\{11\bar{2}0\}$ faces, stacking faults in the (0001) basal plane, and nanopipes of hexagonal form with lateral dimensions ranging from 5 to 25 nm. The grains in the GaN matrix are bounded by lower-energy inversion domains that nucleate at the interfacial defects. The inversion domains have Ga-polarity, different from the N-polarity of the GaN matrix. Upon doping with silicon, the dimensions of the GaN grains bounded by inversion domains decrease from 400–500 nm for undoped GaN to 300–400 and 200–300 nm for GaN with silicon concentrations of 10^{17} and 10^{18} cm^{-3}, respectively. No correlation between the densities of threading dislocations and nanopipes was observed. Presumably, the formation of nanopipes is to a large extent governed by the mechanism of stress relief at the Al_2O_3/GaN interface. There exists a certain intermediate concentration of silicon dopant (10^{17} cm^{-3}) at which the steplike relief of the GaN film surface is smoothed out.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research, project no. 98-02-18109.

REFERENCES

1. S. Nakamura, M. Senoh, S. Nagahama, *et al.*, Jpn. J. Appl. Phys. **35**, L74 (1996).
2. X. H. Wu, L. M. Brown, D. Kapolnek, *et al.*, J. Appl. Phys. **80**, 3228 (1996).
3. J.-L. Rouvier, M. Arlery, and A. Bourret, in *Proceedings of Royal Microscopical Society Conference on Microscopy of Semiconducting Materials, Oxford, 1997*, Inst. Phys. Conf. Ser. **157**, 173 (1997).
4. D. Cherns, W. T. Young, M. A. Saunders, *et al.*, in *Proceedings of Royal Microscopical Society Conference on Microscopy of Semiconducting Materials, Oxford, 1997*, Inst. Phys. Conf. Ser. **157**, 187 (1997).
5. V. Potin, P. Ruterana, G. Nouet, *et al.*, in *Proceedings of Royal Microscopical Society Conference on Microscopy of Semiconducting Materials, Oxford, 1997*, Inst. Phys. Conf. Ser. **157**, 191 (1997).
6. D. M. Tricker, M. K. H. Natusch, C. B. Boothroyd, *et al.*, in *Proceedings of Royal Microscopical Society Conference on Microscopy of Semiconducting Materials, Oxford, 1997*, Inst. Phys. Conf. Ser. **157**, 217 (1997).
7. S. Ruvimov, Z. Liliental-Weber, T. Suski, *et al.*, Appl. Phys. Lett. **69**, 990 (1996).
8. Z. Liliental-Weber, Y. Chen, S. Ruvimov, and J. Washburn, Phys. Rev. Lett. **79**, 2835 (1997).

Translated by D. Mashovets

Semiconductors, Vol. 34, No. 8, 2000, pp. 872–879. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 909–916.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Properties of Precisely Compensated Semiconductors

S. Zh. Karazhanov

Beruni Physicotechnical Institute, Academy of Sciences of Uzbekistan, Tashkent, 700084 Uzbekistan

Submitted April 5, 1999; accepted for publication January 27, 2000

Abstract—Properties of precisely compensated semiconductors were studied; a phenomenon consisting of a drastic increase in resistivity by several orders of magnitude as the concentration of deep-level impurities increased was observed. It is shown that an anomalous increase in the lifetime of charge carriers, in photoconductivity, and in dark resistivity are related effects. It is demonstrated that, in this case, the semiconductor becomes sensitive to variations in temperature and in the intensity of illumination in the range of interband and impurity absorption; however, these effects are inertial. The highest value of the carriers' lifetime is limited by the band-to-band Auger recombination, for which an empirical expression is suggested. The origin of the above effects is examined and explained in the context of the Shockley–Read–Hall theory of recombination by the example of silicon doped with indium. It is shown that the degree of compensation of a semiconductor can be determined from the temperature dependence of the equilibrium charge-carrier concentration. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

According to generally accepted concepts, an increase in the concentration of deep-level impurities brings about an increase in the rate of recombination of charge carriers and, correspondingly, a decrease in the carriers' lifetime, which constrains the performance of a number of semiconductor devices (for example, it affects the efficiency of solar cells, the dark saturation current of diodes, and so on). These concepts changed drastically after the effects of a giant increase in the lifetime of electrons (τ_n) and holes (τ_p) [1–4] and in the photoconductivity [5] were discovered. The essence of these effects consists in the fact that, as the concentration of recombination centers N_t increases in a narrow range of N_t close to the concentration of shallow-level impurities, the lifetimes τ_n and τ_p and the photoconductivity increase by several orders of magnitude. As a result, the recombination rate U for the charge carriers decreases. In this case, other recombination processes (in particular, band-to-band Auger recombination, which was not considered in [1–5]), rather than those involving the deep impurity levels, may become predominant. In addition, the results reported in [1–5] are valid only for low illumination intensities, such that the concentration of excess carriers is much lower than the majority-carrier concentration.

The objective of this work was to study the lifetimes of electrons and holes in semiconductors with allowance made for the Auger recombination and under the conditions of an arbitrary intensity of excitation. We found that, in this case, an increase in the dark resistivity and a decrease in the total charge-carrier concentration by several orders of magnitude can be observed; an increased sensitivity of the semiconductor to variations in the illumination intensity in the range of the interband and impurity absorption can also be observed.

2. THEORETICAL MODEL

We consider single-crystal silicon that contains shallow donors with concentration N_d and deep acceptors, which have the concentration N_t and are represented by indium impurity. Since the indium atom in silicon is a single-level center, the rate of recombination (U) via this center, as well as the corresponding lifetimes of electrons (τ_n^*) and holes (τ_p^*), can be estimated in the context of the Shockley–Read–Hall recombination theory as

$$U = \frac{N_t C_n C_p (np - n_i^2)}{C_n(n + n_1) + C_p(p + p_1)}, \quad (1)$$

$$\tau_n^* = \frac{\Delta n}{U}, \quad (2)$$

$$\tau_p^* = \frac{\Delta p}{U}. \quad (3)$$

Here, $n = n_0 + \Delta n$ and $p = p_0 + \Delta p$, n_0 and p_0, and Δn and Δp are the total, equilibrium, and excess concentrations of electrons and holes, respectively; the recombination coefficients for electrons and holes C_n and C_p were estimated [6–8] using the empirical formulas

$$C_n = 6 \times 10^{-9} T^{-0.5}, \quad (4)$$

$$C_p = 8 \times 10^{-4} T^{-1.5}; \quad (5)$$

$n_i = (N_c N_v)\exp[-E_g/2kT]$ is the intrinsic concentration of charge carriers; T is temperature of the sample; $N_c =$

$3 \times 10^{19}(T/300)^{1.5}$ and $N_v = 10^{19}(T/300)^{1.5}$ are the densities of states of electrons in the conduction band and holes in the valence band; k is the Boltzmann constant; and E_g is the band gap that varies with the temperature [9] and doping level [10] as

$$E_g(T, N_d) \approx E_g(T = 0, N_d = 0) - 0.00024T - 0.148(N_d \times 10^{-20})^{0.325}, \quad (6)$$

where $n_1 = N_c \exp[(\Delta E - E_g)/kT]$ and $p_1 = N_v \exp[-\Delta E/kT]$ are the Shockley–Read statistical factors, and $\Delta E = 0.156$ eV is the depth of the impurity energy level measured from the top of the valence band.

At present, we may posit that the lifetime of charge carriers in single-crystal silicon, which is almost devoid of deep-level impurities, is controlled by Auger recombination [11]. In the case where the concentration of deep-level impurities is fairly high, the lifetimes are controlled by two recombination processes; one of these is recombination via single-energy deep-level centers and the other is band-to-band Auger recombination. As a result, the total lifetimes τ_n and τ_p are given by [11]

$$\frac{1}{\tau_n} = \frac{1}{\tau_n^*} + \frac{1}{\tau_A}, \quad (7)$$

$$\frac{1}{\tau_p} = \frac{1}{\tau_p^*} + \frac{1}{\tau_A}, \quad (8)$$

where τ_A is the lifetime controlled by band-to-band Auger recombination. Using the experimental data [11], we can easily derive the following empirical expression for τ_A:

$$\tau_A = 2.86 \times 10^{17} N_d^{-1.3} \exp(0.36 N_d^{-0.1} T); \quad (9)$$

this formula is valid in the ranges of $N_d = 10^{16}$–10^{18} cm^{-3} and $T = 100$–500 K.

The equilibrium and excess concentrations of electrons and holes are determined from the equation for the rates of (i) carriers' capture by impurity levels [1] and (ii) their reverse thermal emission from these levels and from the condition for complete electroneutrality; in the case under consideration, the latter condition can be written as

$$p + N_d = n + N_t^-. \quad (10)$$

Here, $N_t^- = N_t - N_t^0$, where N_t^0 and N_t^- are the concentrations of neutral and charged impurities, respectively. The relation between the quantities N_t, N_t^0, and N_t^- can be derived from the expressions $N_t^0 = N_t(1 - f)$ and $N_t^- = N_t f$, where f is the occupancy of the deep impurity state with electrons; according to the Shockley–Read–Hall recombination theory, we have

$$f = \frac{C_n n + C_p p_1}{C_n(n + n_1) + C_p(p + p_1)}. \quad (11)$$

The total mobilities μ_n and μ_p of the charge carriers are controlled by scattering from thermal lattice vibrations (μ_n^* and μ_p^*) and also from neutral (μ_A) and charged (μ_I) impurities (see, for example, [12]); thus, we have

$$\frac{1}{\mu_n} = \frac{1}{\mu_n^*} + \frac{1}{\mu_A} + \frac{1}{\mu_I}, \quad (12)$$

$$\frac{1}{\mu_p} = \frac{1}{\mu_p^*} + \frac{1}{\mu_A} + \frac{1}{\mu_I}, \quad (13)$$

where

$$\mu_A = 1.56 \times 10^{20}(N_t^0)^{-1}, \quad (14)$$

$$\mu_I = \frac{3.16 \times 10^{15} T^{1.5}}{N_t \ln(1 + 2 \times 10^4 T N_t^{-1/3})}, \quad (15)$$

$$\mu_n^* = 1300(300/T)^2, \quad (16)$$

$$\mu_p^* = 500(300/T)^{2.7}. \quad (17)$$

Resistivity is defined by the conventional formula

$$\rho = (q\mu_n n + q\mu_p p)^{-1}, \quad (18)$$

where q is the elementary charge.

We now consider the conventional single-crystal silicon and evaluate the dependences of the lifetimes of majority and minority charge carriers, mobility, and the resistivity of this material on the indium concentration in the range of $N_t = 10^{16}$–10^{18} cm^{-3} for various values of the shallow-donor concentration within $N_d = 10^{16}$–10^{18} cm^{-3} in a wide temperature range of $T = 200$–500 K.

3. RESULTS AND DISCUSSION

3.1. Lifetimes of Electrons and Holes

The lifetimes of electrons (τ_n) and holes (τ_p) are calculated with formulas (1)–(11) under the conditions of low injection level $\Delta n = 0.1 n_i$ (n_i is the intrinsic electron concentration) at $T = 260$–500 K for $N_d = 10^{16}$, 10^{17}, and 10^{18} cm^{-3}. Figure 1 shows the dependence of τ_n on N_t at $T = 300$ K for the concentration of shallow-level donors $N_d = 10^{17}$ cm^{-3}. An analysis of the obtained results showed that τ_n and τ_p initially increased with increasing deep-trap concentration N_t for $N_t < N_d$ and then decreased for $N_t > N_d$ (Fig. 1). The largest values $\tau_{n\max}$ and $\tau_{p\max}$ are observed for $N_t = N_d$. This signifies that

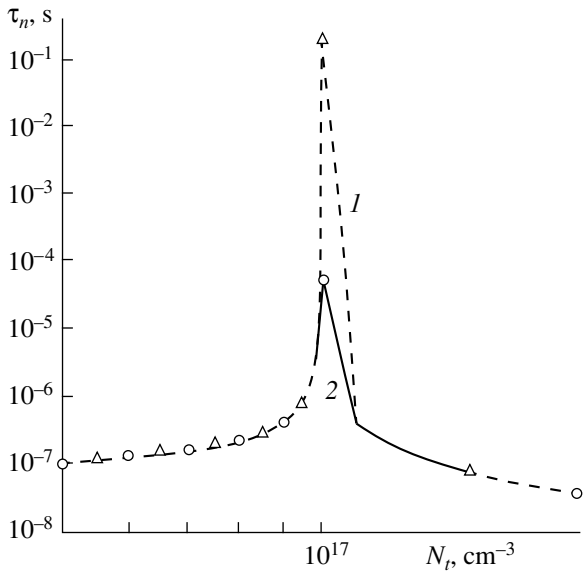


Fig. 1. The dependences of electron lifetime on the recombination-center concentration N_t at $T = 300$ K for $\Delta n = 0.1n_i$ and the doping level of $N_d = 10^{17}$ cm^{-3} (*1*) without and (*2*) with consideration for the band-to-band Auger recombination.

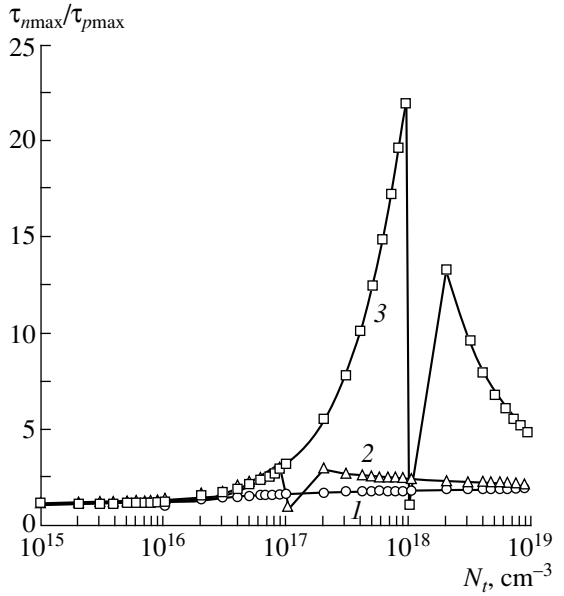


Fig. 2. The dependences of the ratio between the maximal lifetimes of electrons $\tau_{n\max}$ and holes $\tau_{p\max}$ on the concentration of deep-level centers N_t for $\Delta n = 0.1n_i$ and the concentrations of shallow-level donors N_d = (*1*) 10^{16}, (*2*) 10^{17}, and (*3*) 10^{18} cm^{-3}.

the observed effect constitutes a characteristic property of a precisely compensated semiconductor. This inference is consistent with the results obtained previously [1–5].

Figure 1 shows only the dependence $\tau_n(N_t)$, because the dependence $\tau_p(N_t)$ is similar and qualitatively replicates the trend in $\tau_n(N_t)$. However, the numerical values of τ_n and τ_p may differ markedly from one another (Fig. 2). We studied the dependence of the ratio $\tau_{n\max}/\tau_{p\max}$ on N_t at $T = 300$ K for $\Delta n = 0.1n_i$ and $N_d = 10^{16}$, 10^{17}, and 10^{18} cm^{-3} (Fig. 2). An analysis of the data in Fig. 2 shows that, for $N_d \le 10^{17}$ cm^{-3}, $\tau_n \approx \tau_p$ in the entire range of N_d under consideration; at the same time, τ_n and τ_p differ markedly from one another for $N_d > 10^{17}$ cm^{-3}. Consequently, in such cases, it is necessary to take into account the difference between the lifetimes of electrons and holes in a theoretical study of transport phenomena in a silicon structure.

In order to gain insight into the causes of this phenomenon (Fig. 2), we divide τ_n^* (2) by τ_p^* (3):

$$\frac{\tau_n^*}{\tau_p^*} = \frac{\Delta n}{\Delta p}. \tag{19}$$

This formula signifies that the difference between τ_n^* and τ_p^* is due to the difference between excess concentrations Δn and Δp that are related to trapping by deep-level impurities. This can be easily verified by analyzing the relation between Δn and Δp,

$$\Delta n \approx \Delta p\left[1 + \frac{N_t p_1}{(p + p_1)(p_0 + p_1)}\right], \tag{20}$$

obtained from the condition for electroneutrality. We perform the analysis for the case of $p_0 \ll p_1$ and $N_d > N_t$. Expression (20) then becomes simplified, and, for $\Delta p < p_1$, we have $\Delta n \approx \Delta p(1 + N_t/p_1)$. Hence, for $N_t < p_1$, Δn differs only slightly from Δp; correspondingly, the difference between τ_n^* and τ_p^* is insignificant. Since we have $p_1 \approx 4.7 \times 10^{16}$ cm^{-3} for indium impurity in silicon, any significant difference between τ_n^* and τ_p^* should not be expected for $N_t = N_d = 10^{16}$ cm^{-3}, which, in fact, can be clearly seen in Fig. 2. For $N_t > p_1$ and $N_d > p_1$, the differences between Δn and Δp (20) and also between τ_n^* and τ_p^* (19) increase for $N_t < N_d$, which accounts for the results shown in Fig. 2 for $N_t < N_d = 10^{17}$ and 10^{18} cm^{-3}.

For $N_t \ge N_d$, p_0 increases sharply with increasing N_t and p_0 eventually becomes larger than p_1. Expression (20) is now simplified and can be written as

$$\Delta n \approx \Delta p\left(1 + \frac{N_t p_1}{p_0^2}\right). \tag{21}$$

Since p_0 increases with increasing N_t for $N_t > N_d$, the differences between Δn and Δp and between τ_n^* and τ_p^* decrease, which accounts for the results shown in Fig. 2 in the case of $N_t > N_d = 10^{17}$ and 10^{18} cm^{-3}.

It can be easily seen in Fig. 2 that the ratio $\tau_{n\max}/\tau_{p\max}$ decreases sharply to unity for $N_t = N_d = 10^{17}$ and 10^{18} cm^{-3}. Such an effect is possible in the case where the charge-carrier lifetime related to deep-level impurities becomes larger than the lifetime controlled by band-to-band Auger recombination. In this situation, it is Auger recombination that yields $\tau_n \approx \tau_p$.

It is noteworthy that the results illustrated in Figs. 1 and 2 are consistent only qualitatively (and not quantitatively) with results reported previously [1–5]. The largest values of lifetimes $\tau_{n\max}$ and $\tau_{p\max}$ were studied in [1–5] without considering the band-to-band Auger recombination, while the latter was taken into account in this study. It was found that the values of $\tau_{n\max}$ and $\tau_{p\max}$ estimated with τ_A disregarded were about three orders of magnitude larger than those calculated taking into account the lifetime τ_A (Fig. 1).

It is worth noting that the lifetimes τ_n^* (2) and τ_p^* (3) were studied in [1] for low excitation levels (Δn and $\Delta p \ll n_0$ and p_0) and almost all of the most important cases were considered analytically. In connection with this, we concentrated our attention on numerical simulation of the dependences of τ_n (7) and τ_p (8) on the concentration of excess carriers Δn and Δp for large Δn and Δp at T = 300 K and for N_t and N_d in the range of 10^{16}–10^{18} cm^{-3}, in which case the dependences of τ_n and τ_p on N_t become nonlinear. The results are shown in Fig. 3 in the form of the dependence $\tau_n(N_t)$ for $N_d = 10^{17}$ cm^{-3} and $\Delta n/n_i$ = 0.1, 1, and 10^4. It can be seen from Fig. 3 that $\tau_{n\max}$ (and, similarly, $\tau_{p\max}$) decreases with an increasing excitation level. For $\Delta n \approx 10^4 n_i$, the peaks in the curves $\tau_n(N_t)$ and $\tau_p(N_t)$ almost cease to exist. An analysis of the curves in Fig. 3 shows that $\tau_{n\max}$ and $\tau_{p\max}$ are very sensitive to variations in the intensity of weak optical radiation. This fact should be also taken into account in studying the transport properties of semiconductor structures.

3.2. Equilibrium Concentration of Charge Carriers

In this subsection, we interpret the obtained dependences of τ_n and τ_p on N_t, which, in contrast to inferences in [1–5], are related to a decrease in the total concentration of electrons and holes $n_0 + p_0$. An analysis of expressions (1)–(3) shows that the cause of the obtained dependences of τ_n^* and τ_p^* on N_t consists in the fact that $n_0 + p_0$ decreases by several orders of magnitude. We studied the dependence of the concentrations of electrons and holes on the concentration of

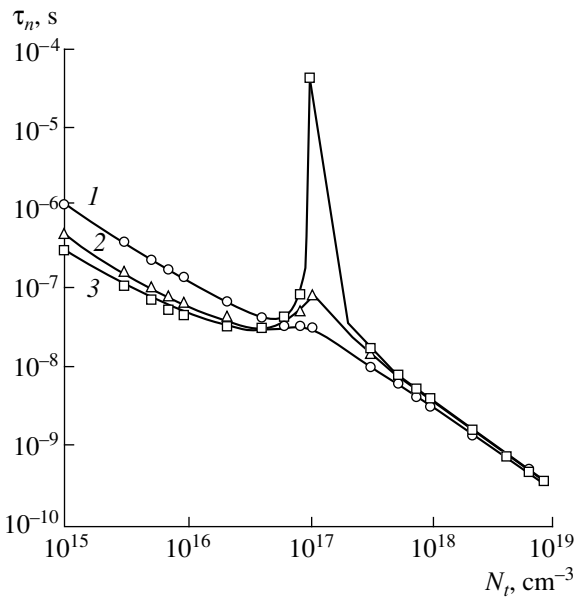


Fig. 3. The dependences of electron lifetime on the recombination-center concentration N_t at T = 300 K for the doping level of $N_d = 10^{17}$ cm^{-3} and the injection levels of $\Delta n/n_i$ = (*1*) 10^4, (*2*) 10, and (*3*) 0.1.

recombination centers; the results are shown in Fig. 4. If $N_d > N_t$, the majority charge carriers are electrons; correspondingly, we have $p_0 \ll n_0 \approx N_d$. In this case, n_0 decreases only slightly with increasing N_t. If $N_t \le N_d$, the total concentration $n_0 + p_0$ decreases sharply by several orders of magnitude and becomes lower than that for $N_t < N_d$ and $N_t > N_d$. In the case of $N_t = N_d$, all free electrons supplied by shallow-level donors are trapped by deep-level acceptors. Since the number of recombining free electrons and holes is small, the rate of recombination via deep-level impurities N_t (1) is low and the corresponding lifetime is large (Fig. 1).

An analysis of the data in Fig. 4 shows that, for $N_d < N_t$, the holes thermally generated from the deep level of impurity become the majority charge carriers and an conductivity-type inversion occurs. As a result, the dependence of the total concentration $n_0 + p_0$ passes through a minimum. The latter ceases to exist if the level of the deep-lying acceptor center is located above the midgap. In this case, the majority-carrier concentration decreases sharply to the value defined by electrons generated thermally from the level of the deep-lying impurity center; however, no minimum in the dependence of $n_0 + p_0$ on N_t would be observed. This means that the results shown in Figs. 1–4 depend strongly on the position of the energy level of the deep-lying acceptor center in the band gap of semiconductors doped also with low-level donors. In what follows, we analyze certain important cases.

(i) If the energy level of deep-lying acceptor is located within the lower half of the band gap (for exam-

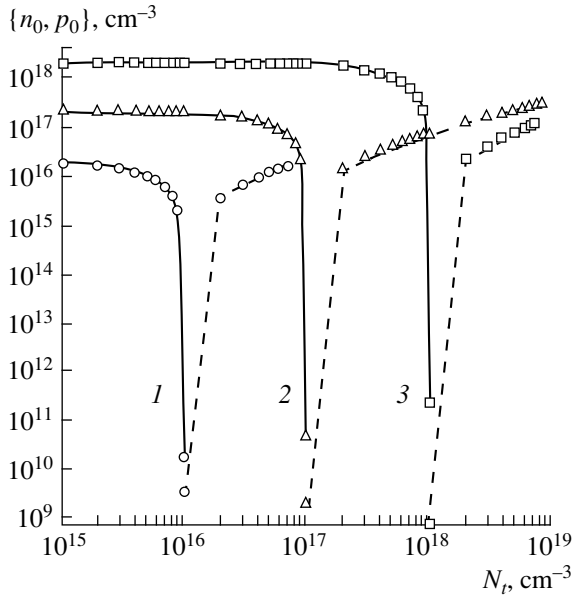


Fig. 4. Dependences of the concentration of equilibrium electrons (the solid-line curves) and holes (the dashed-line curves) on the recombination-center concentration N_t at T = 300 K for $\Delta n = 0.1 n_i$ and the doping levels N_d = (*1*) 10^{16}, (*2*) 10^{17}, and (*3*) 10^{18} cm^{-3}.

ple, in the case of indium or thallium impurities in silicon), we have $n_1 \ll p_1$. Correspondingly, the comparison of lifetimes of electrons τ_n^* and holes τ_p^* for low injection levels [$\Delta n \approx \Delta p \ll (p_i, N_d)$] in the case of $N_d > N_t$,

$$\tau^* \approx \tau_{n0} \frac{p_1}{n_0 + p_0}; \tag{22}$$

in the case of $N_d < N_t$,

$$\tau^* \approx \tau_{n0} \left(1 + \frac{p_1}{p_0}\right); \tag{23}$$

and, in the case of $N_d \approx N_t$,

$$\tau^* \approx \tau_{n0} \frac{p_1}{2 p_i}, \tag{24}$$

showed that the lifetime is larger by several orders of magnitude for $N_d = N_t$ than for $N_d > N_t$ and $N_d < N_t$, which accounts for the results shown in Fig. 1. The rate of band-to-band Auger recombination may then become the dominant factor and may control the largest value of the lifetime (Fig. 1, the solid line).

As mentioned above, expression (24) was derived for low injection levels, i.e., for $\Delta n \approx \Delta_p \leq n_i$. For large Δn and Δp ($\Delta n \approx \Delta p > n_i$), this expression can be written as $\tau^* \approx \tau_{n0} 0.5 p_1 (p_0 + \Delta p)^{-1}$. Hence, it follows that the maximal value of lifetime decreases with an increasing excitation level and vanishes completely for Δp comparable to p_1, which accounts for the results shown in Fig. 3.

(ii) If the level of the deep-lying acceptor is located within the upper half of the band gap, as is the case with E centers in silicon, we have $n_1 \gg p_1$. As a result, the expressions for the lifetimes of electrons (2) and holes (3) may be simplified as

$$\tau_n^* \approx \tau_{p0} \frac{n + n_1}{n_0 \Delta p / \Delta n + p}, \tag{25}$$

$$\tau_p^* \approx \tau_{p0} \frac{n + n_1}{n + p_0 \Delta n / \Delta p}. \tag{26}$$

For low concentrations of deep-level impurity ($N_t < N_d$), in view of expressions (19) and (20), the excess concentrations and lifetimes are almost the same for electrons and holes. Consequently,

$$\tau_n^* \approx \tau_p^* \approx \tau_{p0} \frac{n + n_1}{n_0 + p_0}. \tag{27}$$

The concentration n_0 decreases with increasing N_t (Fig. 4), and, according to relationship (20), the ratio $\Delta n/\Delta p$ increases. Consequently, the lifetime of electrons (25) increases, whereas the lifetime of holes (26) decreases. Once $N_t > N_d$, the equilibrium concentration of holes p_0 becomes larger than p_1 and increases with increasing N_t, which, in view of (20), causes the ratio $\Delta n/\Delta p$ to decrease. Consequently, τ_n^* (25) decreases, whereas τ_p^* (26) may increase. We performed the study of lifetimes related to the E center in silicon; the results are consistent with the above inferences.

(iii) If a deep-level acceptor is a trapping center whose level is located in close proximity to the conduction-band bottom (as, e.g., in the case of the A center in silicon), we have $p_1 \ll (n_1, p_0)$. In this case, according to (20), the ratio $\Delta n/\Delta p$ decreases with increasing N_t. Consequently, τ_n^* and τ_p^* may diminish with increasing N_t and do not exhibit a maximum. We studied the dependence of lifetimes on the concentration of A centers in silicon; the results are consistent with the above inference.

In addition, we studied the temperature dependence of n_0 and p_0 for the three cases of $N_t = N_d = 10^{16}$, 10^{17}, and 10^{18} cm^{-3}, and also for the case of $N_d \neq N_t$, namely, $N_d = 1.7 \times 10^{13}$ cm^{-3} and $N_t = 8.15 \times 10^{15}$ cm^{-3} (see Fig. 5). For $N_t \neq N_d$, the dependence of the charge-carrier equilibrium concentration is a steadily decreasing function of reciprocal absolute temperature, whereas, for $N_t = N_d$, this dependence is more complex (Fig. 5). This distinction may be used as an experimental indication of whether the semiconductor is or is not precisely compensated.

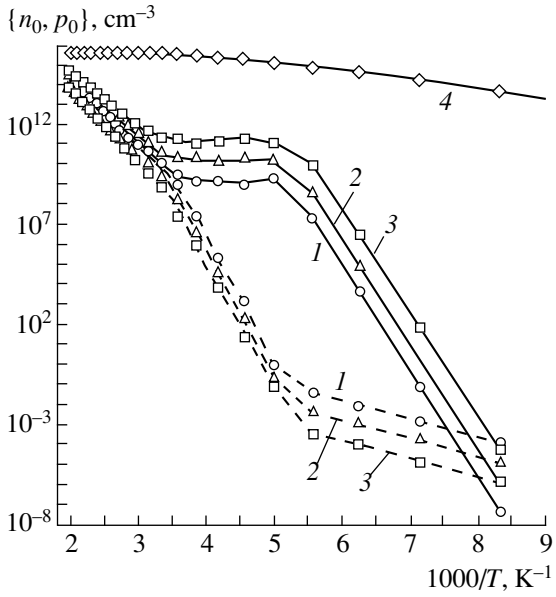


Fig. 5. The temperature dependences of the equilibrium concentration of electrons (the solid-line curves) and holes (the dashed-line curves) for $N_d = N_t = (1)\ 10^{16}$, $(2)\ 10^{17}$, and $(3)\ 10^{18}$ cm^{-3} and for $(4)\ N_d = 1.7 \times 10^{13}$ cm^{-3} and $N_t = 8.15 \times 10^{15}$ cm^{-3} [6–8].

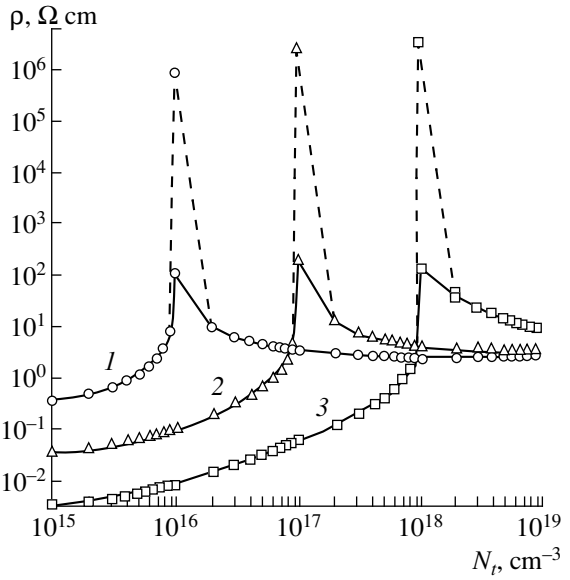


Fig. 6. Dependences of resistivity on the recombination-center concentration N_t at $T = 300$ K for the doping levels of $N_d = (1)\ 10^{16}$, $(2)\ 10^{17}$, and $(3)\ 10^{18}$ cm^{-3} and for the injection levels of $\Delta n/n_i = 0.1$ (the dashed-line curve) and 10^4 (the solid-line curve).

3.3. Resistivity

It is known that a decrease in the charge-carrier concentration $n_0 + p_0$ by several orders of magnitude (Fig. 4) signifies that the resistivity of a semiconductor increases. We studied the dependence of resistivity on the trap concentration N_t taking into account the temperature dependence of mobilities μ_n and μ_p [see (12)–(17)] at $T = 300$ K for $N_d = 10^{16}$, 10^{17}, and 10^{18} cm^{-3} and for the injection levels amounting to $\Delta n/n_i = 0.1$ and 10^4. The results are shown in Fig. 6. It can be seen that, for $N_d < N_t$, resistivity of a semiconductor ρ increases only slightly with increasing N_t and increases sharply (by several orders of magnitude) for $N_t = N_d$. If $N_t < N_d$, ρ decreases only slightly with increasing N_t. This special feature of the dependence $\rho(N_t)$ can be also used to determine the degree of compensation of a semiconductor. The performed study of temperature dependence of resistivity shows that $\rho(T)$ reproduces the dependence $n_0(T) + p_0(T)$. The dependences of mobilities μ_n and μ_p on the concentrations of neutral and charged impurity and on temperature do not introduce any radical changes in the trend of $\rho(T)$. The temperature coefficient of the maximal resistivity $Q_T \equiv \rho_{\max}^{-1} d\rho_{\max}/dT$ at $T > 300$ K is comparable to the corresponding coefficient for thermoresistors made of silicon doped with manganese [13].

3.4. Photosensitivity

On analyzing (1)–(11), we can easily verify that large values of lifetimes and dark resistivity are extremely sensitive to variations in the concentration of excess charge carriers; i.e., to variations in the photoexcitation intensity (Figs. 3, 6). The value of $\rho_{\max}$ decreases with increasing Δn within the range of $n_i \le \Delta n \le N_d$ (Fig. 6); this range can be expanded by increasing the level of doping with shallow-level donors N_d. The lower bound of photosensitivity can be also shifted by decreasing n_i. As a rule, this can be accomplished by employing an external pressure or using a semiconductor with a wider band gap.

An analysis of formulas (1)–(3) shows that the maximal value of the charge-carrier lifetime decreases with an increasing concentration of excess charge carriers (Fig. 3) in the range of $n_0 \le \Delta n \le p_1$. In this case, the smaller the depth of the recombination-center energy level, the larger is p_1 and, correspondingly, the larger is the upper limit of sensitivity to variations in Δn.

We studied the dependences of $\tau_{n\max}$, $\tau_{p\max}$, and $\rho_{\max}$ on the intensity of impurity-absorption illumination with photons of the energy $h\nu \ge E_g - \Delta E$; these photons induce the photostimulated transition of electrons from the impurity level to the conduction band. To this end, we modified the expressions for the recombination rate U (1) and for the charge-carrier lifetimes τ_n^* (2) and τ_p^*

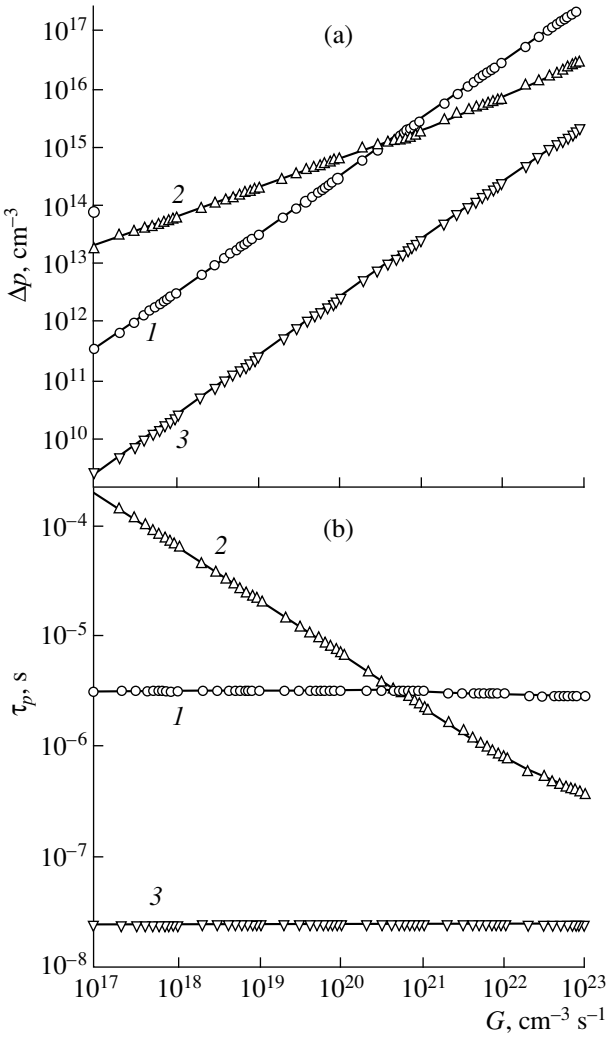


Fig. 7. Dependences of (a) excess concentration and (b) the lifetime of holes on the generation rate G of photocarriers in silicon with concentration of shallow-level donors N_d = 10^{16} cm^{-3} and the concentration of deep-level acceptors N_t = (*1*) 10^{15}, (*2*) 10^{16}, and (*3*) 10^{17} cm^{-3} at T = 300 K.

(3). In this case, similarly to [14], we use the following expressions for p_1 and n_1 in (1):

$$n_1^* = n_1 + \tau_{n0} G_n, \tag{28}$$

$$p_1^* = p_1 + \tau_{p0} G_p. \tag{29}$$

Here, G_n and G_p are the generation rates for electrons excited from the impurity level to the conduction band and from the valence band to the impurity level.

Estimation based on formulas (28) and (29) showed that $p_1 \gg \tau_{p0} G_p$. This signifies that the thermally stimulated electron exchange between the level of deep-lying impurity and the valence band dominates over the photostimulated exchange with the same level involved. In addition, we have $n_1 \ll \tau_{n0} G_n$. Correspondingly, the rate of photostimulated electron exchange between the deep-level impurity and the conduction band predominates over the rate of thermally stimulated exchange.

The study was performed for the temperature range of T = 100–500 K and for the doping levels of N_d = 10^{16}–10^{18} cm^{-3} and for the recombination-center concentrations of N_t = 10^{16}–10^{18} cm^{-3}. An analysis of the results demonstrated high sensitivity of maximal values of lifetime and resistivity to variations in the intensity (ϕ) of illumination within the range of impurity absorption, with the sensitivity coefficients given by $Q_\tau \equiv \tau_{n\max}^{-1} d\tau_{n\max}/d\phi \approx 12$ and $Q_\rho \equiv \rho_{n\max}^{-1} d\rho_{n\max}/d\phi \approx 10$. It is noteworthy that the values of Q_τ and Q_ρ are of the same order of magnitude as the corresponding coefficients for photoresistors made of silicon doped with manganese [13].

It is worth noting that the results shown in Figs. 1–6 were obtained for different excess concentrations of minority charge carriers Δp without specifying how these carrier were generated. In fact, the value of Δp can be varied using the injection from the contact or illumination. We further studied the dependence of excess concentration Δp and the lifetime τ_p on the illumination intensity for silicon with the concentration of shallow-level donors of N_d = 10^{16} cm^{-3} and with the concentrations of deep-level indium impurity in the range of N_t = 10^{15}, 10^{16}, and 10^{17} cm^{-3} at room temperature. To this end, we solved the time-independent kinetic equation by equating the recombination rate U (1) with the photogeneration rate G for charge carriers ($U = G$); this equation was solved simultaneously with the electroneutrality equation (10) using iterations. This method was used to determine the concentration of holes; then, with formulas (1)–(11), the hole lifetime τ_p was determined as a function of the photogeneration rate G for charge carriers. The results are shown in Fig. 7. It can be seen from Fig. 7 that the excess concentration and lifetime of charge carriers in a precisely compensated semiconductor ($N_t = N_d$) are larger than those in undercompensated ($N_t < N_d$) and overcompensated ($N_t > N_d$) semiconductors. Such an effect is especially well pronounced under illumination with intensity much lower than that of the AM1-illumination intensity of 100 mW/cm^2. Consequently, the semiconductor is sensitive to variations in the low-intensity illumination, which validates the results obtained in this study and is consistent with the data reported in [5].

4. CONCLUSION

Thus, the dark resistivity of a semiconductor may increase by several orders of magnitude with increasing concentration of deep-level impurity. A comparison of the results of this study with previous theoretical results [1–5] shows that the phenomena of an anomalous increase in the dark resistivity, in the lifetime of charge carriers, and in photoconductivity are related and are observed only in a precisely compensated semiconduc-

tor. Such a semiconductor is also found to be highly sensitive to variations in temperature and in the intensity of illumination with light in the ranges of band-to-band and impurity absorption. The origin of the above phenomena is related to a drastic decrease (by several orders of magnitude) in equilibrium charge-carrier concentration with an increasing concentration of deep-level centers. It is shown that the temperature dependence of the equilibrium charge-carrier concentration in a precisely compensated semiconductor ($N_d = N_t$) differs radically from the corresponding dependence in an uncompensated semiconductor. This distinction can be used as an indicator to characterize the degree of compensation.

It should be emphasized that, under the condition of $N_d = N_t$, band-to-band Auger recombination becomes very important without exception and reduces the maximal value of the lifetime by several orders of magnitude (see the dashed- and solid-line curves in Fig. 1).

The phenomenon under consideration depends strongly on the position of the energy level of the deep-lying acceptor in the band gap. If this level is in the lower half of the band gap, the dependences of the lifetimes of electrons τ_n and holes τ_p on the acceptor concentration exhibit a maximum. If the acceptor level is located in the upper half of the band gap close to the midgap, the dependence $\tau_n(N_t)$ may exhibit a maximum, whereas $\tau_p(N_t)$ may go through a minimum. If the deep level is located closer to the conduction band, both $\tau_n(N_t)$ and $\tau_p(N_t)$ may be steadily decreasing functions. These inferences are valid for a semiconductor with shallow-level donors and deep-level acceptors provided that the above conditions are satisfied.

In conclusion, we note that the high sensitivity of a completely compensated semiconductor to external effects is widely used in the technology of thermally sensitive resistors, photoresistors, and magnetoresistors [13], as well as of temperature sensors and pressure and magnetic transducers [15].

ACKNOWLEDGMENTS

This work was supported by the Academy of Sciences of Uzbekistan. I thank DAAD (Germany) for reimbursement of the expenses related to my travel to the Institut für Solarenergieforschung Hameln (Emmerthal, Germany). I am grateful to Professor P.T. Landsberg (Southampton, UK) and Professor A.Yu. Leiderman for their helpful participation in discussions, and also to É.N. Tsoi and A.S. Kirakosyan (Physicotechnical Institute, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan) for technical assistance and critical comments.

This study was performed in the framework of the NATO Linkage Grants according to the contract PST SLG 975758.

REFERENCES

1. A. A. Drugova and V. A. Kholodnov, Pis'ma Zh. Tekh. Fiz. **18** (1), 23 (1992) [Sov. Tech. Phys. Lett. **18**, 8 (1992)].
2. V. A. Kholodnov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30** (6), 1011 (1996) [Semiconductors **30**, 538 (1996)].
3. A. A. Drugova and V. A. Kholodnov, Solid-State Electron. **38** (6), 1247 (1995).
4. V. A. Kholodnov and P. S. Serebrennikov, Pis'ma Zh. Tekh. Fiz. **23** (7), 39 (1997) [Tech. Phys. Lett. **23**, 268 (1997)].
5. V. A. Kholodnov, Pis'ma Zh. Éksp. Teor. Fiz. **67** (9), 655 (1998) [JETP Lett. **67**, 685 (1998)].
6. A. G. Milnes, *Deep Impurities in Semiconductors* (Wiley, New York, 1973; Mir, Moscow, 1977).
7. J. S. Blakemore and C. E. Sarver, Phys. Rev. **173**, 767 (1968).
8. G. K. Wertheim, Phys. Rev. **109**, 1086 (1958).
9. J. D. Arora and P. C. Mathur, J. Appl. Phys. **52** (5), 3646 (1981).
10. D. J. Roulston, N. D. Arora, and S. C. Chamberlain, IEEE Trans. Electron Devices **29** (2), 284 (1982).
11. P. P. Altermatt, J. Schmidt, G. Heiser, and A. G. Aberle, J. Appl. Phys. **82**, 4938 (1997).
12. K. V. Shalimova, *The Physics of Semiconductors* (Énergoatomizdat, Moscow, 1985).
13. A. N. Marchenko, *Controlled Semiconductor Resistors* (Énergiya, Moscow, 1978).
14. M. J. Keevers and M. A. Green, J. Appl. Phys. **75** (8), 4022 (1994).
15. K. S. Ayupov and N. F. Zikrillaev, Dokl. Akad. Nauk Resp. Uzb., No. 8/9, 41 (1992).

Translated by A. Spitsyn

Semiconductors, Vol. 34, No. 8, 2000, pp. 880–885. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 917–922.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Electrical Properties of Semiconductors with Pair Defects

S. Zh. Karazhanov and É. V. Kanaki

Beruni Physicotechnical Institute, Academy of Sciences of Uzbekistan, Tashkent, 700084 Uzbekistan

Submitted July 19, 1999; accepted for publication January 27, 2000

Abstract—Electrical properties of semiconductors with pair defects were investigated. It was demonstrated that the electron lifetime, resistivity, and Hall factor exhibit a pronounced nonmonotonic dependence on the pair defect concentration. A mechanism relating this phenomenon to a precise compensation of the semiconductor is suggested. It is shown that an abrupt drop of the total electron and hole density occurs in this case and the semiconductor becomes sensitive to external effects such as weak illumination intensity in the regions of the impurity and band-to-band absorption, temperature, etc. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

For many semiconductor compounds, a noticeable fraction of point defect is represented by the impurity complexes. In the simplest case, these are the pair defects, for example, the donor–acceptor pairs and impurity–vacancy pairs [1–3]. Intracenter transitions, configuration rearrangements, and an electronic spectrum rich in features are characteristic of these defects. These phenomena make it possible to use complex defects purposefully for controlling the electrical and photoelectric properties of semiconductor materials and fabricating devices of the new type.

Until recently, it has been generally accepted that the electron (τ_n) and hole (τ_p) lifetimes decrease steadily with increasing concentration of deep-level impurities (N_R). However, more detailed investigations [4–7] of recombination via the single-charged [4–6] and double-charged [7] centers demonstrated that the N_R-dependences of τ_n and τ_p can be nonmonotonic and N_R can increase by several orders of magnitude in a certain range of N_R. It is our intention to show that a similar effect is also possible at recombination of electrons and holes via pair defects. A similar problem for the double-charged center, which can be neutral, negatively single-charged, or negatively double-charged, was considered previously. It was demonstrated that the N_R-dependences of the lifetimes τ_n and τ_p exhibit two maxima [7]. In contrast with [7], we consider the pair defect, which can be positively charged, neutral, or negatively charged. In this case, the dependence of the electron lifetime on the double-defect concentration exhibits a single maximum, while a similar dependence of the hole lifetime passes through a minimum.

We note that the pair defects and their emission are found experimentally for many semiconductors that are widely used in modern optoelectronics, for example, for silicon carbide [2, 3], gallium phosphide [1], etc.

2. RECOMBINATION VIA PAIR DEFECTS

Let us consider the point defect [*AB*] formed by two neighboring point components *A* and *B*. Similarly to the donor–acceptor pair, this defect can have three different charged states q_0, q_1, and q_2. Here, q_0 is the charge of unoccupied defect $[A^0B^0]$ and q_1 is the charge of the defect with one additional electron, which is usually localized at one of the components. If the electron is bound by the *B*-component or by the *A*-component, the defect is denoted as $[A^0B^1]$ or $[A^1B^0]$, respectively. The charge of the defect with two additional electrons, which are bound by both components $[A^1B^1]$, is q_2. Thus, in the simplest case, the triple-charged pair defect can occur in four various quantum states, between which the transitions can occur.

Being unoccupied, the $[A^0B^0]$ defect can capture the electron from the conduction band (*c*) or from the valence band (*v*) and convert to one of the single-charged states $[A^0B^1]$ or $[A^1B^0]$. These processes can be described using the following quasi-chemical reactions:

$$[A^0B^0] + e_c \xrightarrow{C_n^{12}} [A^0B^1], \quad [A^0B^0] + e_c \xrightarrow{C_n^{13}} [A^1B^0],$$

$$[A^0B^0] \xrightarrow{e_p^{12}} [A^0B^1] + h_v, \quad [A^0B^0] \xrightarrow{e_p^{13}} [A^1B^0] + h_v.$$

The kinetic coefficients for corresponding reactions per defect and per free carrier are shown above the arrows. The dimension for all C_n^{ij} and C_p^{ij} coefficients is $[L^3T^{-1}]$; for all e_n^{ij} and e_p^{ij} coefficients, it is $[T^{-1}]$. L is the length dimension, and T is the time dimension.

Being single-charged, the defect can capture one more electron from the allowed band and convert to a double-charged state. However, it can also donate the electron to the allowed band and thereby become unoccupied:

$$[A^0B^1] + e_c \xrightarrow{C_n^{24}} [A^1B^1], \quad [A^1B^0] + e_c \xrightarrow{C_n^{34}} [A^1B^1],$$

1063-7826/00/3408-0880$20.00

$$[A^0B^1] + h_\nu \xrightarrow{e_p^{24}} [A^1B^1], \quad [A^1B^0] + h_\nu \xrightarrow{e_p^{34}} [A^1B^1],$$

$$[A^0B^1] \xrightarrow{e_n^{21}} [A^0B^0] + e_c, \quad [A^1B^0] \xrightarrow{e_p^{31}} [A^0B^0] + e_c,$$

$$[A^0B^1] + h_\nu \xrightarrow{C_p^{21}} [A^0B^0], \quad [A^1B^0] + h_\nu \xrightarrow{C_p^{31}} [A^0B^0].$$

Additionally, the electron can be transferred from one component of the single-charged defect to the other. This intracenter transition is described by the reactions $[A^1B^0] \xrightarrow{\nu_{32}} [A^0B^1]$ and $[A^0B^1] \xrightarrow{\nu_{23}} [A^1B^0]$. The dimension of the coefficients ν_{32} and ν_{23} is $[T^{-1}]$. Finally, being in a double-charged state $[A^1B^1]$, the defect can donate the electron to the allowed band and convert to one of single-charged states:

$$[A^1B^1] + h_\nu \xrightarrow{C_p^{42}} [A^0B^1], \quad [A^1B^1] + h_\nu \xrightarrow{C_p^{43}} [A^1B^0],$$

$$[A^1B^1] \xrightarrow{e_n^{42}} [A^0B^1] + e_c, \quad [A^1B^1] \xrightarrow{e_n^{43}} [A^1B^0] + e_c.$$

All transitions described above are schematically shown in Fig. 1. The probabilities of these transitions are also shown in Fig. 1. Here, a_{ij} is the probability that the defect in the ith state converts to the jth state in unit time. The dimension of the quantity a_{ij} is $[T^{-1}]$. For example, a_{12} is the probability of converting the unoccupied defect $[A^0B^0]$ to the $[A^0B^1]$ state in unit time. The transition can proceed through the electron capture either from the conduction band or from the valence band. For this reason, the total probability equals the sum of probabilities for each of these processes:

$$a_{12} = C_n^{12} n + e_p^{12}.$$

Other coefficients a_{ij} are determined similarly:

$$a_{21} = C_p^{21} p + e_n^{21}, \quad a_{13} = C_n^{13} n + e_p^{13},$$

$$a_{31} = C_p^{31} p + e_n^{31},$$

$$a_{34} = C_n^{34} n + e_p^{34}, \quad a_{43} = C_p^{43} p + e_n^{43},$$

$$a_{24} = C_n^{24} n + e_p^{24}, \quad a_{42} = C_p^{42} p + e_n^{42}.$$

The coefficients $e_{n,p}^{ij}$ are calculated from the Gibbs distribution of electronic states:

$$e_n^{21} = C_n^{12} g_1 g_2^{-1} n_i^2 N_\nu^{-1} \exp[E_1/kT],$$

$$e_n^{31} = C_n^{13} g_1 g_3^{-1} n_i^2 N_\nu^{-1} \exp[(E_1 + E)/kT],$$

$$e_n^{42} = C_n^{24} g_2 g_4^{-1} n_i^2 N_\nu^{-1} \exp[(E_1 + E)/kT],$$

$$e_n^{43} = C_n^{34} g_3 g_4^{-1} n_i^2 N_\nu^{-1} \exp[E_1/kT],$$

$$e_p^{12} = C_p^{21} g_2 g_1^{-1} N_\nu \exp[-E_1/kT],$$

$$e_p^{13} = C_p^{31} g_3 g_1^{-1} N_\nu \exp[-(E_1 + E)/kT],$$

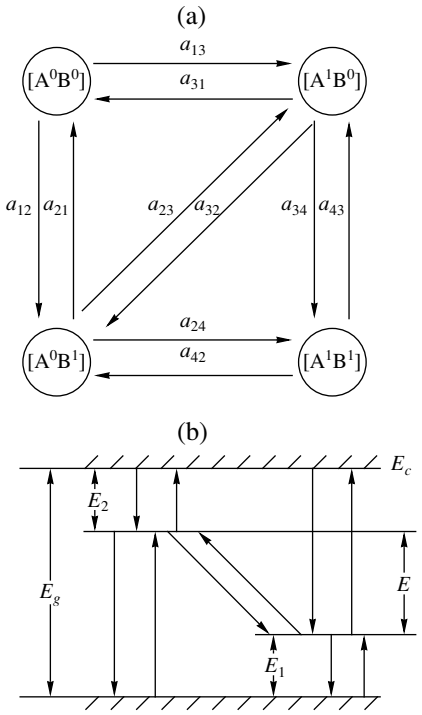


Fig. 1. (a) Schematic representation of mutual transformation of pair-defect states and (b) schematic diagram of energy levels.

$$e_p^{24} = C_n^{24} g_4 g_2^{-1} N_\nu \exp[-(E_1 + E)/kT],$$

$$e_p^{34} = C_n^{43} g_4 g_3^{-1} N_\nu \exp[-E_1/kT],$$

$$\nu_{32} = \nu_{23} \exp[-E/kT].$$

Here, N_c and N_ν are the densities of quantum states in the conduction band and the valence band, respectively, and g_1, g_2, g_3, and g_4 are the degeneracy factors. The schematic diagram of the energy levels is shown in Fig. 1b.

The set of equations, which determines the kinetics of variation in the defect concentration for each state, taking into account all transitions mentioned above and conservation of the total amount of defects, is written as

$$\dot{N}_1 = -(a_{12} + a_{13})N_1 + a_{21}N_2 + a_{31}N_3, \tag{1}$$

$$\dot{N}_2 = -(\alpha_{21} + a_{23} + a_{24})N_2 + a_{12}N_1 + a_{32}N_3 + a_{42}N_4, \tag{2}$$

$$\dot{N}_3 = -(a_{31} + a_{32} + a_{34})N_3 + a_{13}N_1 + a_{23}N_2 + a_{43}N_4, \tag{3}$$

$$N_1 + N_2 + N_3 + N_4 = N_R, \tag{4}$$

where N_R is the total defect concentration.

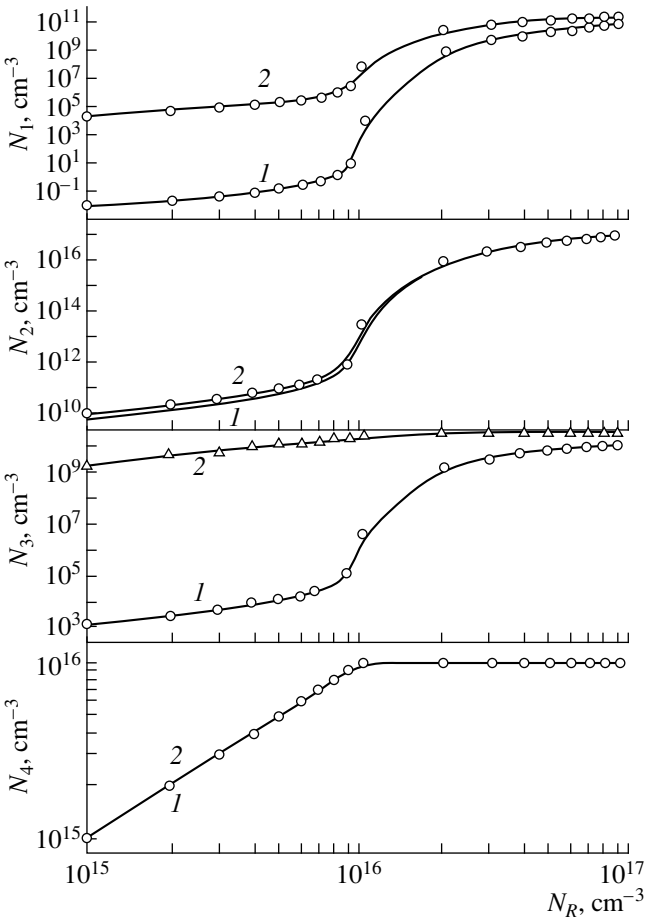


Fig. 2. Dependences of (*1*) equilibrium and (*2*) nonequilibrium concentrations of various types of pair defects on their total concentration N_R at 300 K, $N_d = 10^{16}$ cm^{-3}, $\Delta n = 10^{10}$ cm^{-3}, and $\nu_{32} = 10^7$ s^{-1}.

Expressions for the recombination rate (U) and the electron (τ_n) and hole (τ_p) lifetimes have the form

$$U = N_R(C_n^{12} + C_n^{13})n f_1 + (C_n^{24} n - e_n^{21}) f_2 + (C_n^{34} n - e_n^{31}) f_3 - (e_n^{43} + e_n^{42}) f_4, \quad (5)$$

$$\tau_n = \Delta n/U, \quad \tau_p = \Delta p/U, \quad (6)$$

where $n = n_0 + \Delta n$ and $p = p_0 + \Delta p$ are the total densities, Δn and Δp are the excess densities, and n_0 and p_0 are the equilibrium densities of free electrons and holes.

The relation between the concentrations of free electrons, holes, shallow-level donors (N_d), and positively charged (N_1) and negatively charged (N_4) pair defects is given by the equation of the electroneutrality:

$$p + N_d + N_1 = n + N_4. \quad (7)$$

The estimation of resistivity (ρ), the Hall coefficient (R_H), and Hall mobility (μ_H) was carried out with the formulas

$$\rho = \frac{1}{q(n\mu_n + p\mu_p)}, \quad R_H = \frac{r(p\mu_p^2 - n\mu_n^2)}{n\mu_n + p\mu_p}, \quad \mu_H = \frac{R_H}{\rho}, \quad (8)$$

where q is the elementary charge and μ_n and μ_p are the electron and hole mobilities, respectively.

Although the system (1)–(4) can be solved analytically, the expressions obtained for N_1, N_2, N_3, and N_4 are very cumbersome and we see no reason to present them. The numerical solution for the material of concern is given below.

3. CHOICE OF THE MATERIAL AND PARAMETERS FOR NUMERICAL CALCULATION

We choose silicon as the object for investigation. Silicon is one of the materials most often used in modern electronics and contains pair deep-level centers. Let us suppose that the upper energy level of this center is located at $E_2 = 0.5$ eV below the conduction-band bottom (Fig. 1b). The lower energy level is thereby located at $E_1 = 0.2$ eV above the valence-band top. The band gap for silicon equals $E_g = 1.1$ eV at room temperature. From here, it is easy to find that the E_1 and E_2 levels are separated by $E = E_g - E_1 - E_2 = 0.4$ eV. The following values of the trapping coefficients are used (cm^3 s^{-1}): $C_n^{12} = 10^{-12}$, $C_n^{13} = 10^{-10}$, $C_n^{24} = 10^{-10}$, $C_n^{34} = 10^{-12}$, $C_p^{21} = 10^{-9}$, $C_p^{31} = 10^{-12}$, $C_p^{42} = 10^{-12}$, and $C_p^{43} = 10^{-9}$. The intracenter transition coefficient ν_{32} was varied from 10^4 to 10^{10} s^{-1}; $N_c = 2.8 \times 10^{19}$ cm^{-3}, and $N_v = 1.04 \times 10^{19}$ cm^{-3}.

Let us carry out the estimation for room temperature $T = 300$ K in conditions of complete ionization of shallow donors with the concentration $N_d = 10^{16}$ cm^{-3}.

4. RESULTS AND ANALYSIS

The set of equations (1)–(4) was solved for equilibrium conditions taking into account the electroneutrality equation (7). We investigated the dependence of the equilibrium concentration of the pair defect for each state N_1, N_2, N_3, and N_4 on the total concentration of pair centers N_R for a fixed concentration of shallow-level donors. Similar dependences of the electron (n_0) and hole (p_0) densities, resistivity ρ, Hall coefficient R_H, and Hall mobility μ_H were also investigated. The results are shown in Figs. 2–4. It can be seen from Fig. 2 that, for $N_R < N_d$, most pair defects, both of whose levels are occupied by electrons, are in the N_4 state. If $N_R > N_d$,

most defects are in the N_2 state, with the lower level occupied by one electron and the upper level unoccupied. It is noteworthy that the N_2 and N_4 concentrations are virtually independent of the concentration of excess charge carriers, while N_1 and N_3 substantially depend on it (Fig. 2). Introducing the excess charge carriers in a semiconductor in even a minute amount comparable with the intrinsic concentration n_i brings about a variation of N_1 and N_3 by several orders of magnitude.

The N_R-dependences of the equilibrium electron (n_0) and hole (p_0) concentration are conditioned by the above dependences of the pair defect fractions on their total concentration. The N_R-dependence of the sum $n_0 + p_0$ is shown in Fig. 3a. For $N_R < N_d$, the equilibrium concentration of the majority carriers is defined by the shallow-level donor concentration $n_0 \approx N_d \gg p_0$. The equilibrium concentration slowly decreases with increasing N_R. For $N_R > N_d$, the equilibrium concentration n_0 is defined by the density of electrons thermally generated from the upper level of the deep-level impurity. Since it was assumed in calculations that $E_2 = 0.5$ eV, $n_0 \approx 10^{11}$ cm^{-3}. This is the case for $N_R > N_d$. For $N_R = N_d$, i.e., for a complete compensation, the equilibrium electron density abruptly drops from $n_0 \approx N_d$ to $n_0 \approx 10^{11}$ cm^{-3}. The latter corresponds to electrons thermally generated from the impurity level. For $N_R > N_d$, n_0 slowly decreases with increasing N_R and the location of other levels and the probability of electron exchange between them therewith do not affect n_0.

A similar dependence (Fig. 3a) should lead to corresponding abrupt changes in all electrical parameters of the semiconductor, which are conditioned by the charge-carrier concentration. These are the recombination rate, electron and hole lifetimes, resistivity, the Hall coefficient, Hall mobility, Dember photovoltage, etc.

The N_R-dependence of the recombination rate U is shown in Fig. 3b. It can be seen that U increases with increasing N_R for $N_R < N_d$ and levels off for $N_R \geq N_d$. Analyzing Fig. 2, it is easy to understand the character of the $U(N_R)$ dependence. An increase in U with increasing N_R for $N_R < N_d$ is related to the fact that the trapping rate of free charge carriers and the intensity of recombination increase with an increase in the deep-level impurity concentration. For $N_R > N_d$, as shown in Fig. 2, $N_2 \gg (N_1, N_3,$ and $N_4)$. Consequently, we can simplify the expression for the recombination rate and derive the asymptotic formula

$$U \approx C_n^{24} n N_2. \quad (9)$$

Since N_2 slowly increases (Fig. 2b) as N_R rises, while the free electron concentration slowly decreases (Fig. 3a), the recombination rate remains constant.

Investigations of the $U(N_R)$ dependence for various probabilities of the intracenter transition ν_{32} demonstrated that the curves $U(N_R)$ shift upwards parallel to themselves with an increase in ν_{32} and rise rapidly for small ν_{32}. For large values of ν_{32}, $U(N_R)$ slowly increases and then levels off. Consequently, ν_{32} introduces only quantitative rather than the radical changes in the N_R-dependences of U.

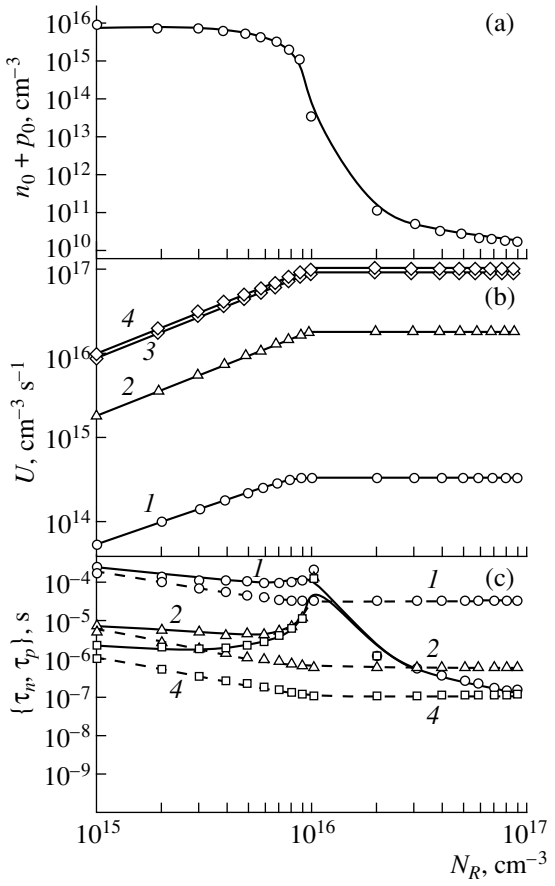


Fig. 3. Dependences of (a) the sum of equilibrium densities $n_0 + p_0$, (b) recombination rate U, and (c) electron τ_n (solid line) and hole τ_p (dashed line) lifetimes on the pair defect concentration N_R at 300 K for $N_d = 10^{16}$ cm^{-3}, $\Delta n = 10^{10}$ cm^{-3}, and the intracenter transition probability $\nu_{32} =$ (*1*) 10^4, (*2*) 10^6, (*3*) 10^8, and (*4*) 10^{10} s^{-1}.

The electron and hole lifetimes for a fixed concentration of excess holes 10^{10} cm^{-3} were also investigated. This condition can be realized by exposing the crystal to light in the impurity absorption region. It is clearly seen that the electron lifetime decreases with increasing N_R (Fig. 3c) for $N_R < N_d$. The reason is that the recombination rate increases (Fig. 3b), while the equilibrium concentration decreases only slightly (Fig. 3a). For $N_R = N_d$, τ_n sharply increases by several orders of magnitude. This is caused by an abrupt decrease in n_0 and, correspondingly, in the number of recombining electrons. The maximum of τ_n is defined by the number of electrons, which are thermally excited from the deep

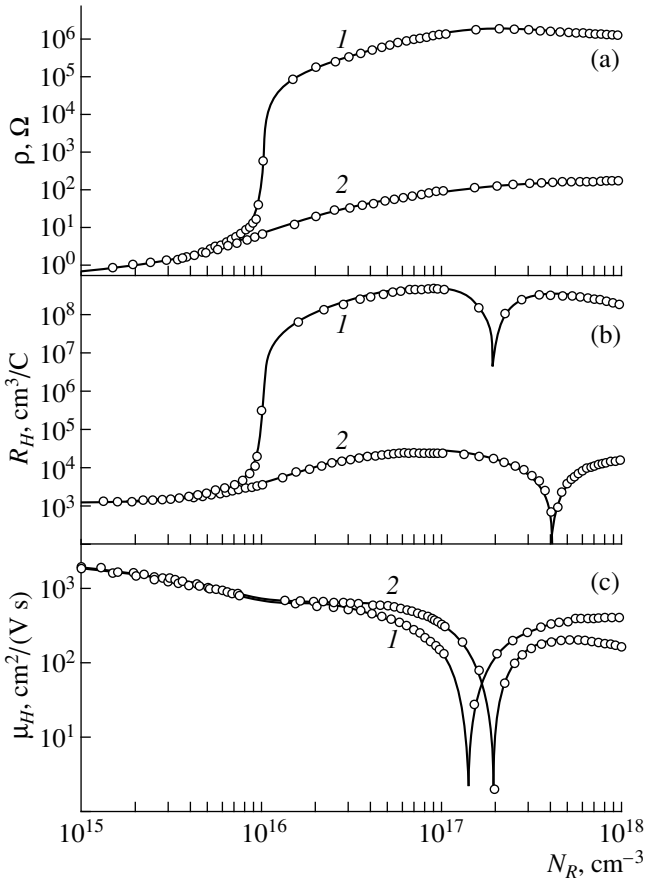


Fig. 4. Dependences of (a) resistivity ρ, (b) Hall coefficient R_H, and (c) Hall mobility μ_H on the pair defect concentration N_R at 300 K. $N_d = 10^{16}$ cm^{-3}, and Δn = (*1*) 0 and (*2*) 10^{10} cm^{-3}.

impurity level to the conduction band, and decreases as the concentration of excess charge carriers increases. Consequently, τ_n is governed by temperature, the illumination intensity in the band-to-band and impurity absorption regions, pressure, and the deep-level impurity type. The maximum in the $\tau_n(N_R)$ dependence disappears completely at very high temperatures and injection levels rendering the carrier concentration comparable with the shallow-level donor concentration. For $N_R > N_d$, τ_n slowly decreases with an increase in N_R. This is caused by a slow decrease in n_0 with U being constant.

The N_R-dependence of the hole lifetime replicates the shape of the N_R-dependence of the $1/U$ quantity (Fig. 3c), since, in the above calculation of all N_R values, the excess hole concentration was assumed to be constant. A large difference between τ_n and τ_p (Fig. 3c) is related to the difference between the excess densities of electrons Δn and holes Δp, which is easily verifiable by dividing the first equation (6) by the second one:

$$\frac{\tau_n}{\tau_p} = \frac{\Delta n}{\Delta p}. \qquad (10)$$

Investigation of the τ_n and τ_p lifetime dependence on the probability of intracenter electronic transition ν_{32} demonstrates (Fig. 3c) that ν_{32} introduces no drastic changes in the $\tau_n(N_R)$ and $\tau_p(N_R)$ dependences.

An abrupt change in the N_R-dependence of $n_0 + p_0$ (Fig. 2a) in the case of complete compensation defines the corresponding abrupt change in other parameters of the semiconductor characterizing its electrical properties. These parameters are the resistivity ρ (Fig. 4a), Hall coefficient R_H (Fig. 4b), and Hall mobility μ_H (Fig. 4c) estimated from formulas (8). The maximal values of resistivity and Hall coefficient decrease with an increase in temperature and concentration of injected charge carriers. The dependences of the electron and hole mobilities on the charged impurity concentration introduce no radical changes in the shape of the $\rho(N_R)$ and $R_H(N_R)$ dependences. In addition, the $R_H(N_R)$ and $\mu_H(N_R)$ dependences are nonmonotonic and pass a minimum at $\mu_n n^2 = \mu_p p^2$. This is caused by the variations in n_0 and p_0 with an increase in N_R.

5. CONCLUSION

Thus, a sudden change in the sum of equilibrium concentrations, resistivity, the Hall coefficient, and charge carrier lifetimes are the related effects, which are typical of precisely compensated semiconductors. These effects are clearly pronounced for semiconductors with a shallow-level donor (acceptor) concentration several orders of magnitude larger than the concentration of electrons (holes) excited to the conduction (valence) band from the nearby deep impurity level. An additional specific feature of a precise compensation of the semiconductor and, correspondingly, a manifestation of the aforementioned effects is the enhanced sensitivity of the semiconductor to variation of external factors. These are the intensity of weak radiation in the band-to-band and impurity absorption regions, temperature, etc.

We note that the effect of a sharp rise of the resistivity was previously observed experimentally for many semiconductors, namely, polycrystalline silicon [8], single-crystal silicon [9, 10], cadmium telluride [11], etc.

It was reported that two maxima are possible in the dependences $\tau_n(N_R)$ and $\tau_p(N_R)$ [7]. It is our opinion that this is possible for a semiconductor with two types of mutually independent deep-level recombination centers with different concentrations N_{R1} and N_{R2}. Two sharp peaks τ_n at $N_d = N_{R1}$ and N_{R2} are possible only in this case. Since we consider only one type of deep-level centers in this study, the precise compensation is possible only for a unique concentration of deep-level cen-

ters at a given shallow-level donor concentration. According to this, a single maximum is observed in the $\tau_n(N_R)$ dependence.

ACKNOWLEDGMENTS

This work was supported by the UNDR Fellowship and the Academy of Sciences of Uzbekistan.

This work was supported by NATO Linkage Grants, contract no. PST SLG 975758.

REFERENCES

1. A. A. Bergh and P. J. Dean, *Light Emitting Diodes* (Clarendon, Oxford, 1976; Mir, Moscow, 1973).
2. L. I. Berezhinskiĭ, F. K. Dzhaparova, N. V. Kitsyuk, and V. E. Rodionov, Ukr. Fiz. Zh. **36**, 513 (1992).
3. V. V. Evstropov, I. Yu. Linkov, I. V. Morozenko, and F. P. Pikus, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 969 (1992) [Sov. Phys. Semicond. **26**, 543 (1992)].
4. V. A. Kholodnov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30** (6), 1011 (1996) [Semiconductors **30**, 538 (1996)].
5. A. A. Drugova and V. A. Kholodnov, Solid-State Electron. **38** (6), 1247 (1995).
6. V. A. Kholodnov, Pis'ma Zh. Éksp. Teor. Fiz. **67** (9), 655 (1998) [JETP Lett. **67**, 685 (1998)].
7. V. A. Kholodnov and P. S. Serebrennikov, Pis'ma Zh. Tekh. Fiz. **23** (7), 39 (1997) [Tech. Phys. Lett. **23**, 268 (1997)].
8. J. Y. W. Seto, J. Appl. Phys. **46** (12), 5247 (1975).
9. A. G. Milnes, *Deep Impurities in Semiconductors* (Wiley, New York, 1973; Mir, Moscow, 1977).
10. A. N. Marchenko, *Controlled Semiconductor Resistors* (Énergiya, Moscow, 1978).
11. Sh. A. Mirsagatov, V. M. Rubinov, and F. N. Dzhamalov, Geliotekhnika, No. 2, 12 (1997).

Translated by N. Korovin

Semiconductors, Vol. 34, No. 8, 2000, pp. 886–888. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 923–925.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Transition of a Noncavity Bistable Exciton System Driven by Correlated External Noise to a Strongly Absorbing State

Yu. V. Gudyma

Fed'kovich State University, ul. Kotsyubinskogo 2, Chernovtsy, 274012 Ukraine

e-mail: general@chdu.cv.ua

Submitted January 10, 2000; accepted for publication February 17, 2000

Abstract—The correlated action of additive and multiplicative noise interferences on an exciton bistable system results in the disappearance of the weakly absorbing state, with the system's behavior turning from reversible to irreversible. The strongly absorbing state of the bistable system shifts to the region of higher exciton concentrations. © *2000 MAIK "Nauka/Interperiodica".*

Interest in studying bistable systems driven by correlated additive and multiplicative noise interferences has grown rapidly in recent years [1, 2]. These investigations are based on the pioneering work [3], in which adequate theoretical tools were developed and the most important relationships for correlated noise were obtained. However, all these calculations deal with simplified model systems, although it is clear that a combined action of environment fluctuations is not a rare occurrence in specific physical problems. We consider such a situation for the example of transitions in a noncavity bistable exciton system. The nonlinear mechanism of the noncavity bistability associated with an exciton transition is related to the screening of excitonic states by a high-density electron–hole plasma or to band-gap renormalization [4, 5]. The noncavity optical bistability associated with an exciton transition in a semiconductor was studied in detail experimentally and theoretically [6]. This problem has drawn interest, because it is a prominent example of a phase transition in systems far from thermodynamic equilibrium and creates ample opportunities for application in ultrafast switching.

In recent years, laser physics has rapidly mastered the range of ultrashort pulses generated by continuous-wave solid-state lasers, which have produced a number of new results, including those in the field of the noncavity exciton bistability [7]. The finite laser pulse repetition rate causes the energy density to fluctuate from one burst to another. Since the light intensity is a multiplicative parameter in the Beer–Lambert law, one would expect that these fluctuations can have a strong impact on the system, because its deterministic nonequilibrium properties are known to be fully controllable by varying the energy flux of incident light. On the other hand, the significantly nonequilibrium character of the process in question results in random fluctuations of exciton velocity and concentration, appearing as additive noise in the kinetic equation for the exciton system.

With the above-mentioned fluctuations neglected, the transport equations for radiation of intensity $I(r)$ and a free quasiparticle density $n(r, t)$ take the form

$$\frac{dI}{dr} = -\alpha(\omega, n)I(r), \tag{1}$$

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial r^2} + \alpha(\omega, n)I(r) - \frac{n}{\tau}, \tag{2}$$

where $\alpha(\omega, n)$ and D are the light absorption and diffusion coefficients, respectively, and τ is the exciton lifetime.

The frequency dependence of the absorption coefficient $\alpha(\omega, n)$ in the case of resonance light absorption is given by

$$\alpha(\omega, n) = \alpha_0\left[1 + \left(\frac{\omega_r' - \omega}{\Delta\omega}\right)^2\right]^{-1}. \tag{3}$$

For high excitation levels, the cooperative interaction phenomena in the system lead the resonance frequency ω_r' to become a function of the excitation concentration n and to decrease with increasing n, which itself depends on the level of optical excitation I: $\omega_r'(n) = \omega_r - an$. In the case of excitons, $\Delta\omega$ and an are the spectral width and shift (in an approximation linear in particle concentration) of the exciton level ω_r.

Let us assume that the surface ($r = 0$) of a semiconductor wafer of thickness l is uniformly illuminated with light of intensity I_0. Assuming that the diffusion length exceeds the wafer thickness and introducing a

1063-7826/00/3408-0886$20.00

thickness-averaged light intensity, we obtain a nonlinear equation of concentration balance

$$\frac{dn}{dt} = I_0 l^{-1}\{1 - \exp[-\alpha(\omega, n)l]\} - n\tau^{-1}. \quad (4)$$

Fixing the frequency of incident light, we can pass in (4) to dimensionless variables. With account of the above-mentioned noise, the dimensionless kinetic equation takes the form

$$\frac{d\eta}{d\theta} = \beta(t)\{1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} - \eta + \gamma(t), \quad (5)$$

where

$$\beta(t) = \beta + \sigma\xi(t), \quad \theta = t/\tau, \quad \lambda = \alpha_0 l,$$

$$\Omega = ((\omega_r - \omega)/\Delta\omega)^2, \quad \eta = an/(\omega_r - \omega),$$

$$\beta = a\tau I_0/(\omega_r - \omega)l.$$

In equation (5), $\gamma(t)$ and $\xi(t)$ are Gaussian white noises with zero means and

$$\langle \xi(t)\xi(t')\rangle = 2\sigma^2\delta(t - t'), \quad (6a)$$

$$\langle \gamma(t)\gamma(t')\rangle = 2\varepsilon^2\delta(t - t'), \quad (6b)$$

$$\langle \xi(t)\gamma(t')\rangle = 2\chi\sigma\varepsilon\delta(t - t'), \quad (6c)$$

where χ denotes the degree of correlation between the multiplicative and additive noises. In terms of generalized functions, the Gaussian white noise is a derivative of the Wiener process. Therefore, equation (5) has the meaning of the Stratonovich stochastic differential equation, which can be associated with the Fokker–Planck equation that defines the evolution of the probability of the system's transition from one state to another [3, 8]:

$$\frac{\partial P(\eta, \theta|\eta')}{\partial\theta} = -\frac{\partial}{\partial\eta}[A(\eta, \theta)P(\eta, \theta|\eta')] + \frac{1}{2}\frac{\partial^2}{\partial\eta^2}[B(\eta, \theta)P(\eta, \theta|\eta')], \quad (7)$$

where

$$A(\eta) = \beta\{1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} - \eta + 2\Omega\lambda\beta\sigma^2(1-\eta)[1 + \Omega(1-\eta)^2]^{-1} \times \exp[-\lambda/(1 + \Omega(1-\eta)^2)] \times \{1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} + 2\Omega\lambda\beta\chi\sigma\varepsilon(1-\eta)((1 + \Omega(1-\eta)^2)^{-1} \times \exp[-\lambda/(1 + \Omega(1-\eta)^2)], \quad (8)$$

$$B(\eta) = \beta^2\sigma^2\{1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} + 2\beta\chi\sigma\varepsilon\{1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} + \varepsilon^2. \quad (9)$$

If microscopic transients of a nonlinear system subjected to an external noise are of no interest, then the states of the system are adequately described by the probability density $P(\eta)$. The time-independent solution of the homogeneous Fokker–Planck equation (7)

$$P_s(\eta) = N\{\beta^2\sigma^2(1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)])^2 + 2\beta\chi\sigma\varepsilon(1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]) + \varepsilon^2\}^{-1/2} \times \exp\left\{\int_0^\eta \{\beta(1 - \exp[-\lambda/(1 + \Omega(1-\eta')^2)]) - \eta'\} \times \{\sigma^2\{1 - \exp[-\lambda/(1 + \Omega(1-\eta')^2)]\}^2 + 2\chi\sigma\varepsilon \times (1 - \exp[-\lambda/(1 + \Omega(1-\eta')^2)]) + \varepsilon^2\}^{-1} d\eta'\right\} \quad (10)$$

corresponds to the diffusion process. Here, N is a constant to be determined from normalization conditions. These are imposed by the physical conditions of a problem, for example, by the variation of the concentration of excitons from zero to that at which their Bose condensation occurs.

The peaks and minimums of the probability density correspond to the stable and unstable steady states, respectively. Thus, the extrema of the stationary probability density can be identified with the macroscopic steady states of the system [8]. Testing equation (10) for extrema yields the following transcendental equation whose roots correspond to the possible steady states of the exciton system:

$$\beta\{1 - \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} \times \{1 - 2\lambda\Omega\sigma^2(1-\eta)(1 + \Omega(1-\eta)^2)^{-1} \times \exp[-\lambda/(1 + \Omega(1-\eta)^2)]\} - 2\beta\lambda\Omega\chi\sigma\varepsilon(1-\eta)(1 + \Omega(1-\eta)^2)^{-1} \times \exp[-\lambda/(1 + \Omega(1-\eta)^2)] = 0. \quad (11)$$

A numerical analysis of equation (11) was performed for the physical parameters corresponding to CdS-like semiconductors illuminated with a laser having light flux intensity at the level of 10^{19}–10^{20} photon/(cm^2 s). In this case, $\Omega = 100$ and $\lambda = 0.055$.

As is known, for the values of the control (cooperative) parameter smaller than a critical value, the build-up of multiplicative noise may initiate an optical bistability in the system [9] (see Fig. 1). An increase in the intensity of the additive noise correlated with the multiplicative noise gives rise to a number of new effects (see Fig. 2). First, the strongly absorbing state of the bistable system shifts to the region of higher exciton concentrations (the effect of steady-state suppression by the additive noise). Second, the weakly absorbing

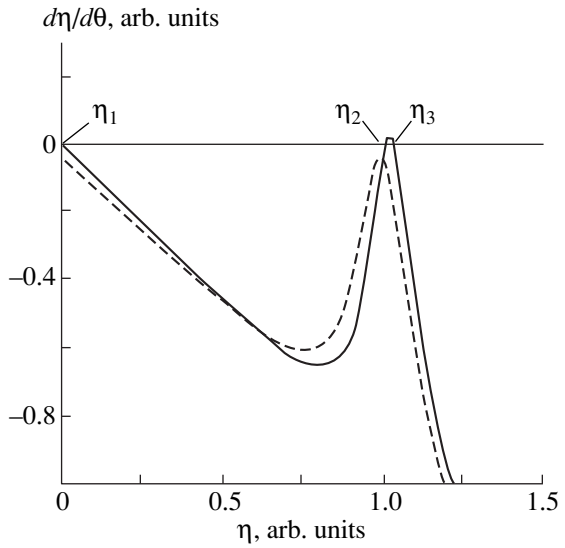


Fig. 1. Initiation of bistability in an exciton system subjected to external multiplicative noises.

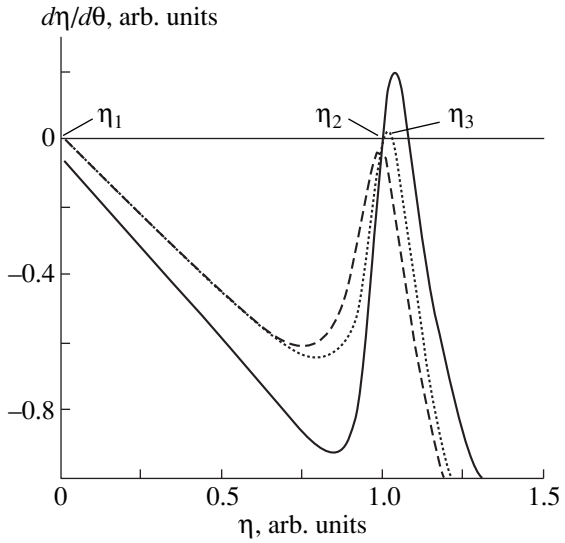


Fig. 2. Transition to the irreversible phenomena under the effect of correlated noises.

state of the system disappears under the effect of the noise. Thereby, the physical processes in the system change from reversible to irreversible. We note that the effect of the noise is significant in the region of physical parameters where the input signal detuning exceeds the spectral width of the exciton level.

Thus, the theory of the effect of the correlated external additive and multiplicative noises not only provides new insight into the origin of the bistability associated with exciton transitions in semiconductors but also demonstrates the occurrence of irreversible phenomena in systems of this kind.

REFERENCES

1. D. Mei, G. Xie, Li Cao, and D. Wu, Phys. Rev. E **59**, 3880 (1999).
2. H. Fu, Li Cao, and D. Wu, Phys. Rev. E **59**, R6235 (1999).
3. D. Wu, Li Cao, and Sh. Ke, Phys. Rev. E **50**, 2496 (1994).
4. K. Bohnert, H. Kalt, and C. Klingshirn, Appl. Phys. Lett. **43**, 1068 (1983).
5. H. Rossmann and F. Henneberger, Phys. Status Solidi B **139**, 185 (1985).
6. P. I. Khadzhi, G. D. Shibarshina, and A. Kh. Rotaru, *Optical Bistability in a System of Coherent Excitons and Biexcitons in Semiconductors* (Shtiintsa, Kishinev, 1988).
7. P. I. Khadzhi, A. M. Rusanov, and S. L. Gaĭvan, Kvantovaya Élektron. (Moscow) **27**, 262 (1999).
8. W. Horsthemke and R. Lefever, *Noise Induced Theory and Applications in Physics, Chemistry and Biology* (Springer, Heidelberg, 1984; Mir, Moscow, 1987).
9. Yu. V. Gudyma, Izv. Vyssh. Uchebn. Zaved. Fiz., No. 2, 94 (1998).

Translated by S. Goupalov

Semiconductors, Vol. 34, No. 8, 2000, pp. 889–890. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 926–928.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Variable-Range-Hopping Conduction via Indium Impurity States in $Pb_{0.78}Sn_{0.22}Te$ Solid Solution

S. A. Nemov*, V. É. Gasumyants*, V. I. Proshin*, Yu. V. Ravich*, and D. A. Potapova**

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

*** Herzen Russian State Pedagogical University, St. Petersburg, 191186 Russia*

Submitted February 8, 2000; accepted for publication February 9, 2000

Abstract—The Seebeck coefficient S was measured in a wide temperature range $T < 100$ K in $Pb_{0.78}Sn_{0.22}Te$ solid solutions doped with 3 at. % of In, with additional Cl doping of up to 3 at. %. The temperature derivative $\partial|S|/\partial T$ changes its sign from negative to positive below 100 K. Theoretical estimations in terms of hopping conduction via highly localized indium-related states show that the transition to variable-range-hopping conduction must occur at temperatures of about 50–100 K, in agreement with the obtained experimental data. © *2000 MAIK "Nauka/Interperiodica".*

Indium impurity in lead telluride and $Pb_{1-x}Sn_xTe$ solid solutions gives rise to highly localized (within ~10 Å) deep resonance electron states [1]. The solubility of indium in $Pb_{1-x}Sn_xTe$ is unusually high for a heterovalent impurity (>20 at. %); thus, it becomes possible to observe and study the hopping conduction via strongly localized states [2–5].

With increasing Sn content in the $Pb_{1-x}Sn_xTe$ solid solution, In-related impurity levels shift from inside the conduction band to its bottom, and, at $x \approx 0.2$, pass into the forbidden band, transforming from quasi-local into local states. For $x = 0.22$, these levels lie near the boundary between the forbidden and conduction bands, thus favoring the observation of transport phenomena associated with hopping conduction.

At temperatures T in the range of 100–400 K, the conduction due to electron hopping between two neighboring impurity atoms is thermally activated with activation energy $\varepsilon_h \approx 35$ meV for In content $N_{In} = 3$ at. % [5]. The dependence of the hopping conductivity on In concentration is also exponential, and its analysis yields the effective radius of the wave function, $a \approx 6$ Å [2]. The obtained parameters enable evaluation of the extent to which the condition that hopping predominantly occurs between nearest neighbors is fulfilled. This condition has the form [6]

$$2/N_{In}^{1/3}a > \varepsilon_h/kT. \quad (1)$$

In particular, for $N_{In} = 3$ at. % and $T = 200$ K, the left-hand side of inequality (1) equals 4.3 and the right-hand side equals 2.1. Thus, at temperatures above 100 K, conduction is due to hopping between nearest neighbor atoms; that is, the average hopping range is temperature-independent. However, with the temperature lowered below 100 K, the inequality (1) is reversed and the variable-range-hopping conduction with weaker temperature dependence described by the Mott law $\sigma \propto \exp[-(T_0/T)^{1/4}]$ must be observed [6].

Our experiments [5] have shown a tendency for the temperature dependence of conduction to become weaker below 100 K. However, the obtained data vary widely from sample to sample, presumably due to the presence of grain boundaries. No conduction following the Mott law could be observed. It would be expected that the zero-current Seebeck effect is less influenced by the sample quality than the electrical conduction, and, therefore, measurements of the thermoelectric power could reveal variable-range-hopping conduction in these materials.

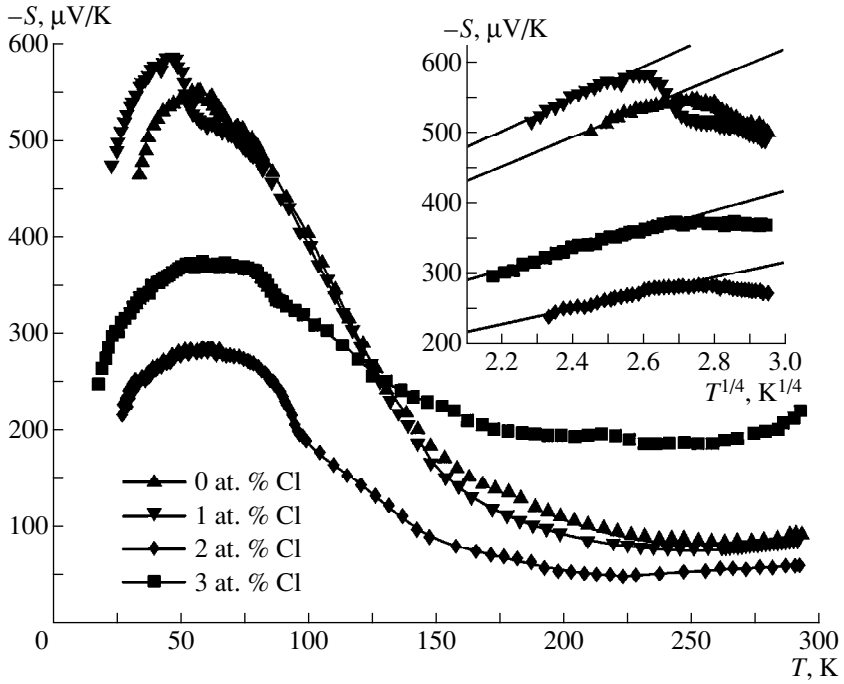
Experimental studies of the thermoelectric power in the temperature range of 100–400 K [3] yielded the temperature dependence of the Seebeck coefficient

$$S = A/T + BT, \quad (2)$$

which is typical of the Fermi glasses with conduction governed by the nearest neighbor hopping [7]. In particular, the value of $|S|$ decreases with increasing temperature at relatively low temperatures (100–200 K for $(Pb_{0.78}Sn_{0.22})_{1-y}In_yTe$), when the first term in (2) is dominant. The transition to the variable-range-hopping conduction must fundamentally change the temperature dependence of the thermoelectric power: according to theory [8], it increases with temperature when the average hopping range is temperature-dependent.

The figure presents the results of thermoelectric power measurements in $(Pb_{0.78}Sn_{0.22})_{0.97}In_{0.03}Te$ with additional chlorine doping (up to 3 at. %) in a wide temperature range including $T < 100$ K. As can be seen, at $T < 50$ K, the Seebeck coefficient increases with temperature, which is characteristic of the variable-range-hopping conduction. The temperature at which the derivative $\partial|S|/\partial T$ reverses its sign is of the same order of magnitude as the above theoretical estimate.

1063-7826/00/3408-0889$20.00 © 2000 MAIK "Nauka/Interperiodica"



Temperature dependence of the Seebeck coefficient S for $(Pb_{0.78}Sn_{0.22})_{0.97}In_{0.03}Te$ doped additionally with chlorine.

Additional doping with chlorine changes the chemical potential, thereby affecting the density of states g at this level and its derivative $dg/d\varepsilon$ with respect to energy at the chemical potential level; this derivative controls the thermoelectric power [7]. An analysis of the results of thermoelectric power measurements at higher temperatures (100–400 K) for additionally doped samples [4] shows that the density of states at the chemical-potential level changes steeply with energy, with nonmonotonic variation of the derivative $dg/d\varepsilon$. In the case where a nearly constant density of states near the chemical-potential level can be assumed, the theory [8] yields the law $S \propto T^{1/2}$ for the temperature dependence of the thermoelectric power. Since, for the materials under study, the assumption of constant density of states is invalid, we cannot expect quantitative agreement of the experimental data with this simple model. Indeed, as can be seen from the insert in the figure, the experimental curve $S(T)$ is closer to $S \propto T^{1/4}$ than to $S \propto T^{1/2}$.

Nevertheless, the general nonmonotonic run of the temperature dependence of the thermoelectric power, including the low-temperature region $T < 100$ K, confirms the assumption that variable-range-hopping conduction via highly localized indium-related impurity states occurs in the materials under study. At $T \approx 100$ K and higher, the conduction is governed by hopping between the nearest neighbors, whereas, at $T \approx 50$ K and below, variable-range-hopping conduction is observed, in accordance with the theoretical estimates.

REFERENCES

1. V. I. Kaĭdanov and Yu. I. Ravich, Usp. Fiz. Nauk **145**, 51 (1985) [Sov. Phys. Usp. **28**, 31 (1985)].
2. Yu. I. Ravich, S. A. Nemov, and V. I. Proshin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1448 (1995) [Semiconductors **29**, 754 (1995)].
3. S. A. Nemov, V. I. Proshin, and Yu. I. Ravich, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 2164 (1996) [Semiconductors **30**, 1128 (1996)].
4. T. G. Abaĭdulina, S. A. Nemov, V. I. Proshin, and Yu. I. Ravich, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 2173 (1996) [Semiconductors **30**, 1133 (1996)].
5. S. A. Nemov, Yu. I. Ravich, V. I. Proshin, and T. G. Abaĭdulina, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 311 (1998) [Semiconductors **32**, 280 (1998)].
6. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984).
7. N. T. E. Whall, J. Phys. C **14**, L887 (1981).
8. I. P. Zvyagin, in *Hopping Transport in Solids*, Ed. by B. Shklovskii and M. Pollak (Elsevier, Amsterdam, 1991), p. 143.

Translated by D. Mashovets

Semiconductors, Vol. 34, No. 8, 2000, pp. 891–893. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 929–931.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Spectral Shift of Photoluminescence Bands of the $(SiC)_{1-x}(AlN)_x$ Epitaxial Films due to Laser Annealing

G. K. Safaraliev, Yu. N. Émirov, M. K. Kurbanov, and B. A. Bilalov

Dagestan State University, Makhachkala, 367025 Russia

Submitted October 4, 1999; accepted for publication February 14, 2000

Abstract—The effect of laser annealing on the photoluminescence properties of $(SiC)_{1-x}(AlN)_x$ epitaxial films was studied. It was proposed that annealing causes the displacement of the Al and N atoms from their lattice sites and the formation of Al_{Si}–N_C donor–acceptor pairs acting as the luminescence centers. According to this model, the increase in the annealing time is accompanied by the formation of donor–acceptor pairs with the shortest interatomic distances at the expense of associations of the distant defects and by a shift of the respective photoluminescence band to the high-energy spectral region. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The excitation of semiconductor materials with radiation of the subthreshold power is one of the conventional means to reveal and study their defect structure [1, 2]. In this paper, the results of photoluminescence (PL) studies are reported for $(SiC)_{1-x}(AlN)_x$ epitaxial films (x = 0.03–0.17) grown on SiC substrates and subjected to laser annealing. The structure of both substrates and epitaxial films corresponds to the 6*H* polytype.

2. EXPERIMENTAL TECHNIQUE

The PL of $(SiC)_{1-x}(AlN)_x$/SiC structures (x = 0.03–0.17) was excited from the film side with radiation from a low-power LG-21 nitrogen laser having a wavelength λ = 337 nm, an average power $P_{av} \approx 3 \times 10^{-3}$ W, and a peak power $P_p \approx 1600$ W. The PL was detected at 77 K with an SDL-2 spectral setup. The $(SiC)_{1-x}(AlN)_x$ epitaxial films were annealed in a "Kvarts-5" installation with a high-power repetitively pulsed nitrogen laser. The latter was characterized by the power density $(1–5) \times 10^7$ W/cm^2, λ = 337 nm, P_p = 15 kW, and a repetition rate f = 1 kHz. Laser annealing was performed in atmospheric air using line scanning of the surface of the epitaxial film with a focused light probe with a spot diameter of 50–60 μm and an average power density at the surface of $P = (2–5) \times 10^3$ W/cm^2. Scanning was performed with a speed of 0.2 cm/s and a step of 50 μm.

3. RESULTS AND DISCUSSION

The PL spectra of $(SiC)_{1-x}(AlN)_x$/SiC structures before annealing appear to be identical for all AlN contents studied (x = 0.03–0.17). The representative spectrum in the form of a band with a peak at the photon energy $h\nu_m$ = 1.9 eV is shown in Fig. 1 (curve *1*). Annealing of the structure causes a shift of the PL band to higher energies and the appearance of a structure in its contour. These features depend on the power density of the laser radiation and are indicative of the formation of new centers with an emission energy close to that of the main peak $h\nu_m$ (curves *2* and *3*). The band at $h\nu_m \approx$ 2.7 eV is observed at the final stage of annealing. A further increase in the power density of the annealing beam ($P > 5 \times 10^3$ W/cm^2) does not produce any modifications in the spectrum.

The PL band at $h\nu_m$ = 1.9 eV in SiC is known to be associated with a carbon vacancy (V_C) [3, 4]. It is believed that, in $(SiC)_{1-x}(AlN)_x$ epitaxial films with a small AlN content ($x < 0.17$) having the same band gap as in the pure SiC [5], the initial PL band at $h\nu_m$ = 1.9 eV is also due to the recombination of the free electron with a hole localized at V_C. Laser radiation with $h\nu_m$ = 3.68 eV is absorbed near the surface of the epitaxial films under study. The power density of radiation $P = (2–5) \times 10^3$ W/cm^2 used for annealing is the limiting value. A further increase in P up to the values $\geq 8 \times 10^3$ W/cm^2 leads to local melting and the evaporation of the material, which manifest themselves in the appearance of grooves at the film surface (see Fig. 2). The subthreshold annealing used ($P = (2–5) \times 10^3$ W/cm^2) causes severe local heating of the film and is accompanied by the displacement of the Al and N atoms from their lattice sites with the subsequent migration of these atoms into the bulk of the film [6]. Since the intensity of the PL band at $h\nu_m$ = 1.9 eV ascribed to V_C decreases as a result of annealing, we believe that it is these vacancies at which Al and N atoms are localized. It is evident that because of the large ionic radius of Al atoms their position in silicon vacancies V_{Si} proves to

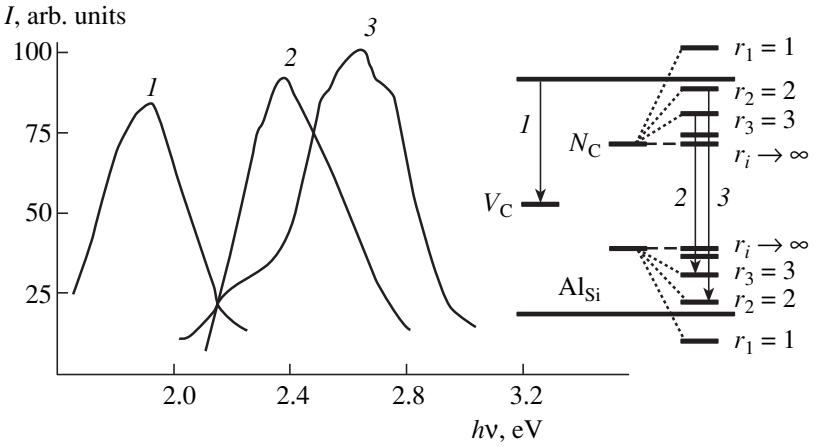


Fig. 1. PL spectra of $(SiC)_{0.93}(AlN)_{0.07}$ epitaxial films at 77 K (*1*) before and (*2* and *3*) after laser annealing with a power density of (*2*) 2.5×10^3 and (*3*) 5×10^3 W/cm^2. The energy level diagram of Al_{Si}–N_C donor–acceptor pairs is shown in the insert. The radiative transitions responsible for the PL bands *1*, *2*, and *3* are shown by the arrows *1*, *2*, and *3*; r_i are the relative interatomic distances in donor–acceptor pairs.

be more advantageous from the energy considerations than at the interstices or in V_C. It is assumed that the Al_{Si}–N_C donor–acceptor pairs formed in this case may give rise to PL bands at higher energies ($h\nu_m \approx$ 2.3–2.7 eV). The positions of these bands depend on the interatomic distance in such donor–acceptor pairs. Similar donor–acceptor pairs with such behavior of the PL band were revealed in SiC [7]. It would appear reasonable to believe that an increase in the laser annealing time would result in the efficient generation of donor–acceptor pairs with the shortest interatomic distances at the expense of pairs of distant defects, provided that the assumption on the relation of the PL centers to the donor–acceptor pairs discussed is valid. Such a redistribution should manifest itself, on the one hand, in a decrease in the intensity of low-energy PL bands associated with the pairs of distant defects and, on the other hand, in the formation of PL bands at higher energies, which are due to the pairs of closely spaced defects. Along with these effects, the appearance of the structure of the PL bands is expected. This structure is evidence of the participation of donor–acceptor pairs with wider ranges of the interatomic distances and hence emission energies in the formation of the integral PL band. Actually, the results obtained confirm our assumption (Fig. 1, curves *2* and *3*).

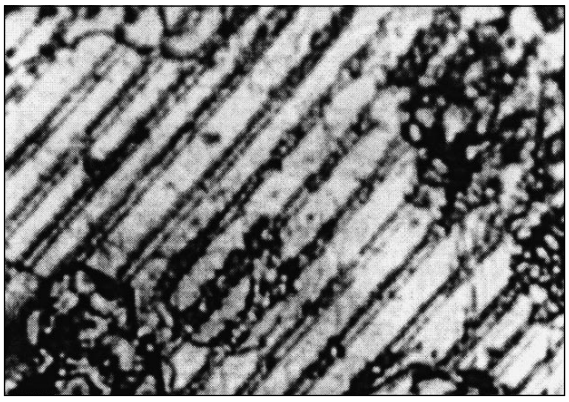


Fig. 2. Photograph of the surface of a $(SiC)_{0.93}(AlN)_{0.07}$ epitaxial film subjected to laser annealing. The parallel lines are due to the line scanning of the laser beam with a power density $P \geq 8 \times 10^3$ W/cm^2.

We emphasize that the probability of the formation of donor–acceptor pairs involved in radiative recombination is reasonably high both in undoped silicon carbide and in SiC with various impurities [7–9].

In the context of the discussed model of the donor–acceptor pairs, the bands at $h\nu_m \approx$ 2.3–2.4 eV correspond to the pairs of distant defects while those at $h\nu_m \approx$ 2.7 eV (Fig. 1, curve *3*) are ascribed to the pairs of the most closely spaced defects. The following two methods were used to estimate the interatomic distances (r_i) in such pairs:

(a) from the theory of donor–acceptor pairs using the formula $h\nu_i = E_g - (E_d + E_a) + e^2/\varepsilon r_i$ (E_d and E_a are the donor and acceptor level depths, ε is the permittivity) and under the assumption that ε depends on the interatomic distance in donor–acceptor pair [10]; and

(b) from the calculation of the exchange interaction energy [11].

Both calculations show that the PL band at $h\nu_m \approx$ 2.7 eV can be associated with donor–acceptor pairs with the distance between the defects in the SiC lattice no closer than 6 Å. The levels of more closely spaced defects shift into the respective band and do not reveal themselves in recombination radiation (see the insert in Fig. 1).

4. CONCLUSION

1. The spectral shift of the PL bands of $(SiC)_{1-x}(AlN)_x$ epitaxial films (x = 0.03–0.17) caused by laser annealing was observed.

2. It was proposed that Al_{Si}–N_C donor–acceptor pairs are formed in the course of laser annealing.

3. The increase in the annealing time was shown to result in the formation of donor–acceptor pairs of closely spaced defects Al_{Si}–N_C at the expense of associations of more distant defects.

ACKNOWLEDGMENTS

We thank Yu.M. Tairov for his helpful comments during the discussion of this paper.

REFERENCES

1. G. P. Yablonskiĭ, Fiz. Tekh. Poluprovodn. (Leningrad) **18**, 918 (1984) [Sov. Phys. Semicond. **18**, 570 (1984)].
2. G. P. Yablonskiĭ, Fiz. Tverd. Tela (Leningrad) **26**, 995 (1984) [Sov. Phys. Solid State **26**, 607 (1984)].
3. V. I. Levin, Yu. M. Tairov, and V. F. Tsvetkov, Fiz. Tekh. Poluprovodn. (Leningrad) **18**, 1194 (1984) [Sov. Phys. Semicond. **18**, 747 (1984)].
4. Yu. A. Vodakov, G. A. Lomakina, E. N. Mokhov, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 2153 (1986) [Sov. Phys. Semicond. **20**, 1347 (1986)].
5. *Materials for Electrical Engineering: A Reference Book* (Énergoatomizdat, Leningrad, 1988), Vol. 3, p. 464.
6. Yu. N. Émirov, G. K. Safaraliev, S. A. Ashurbekov, and M. K. Kurbanov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **28**, 1991 (1994) [Semiconductors **28**, 1098 (1994)].
7. Yu. A. Vodakov, G. A. Lomakina, E. N. Mokhov, *et al.*, in *Problems in Physics and Technology of Wide-Gap Semiconductors* (Leningrad, 1979), p. 164.
8. A. S. Berezhnoĭ, *Silicon and Its Binary Systems* (Akad. Nauk Ukr. SSR, Kiev, 1958).
9. A. I. Venger and Yu. A. Vodakov, Pis'ma Zh. Tekh. Fiz. **6**, 1319 (1981) [Sov. Tech. Phys. Lett. **6**, 566 (1981)].
10. K. Srinivasan and R. Srinivasan, in *Proceedings of the Nuclear Physics and Solid State Physics Symposium, Bombay, 1972*, Vol. 14c, p. 273.
11. A. N. Georgobiani, A. N. Gruzintsev, Yu. V. Ozerov, and I. M. Tiginyanu, Tr. Fiz. Inst. Akad. Nauk SSSR, No. 163, 39 (1985).

Translated by Yu. Aleshchenko

Semiconductors, Vol. 34, No. 8, 2000, pp. 894–896. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 932–934.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Gallium-induced Deep Level in $Pb_{1-x}Ge_xTe$ Alloys

E. P. Skipetrov*, E. A. Zvereva*, V. V. Belousov*, L. A. Skipetrova*, and E. I. Slyn'ko**

** Moscow State University (Department of Physics), Vorob'evy gory, Moscow, 119899 Russia*

e-mail: skip@mig.phys.msu.su

*** Institute of Materials Science Problems, Ukrainian National Academy of Sciences, Chernovtsy, 274001 Ukraine*

Submitted February 1, 2000; accepted for publication February 14, 2000

Abstract—Galvanomagnetic properties of gallium-doped $Pb_{1-x}Ge_xTe$ ($0.04 \le x \le 0.08$) alloys were studied. It is shown that gallium doping induces a deep impurity level in the alloy band gap; the energy position of this level depends on the germanium content in the alloy. A model assuming that doping with gallium results in the emergence of two defect levels in the energy spectrum of the alloys is proposed. *© 2000 MAIK "Nauka/Interperiodica".*

It is well known that the doping of lead telluride with gallium results in a series of unusual properties, the most important of which are pinning of the Fermi level in the band gap and persistent photoconductivity at low temperatures [1]. It is possible to explain most of the experimental results by assuming that doping with gallium introduces a deep level of the Jahn–Teller type in the PbTe band gap. This level is located approximately at an energy of 70 meV below the conduction band bottom. At the same time, there are experimental data indicating that a shallow metastable level induced by gallium can exist both below the conduction band edge [2] and as a resonance impurity level in the conduction band [3, 4].

The situation becomes even more uncertain when we regard the energy spectrum of the gallium-doped alloys on the lead telluride basis. A small amount of fragmentary data presented in [5–8] does not give, in particular, an indication of the existence of an impurity level (or levels) in the energy spectrum of $Pb_{1-x}Sn_xTe$:Ga and $Pb_{1-x}Ge_xTe$:Ga alloys and especially of the energy variation of this level with the alloy composition. Due to this fact, the main objective of this paper was to study the galvanomagnetic phenomena in $Pb_{1-x}Ge_xTe$:Ga alloys with various contents of germanium with the aim of detecting the impurity-induced deep defect levels and determining the energy spectrum of charge carriers in the gallium-doped alloys.

Single-crystal samples of $Pb_{1-x}Ge_xTe$ ($0.04 \le x \le 0.08$) doped with gallium ($C_{Ga} \approx 1.5$–3 at. %) and studied here were synthesized by sublimation from the vapor phase. The germanium content in the doped samples was determined by X-ray diffraction analysis, and the impurity concentration was controlled by gallium charge during growth with regard to the impurity distribution along the ingot. Temperature dependences of the resistivity and Hall coefficient ($B \le 0.1$ T) were measured in the temperature range of $4.2 \le T \le 300$ K.

Typical temperature dependences of dark resistivity in the samples with a different content of germanium are shown in Fig. 1. It is clearly seen that there exist two main features: an exponential portion in the high temperature region related to the impurity conductivity, which indicates that a deep impurity level of gallium exists in the alloy band gap, and an anomalously abrupt resistivity decrease at lower temperatures. The latter is related to the low-temperature transition of the alloys into the rhombohedral phase, which is typical of undoped $Pb_{1-x}Ge_xTe$ alloys [8, 9].

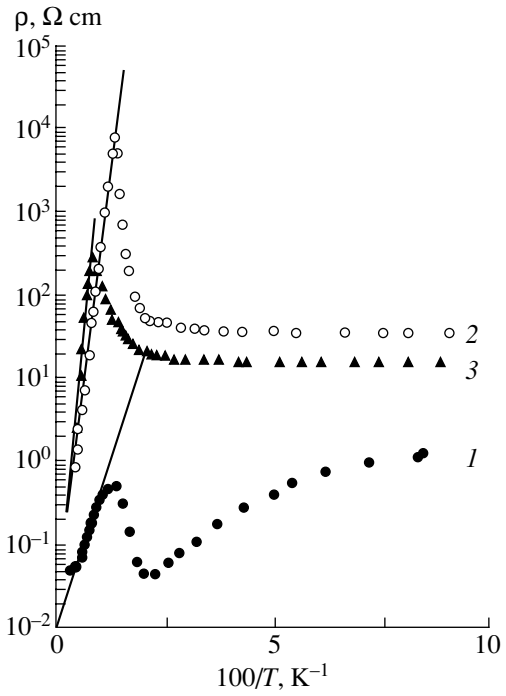


Fig. 1. Temperature dependences of the resistivity in $Pb_{1-x}Ge_xTe$:Ga alloys with germanium content x = (*1*) 0.04, (*2*) 0.06, and (*3*) 0.08.

1063-7826/00/3408-0894$20.00

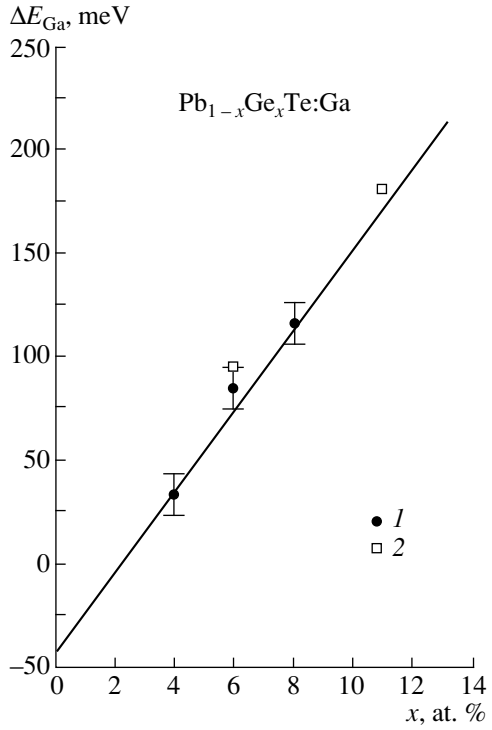


Fig. 2. Activation energy dependence of the gallium impurity level in $Pb_{1-x}Ge_xTe$:Ga on the germanium content in the alloy: (*1*) our data, (*2*) calculation according to [8].

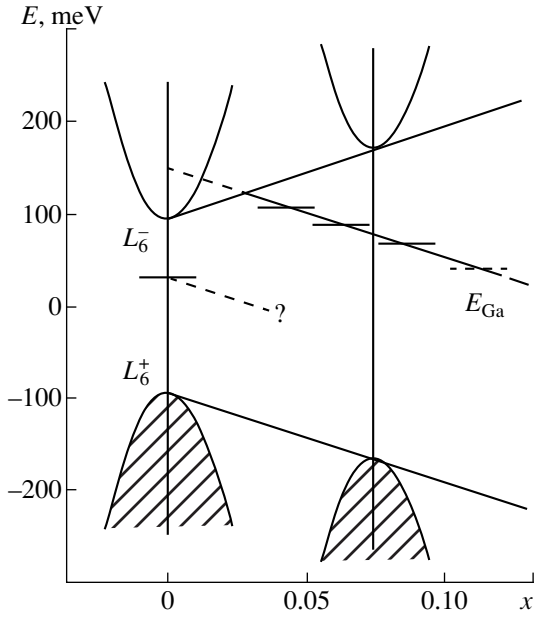


Fig. 3. Diagram of the energy spectrum transformation of charge carriers in the cubic phase of $Pb_{1-x}Ge_xTe$:Ga with a germanium content variation in the alloys.

As the content of germanium in alloys increases, an anomalous peak in the $\rho(1/T)$ curve is monotonically shifted to the high-temperature region. This correlates with the dependence of the phase transition temperature T_c on germanium content x in the undoped $Pb_{1-x}Ge_xTe$ alloys [9, 10]. However, the T_c value obtained from the peak position differs approximately by 80 K from the well-known values of the phase transition temperature in the undoped $Pb_{1-x}Ge_xTe$ alloys. It was found previously [10] that doping $Pb_{1-x}Ge_xTe$ alloys ($x = 0.05$) with indium lowered the phase transition temperature by 50 K as 2.5 at. % of the impurity was introduced. It is quite possible, therefore, that, as in $Pb_{1-x}Ge_xTe$ alloys doped with indium, a marked decrease in the phase transition temperature is related to the emergence of chaotically frozen-out polarized defects with randomly oriented dipole moments. Gallium atoms or complexes containing these atoms may play the role of such defects [10, 11]. It is assumed that due to polarization by these defects of some crystal volume they may reduce the molecular field and lower the phase transition temperature.

We determined the activation energy of the gallium impurity level in the alloys from the slope of the $\rho(1/T)$ curves in the exponential portion using the relation $\rho \approx \exp(\Delta E_{Ga}/kT)$. The results of our calculation are indicated in Fig. 2 by circles *1* (squares *2* refer to the calculations of $\rho(1/T)$ dependences obtained previously [8]). It is found that the activation energy increases linearly as the content of germanium in alloys increases and variation of the activation energy with x practically coincides with an increase in the band gap in the alloys with x [12]: $d\Delta E_{Ga}/dx \approx 20$ meV/at. %.

The experimental results allow us to propose a model of the energy spectrum transformation of $Pb_{1-x}Ge_xTe$:Ga alloys with alloy composition (Fig. 3). According to this model, doping with gallium leads to the emergence of a deep impurity level of gallium in the energy gap of the alloys. This level shifts almost linearly with reference to the conduction band bottom L_6^- as the germanium content in the alloy increases: $E_{Ga} - L_6^- \approx (45–1980x)$ meV. Extrapolation of the $\Delta E_{Ga}(x)$ dependence to $x = 0$ shows (Fig. 2) that, in PbTe:Ga, this level should be a resonance level and should lie approximately 45 meV above the bottom of the conduction band. This fact allows us to assume that doping of lead telluride and alloys on its basis with gallium should result in the emergence of two defect levels in the energy spectrum. In lead telluride, it is a well-known deep level which lies 70 meV below the conduction band bottom [1] and resonance level within the conduction band. A similar situation is realized, apparently, in $Pb_{1-x}Sn_xTe$ alloys doped with a gadolinium impurity with variable valence [13].

It is necessary to note that the diagram shown in Fig. 3 was obtained by estimating the activation energy of the deep level from temperature dependences of the resistivity in the cubic phase of the alloys. Due to this, in the low-temperature region for the alloys with a ger-

manium content $x > 0.01$, this diagram may be considered as virtual, because the crystal structure at low temperatures ($T < T_c$) is not cubic. However, for PbTe and $Pb_{1-x}Ge_xTe$ alloys with a germanium content $x < 0.01$, which have a cubic lattice over the entire temperature range, this energy diagram is quite realistic. Therefore, our conclusion about the existence of the gallium-induced resonance level in the conduction band of these materials is, from our point of view, beyond question.

In conclusion, in accordance with the diagram shown in Fig. 3, the energy position of the gallium level with respect to the valence band top remains practically unchanged. Unfortunately, the validity of this conclusion depends on the precision of the experimental determination of the energy gap variation with the germanium content in the alloys. However, data for the cubic phase of $Pb_{1-x}Ge_xTe$ alloys are practically lacking, and for the energy diagram formation we had to use the value of dE_g/dx average for these alloys in the rhombohedral phase [12].

REFERENCES

1. B. A. Akimov, A. V. Dmitriev, D. R. Khokhlov, and L. I. Ryabova, Phys. Status Solidi A **137**, 9 (1993).
2. A. I. Belogorokhov, I. I. Ivanchik, D. R. Khokhlov, and S. Ponomarev, Braz. J. Phys. **26**, 308 (1996).
3. G. S. Bushmarina, B. F. Gruzinov, T. T. Dedegkaev, *et al.*, Neorg. Mater. **16**, 2136 (1980).
4. Z. Feit, D. Eger, and A. Zemel, Phys. Rev. B **31**, 3903 (1985).
5. G. S. Bushmarina, B. F. Gruzinov, I. A. Drabkin, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **11**, 1874 (1977) [Sov. Phys. Semicond. **11**, 1098 (1977)].
6. B. A. Akimov, N. B. Brandt, L. I. Ryabova, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **31**, 304 (1980) [JETP Lett. **31**, 279 (1980)].
7. A. A. Averkin, G. S. Bushmarina, I. A. Drabkin, and Yu. Z. Sanfirov, Fiz. Tekh. Poluprovodn. (Leningrad) **15**, 197 (1981) [Sov. Phys. Semicond. **15**, 117 (1981)].
8. B. A. Akimov, A. V. Albul, I. I. Ivanchik, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **27**, 351 (1993) [Semiconductors **27**, 194 (1993)].
9. S. Takaoka and K. Murase, Phys. Rev. B **20**, 2823 (1979).
10. A. I. Lebedev and Kh. A. Abdullin, Fiz. Tekh. Poluprovodn. (Leningrad) **18**, 624 (1984) [Sov. Phys. Semicond. **18**, 388 (1984)].
11. A. P. Levanyuk, V. V. Osipov, A. S. Sigov, and A. A. Sobyanin, Zh. Éksp. Teor. Fiz. **76**, 345 (1979) [Sov. Phys. JETP **49**, 176 (1979)].
12. G. A. Antcliffe and R. A. Chapman, Appl. Phys. Lett. **26**, 576 (1975).
13. T. Story, M. Gorska, A. Lusakowski, *et al.*, Phys. Rev. Lett. **77**, 3447 (1996).

Translated by I. Kucherenko

Semiconductors, Vol. 34, No. 8, 2000, pp. 897–901. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 935–939.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Hole Concentration and Thermoelectric Figure of Merit for $Pb_{1-x}Sn_xTe$:Te Solid Solutions

G. T. Alekseeva*, M. V. Vedernikov*, E. A. Gurieva*, L. V. Prokof'eva*, and Yu. I. Ravich**

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

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

Submitted February 9, 2000; accepted for publication February 14, 2000

Abstract—The maximum densities of holes generated by cation vacancies, as well as thermoelectric parameters of $(Pb_{1-x}Sn_x)_{1-y}Te_y$ solid solutions with tin content x in the range from 0.4 to 0.6, were investigated. It is shown that each vacancy produces four holes in the valence band and that only for small x can the concept of doubly charged vacancy be used. The maximum thermoelectric figure of merit Z is $(1.0–1.1) \times 10^{-3}$ K^{-1} at T = 800–850 K. The relatively high value of Z achieved without doping is due to the high electrical conductivity provided, first, by the small effective mass of holes and, second, by the high electrical activity of the vacancies.

Thermoelectric generators based on IV–VI compounds and operating in the temperature range of 500–900 K are used in space technology and find military and industrial applications. The best of the known materials are PbTe (*n*- and *p*-type legs) and doped GeTe (*p*-type leg). No problems arise with *n*-PbTe application, as its thermoelectric figure of merit (parameter $Z = S^2\sigma/\kappa$, where S, σ, and κ are the Seebeck coefficient and electrical and thermal conductivities, respectively) is sufficiently large both at higher and lower temperatures. This permits the use of this material in a wider temperature range, with high average values of Z achieved by fabricating the *n*-leg of a thermoelectric device from several sections with different electron concentrations [1]. The situation with *p*-type materials is more complicated. High-performance *p*-type PbTe is obtained by heavily doping PbTe with Na [2]. However, this impairs the mechanical properties of the material [3], and the high hole concentration, necessary for optimization of thermoelectric performance, tends to decline with time, which leads to a gradual decrease in power and shortens the service life of a thermoelectric power device. GeTe-based materials are expensive, their melting point (~700°C) is not sufficiently high, and they have high volatility; as a result, they are of no practical use above 450°C.

Therefore, in many cases, other *p*-type materials that have lower figures of merit but are more favorable in other respects can be used. $Pb_{1-x}Sn_xTe$ solid solutions can serve as materials of this kind; in contrast to PbTe, they have a wide homogeneity region at high tin concentrations, shifted toward an excess of chalcogen [4]. A high hole concentration can be obtained in these materials by introducing only overstoichiometric Te, i.e., creating metal vacancies. This method rules out doping with the electrically active sodium and, consequently, eliminates all the detrimental factors associated with its presence. In this connection, much importance is acquired by the electrical activity of cation vacancies, i.e., the ratio of the hole to vacancy concentration (parameter c), which is, however, rather difficult to determine.

The experimentally determined value of c strongly depends on the Hall factor r used to determine the hole concentration from the Hall data. The value of r can be found by measuring the magnitude of the Hall effect in both weak and strong magnetic fields. By this method, $r = 0.6$ was obtained for SnTe [5], from which it follows that each tin vacancy is a doubly ionized acceptor ($c = 2$). Later, based on an analysis of data on galvanomagnetic effects in strong (quantizing) magnetic fields, it was concluded [6, 7] that the Hall factor is close to unity and that, hence, each metal vacancy introduces four holes to the SnTe valence band ($c = 4$).

For lead telluride, analysis of the hole concentration as a function of the partial pressure of Te_2 molecules in vapor, as a rule, leads to the conclusion that vacancies are single-charged [8]. At the same time, analysis of the electron energy spectrum in the presence of point defects [9, 10] suggests that each lead vacancy produces two holes. In contrast to [8, 11], physicochemical

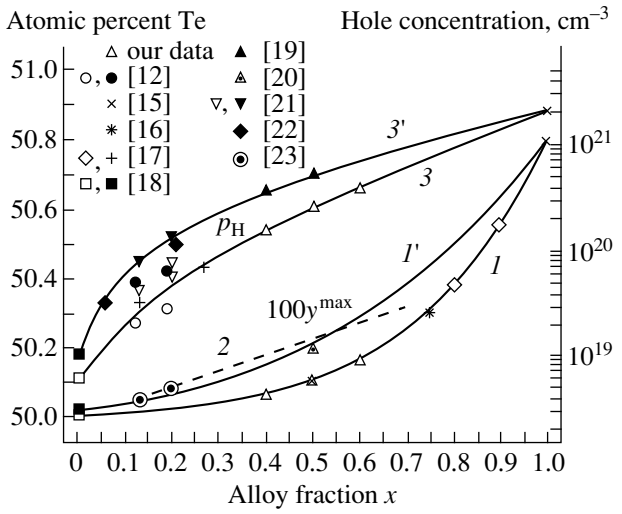


Fig. 1. Te-side boundary of the homogeneity region and the corresponding hole concentration (T = 77 K) as a function of the molarity x of the $(Pb_{1-x}Sn_x)_{1-y}Te_y$ solid solution. Sample annealing temperature were (*1*, *2*, and *3*) 350–500; and (*1*' and *3*') 700–750°C. Symbols denote our data and published data.

studies of $Pb_{1-x}Sn_xTe$ solid solutions with x in the range of 0.12–0.28 support the latter conclusion, indicating doubly ionized vacancies in PbTe as well [12]. Thus, the problem of the electrical activity of metal vacancies in $Pb_{1-x}Sn_xTe$ for different values of x requires further study and is one of the objectives of this work.

In this study, the maximal Hall concentrations of holes produced by metal vacancies were determined experimentally and thermoelectric properties of $Pb_{1-x}Sn_xTe$ with x = 0.4–0.6 were studied in the temperature range T = 300–850 K. This material was chosen as the most promising for solving thermoelectricity problems, with a tin content sufficiently high to make the lattice thermal conductivity as low as possible and, simultaneously, to obtain a high hole concentration, comparable with the optimal value. The high hole concentration suppresses to the required extent the influence of the intrinsic conductivity on Z in the narrow-gap $Pb_{1-x}Sn_xTe$ materials at operating temperatures and enables the evaluation of the effect of the complex hole spectrum on thermoelectric properties.

$(Pb_{1-x}Sn_x)_{1-y}Te_y$ samples with x = 0.4–0.6 and a Te content y of up to 50.45 at. % were fabricated using conventional powder-metallurgy technology [13]. Assuming a Hall factor r = 1, in accordance with data for PbTe [14] and SnTe [6], we found from the experimental Hall data that the highest possible room-temperature hole concentration is achieved on adding excess Te grows with increasing x in the range (0.6–2.2) × 10^{20} cm^{-3} for the compositions under study. Experimental data for x = 0.5 at 77 K show that the ratio of the Hall concentration of holes (with r = 1) to the density of overstoichiometric Te atoms, previously determined for $Sn_{1-y}Te_y$ to be 4 [6, 15], is about the same for the solid solution with 50.025 at. % Te. For 50.075 at. % Te, this ratio is 3.4. A possible explanation for the lower value may be that not all of the excess Te atoms were dissolved in the lattice. According to our experimental data, the maximal hole concentration at x = 0.5 and 77 K is 2.5×10^{20} cm^{-3}. Assuming that the ratio c = 4 is valid for the entire doping range, we determined the limit of the homogeneity region at 500°C (sample annealing temperature) to be about 50.103 at. % Te. This value agrees well with X-ray data for the material (50.1 at. % Te at 600°C) [16], thus confirming once more the assumption of quadruple ionization of metal vacancies in $Pb_{1-x}Sn_xTe$ solid solutions near x = 0.5.

For solid solutions with x = 0.4 and 0.6, the maximal hole concentrations at 77 K are 1.5×10^{20} and 3.8×10^{20} cm^{-3}, respectively, which corresponds to y^{max} of 50.062 and 50.156 at. %, respectively. To construct the composition dependence of y^{max} for the entire system (at 350–500°C), we used published X-ray data for solid solutions with x = 0.8 and x = 0.9 (356°C) [17] and for SnTe (500°C) [15]. As for the y^{max} value for PbTe, we used the data of [18] with some correction, with hole concentration determined at liquid-nitrogen, instead of room temperature [19]. As is known, these data have been obtained assuming singly ionized vacancies in PbTe. Below, we discuss how the $y^{max}(x)$ dependence changes if c = 2 is taken for PbTe; for now, we use the data obtained assuming that c = 1. Figure 1 (curve *1*) shows the corresponding dependence $y^{max}(x)$. The y^{max} value for x = 0.75 at 600°C, derived from X-ray data [16], agrees with this curve. A larger value of y^{max} for x = 0.5 (580°C) was obtained from the temperature dependence of Te vapor partial pressure by the optical absorption method [20]. The dashed straight line (line *2*) in Fig. 1 yields values of y^{max} obtained for alloys with $x < 0.8$ assuming a linear variation of the lattice constant with composition, as in [17]. This line is seen to run noticeably higher than the experimental curve; therefore, the use of corresponding data to determine the number of holes per vacancy (c parameter) in these alloys [17] would yield underestimated values.

Let us now find the composition range of the $Pb_{1-x}Sn_xTe$ solid solution in which the recharging of metal vacancies produces quadruply ionized defects. For this purpose, we plot, in Fig. 1, our data and published data [12, 15, 17, 18, 21] on the Hall concentration of holes p_H (77 K) corresponding to the Te-side boundary of the homogeneity region at 500°C for alloys with different x. Averaging the results, we obtain curve *3* in Fig. 1. Then, using the hole concentrations corresponding to curve *3*, we calculate the values of y^{max} for c = 2 and c = 4 and compare them with curve *1*. The resulting curves *4* and *5* are shown in Fig. 2

together with the enlarged initial portion of curve *1* in Fig. 1. Curve *4* rises steeply from the very beginning, giving no reason to assume only double ionization of vacancies even at a low content of tin. The curve corresponding to $c = 1$ would rise even more steeply. For small x, curve *5* gives noticeably lower y^{max} values than the experimental curve *1*, and it merges with curve *1* only at an x of about 0.35.

Hence, introduction of even a small amount of tin changes the acceptor behavior of defects in the solid solution. The assumption of identical behavior of metal vacancies in solid solutions of different compositions is unjustified. It is necessary to take into account the conditions of vacancy formation in the crystal, i.e., their positions in the lattice, the nearest neighbor configuration, and possible interactions of the vacancies with one another and atoms of the base crystal components. In other words, account should be taken of all the main factors governing the electronic structure of the defect.

As stated above, the homogeneity region for PbTe was estimated assuming singly ionized vacancies. If the dependence $y^{max}(x)$ is constructed assuming doubly ionized defects, the portion of the curve *1* near $x = 0$ (Fig. 2) will run slightly lower (shown by the dashed line) but will make no significant corrections to the results obtained above. Only the ratio of contributions of doubly and quadruply ionized vacancies to the Hall concentration of holes in the alloys with $x < 0.25$ will change. Thus, it can be assumed for $x > 0.3$ that each vacancy supplies four holes to the valence band.

The obtained values of the parameter c in $Pb_{1-x}Sn_xTe$ solid solutions with $x > 0.3$ were used to plot the Te-side boundary of the homogeneity region at higher temperatures of about 700–750°C (Fig. 1, curve *1*'). In doing so, we used the averaged dependence of the Hall concentration of holes on x, constructed using published data [12, 15, 18, 19, 21, 22] for the same temperatures (curve *3*'). The experimental values of y^{max} for the solid solutions with $x = 0.13$ and $x = 0.20$ at intermediate temperatures [23], presented in Fig. 1, are overestimated, similarly to the above value at $x = 0.5$ [20], obtained by the same method. However, the trends in the behavior of the parameter c, established in [23] using also p_H values higher than those corresponding to curve *3*', agree qualitatively with our data. Namely, in both alloys, c exceeds 2 (2.31 and 2.79, respectively) and the parameter c increases with x.

The unusual changes in properties of *p*-type lead salts on isoelectronic substitution of lead for tin has previously been observed. For example, such a change was observed in dilute Na-doped $Pb_{1-x}Sn_xTe$ (x = 0.005–0.01) solid solutions and was explained by transformation of metal vacancies into antisite Te positions, with the corresponding transformation of the electronic structure of the defects [24]. An even more unusual situation is observed in $Pb_{1-x}Sn_xSe$ and $Pb_{1-x}Sn_xS$ systems, where, under the effect of acceptor centers, tin

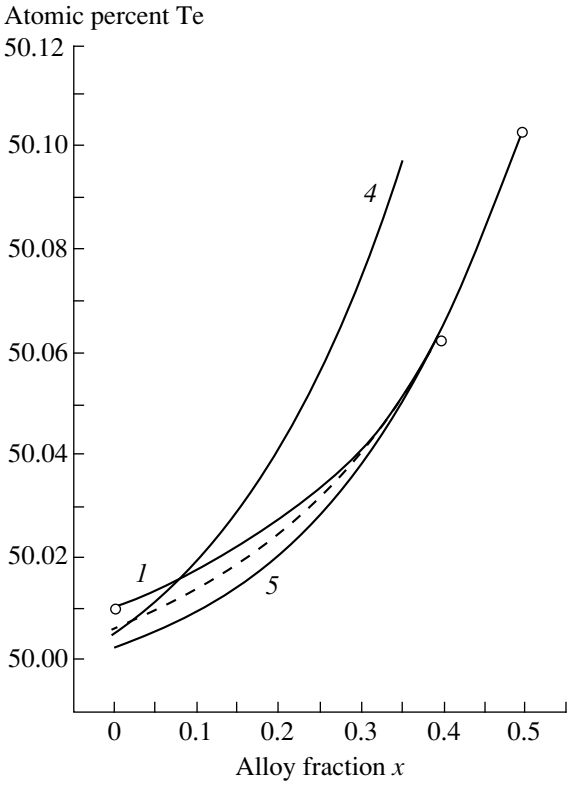


Fig. 2. Te-side boundary of the homogeneity region for $(Pb_{1-x}Sn_x)_{1-y}Te_y$ solid solutions, $x \le 0.5$, on an enlarged scale. Curve *1* is experimental and curves *4* and *5* correspond to calculations for $c = 2$ and 4, respectively.

atoms convert to a charge state +4 on capturing each two free holes from the valence band [25, 26]. Our new results are consistent with these concepts.

The thermoelectric properties of the solid solutions were studied for samples with $x = 0.4$, 0.5, and 0.6 in the temperature range of 300–850 K. Figures 3 and 4 show temperature dependences of thermoelectric power and electrical and thermal conductivities. Using these data, the temperature dependence of the thermoelectric figure of merit Z was calculated and plotted. For comparison, similar data for *p*-type PbTe with 2 at. % Na are presented. The change in the kinetic properties on introduction of tin is primarily due to the band-gap narrowing and the ensuing diminution of the effective hole mass. As a result, the electrical conductivity of these alloys is relatively high, being no lower than that in *p*-type PbTe at comparable hole concentrations, even though additional (and substantial) impurity scattering occurs in the alloys, as indicated by a much weaker $\sigma(T)$ dependence. As follows from the results presented above, a higher electrical conductivity can also be ensured by high electrical activity of the metal vacancies ($c = 4$).

Owing to the high electrical conductivity and the weaker temperature dependence of not only σ but also κ (due to the impurity component of phonon scatter-

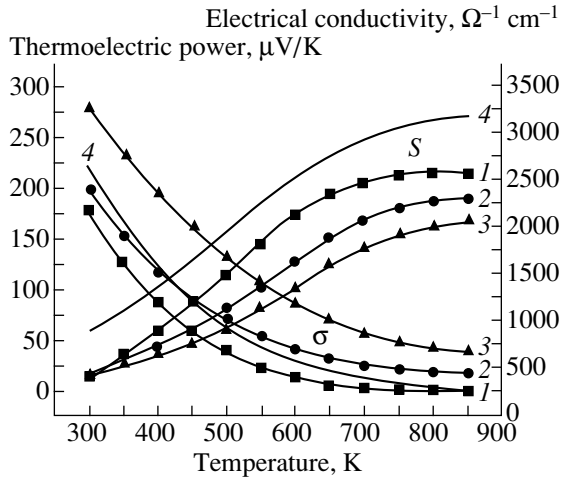


Fig. 3. Temperature dependences of the thermoelectric power (S) and electrical conductivity (σ) for $Pb_{1-x}Sn_xTe$:Te solid solutions with x = (*1*) 0.4, (*2*) 0.5, and (*3*) 0.6, and with (*4*) PbTe:Na.

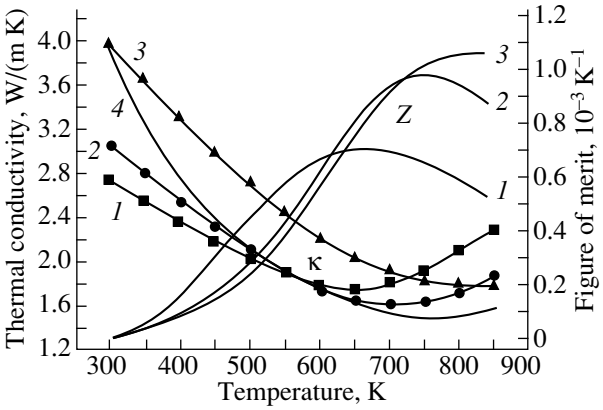


Fig. 4. Temperature dependences of the thermal conductivity (κ) and thermoelectric figure of merit (Z) for $Pb_{1-x}Sn_xTe$ solid solutions. Curve numbers are the same as in Fig. 3.

ing), the decrease in the total thermal conductivity of the samples is insignificant (alloys with $x = 0.4$ and 0.5 at low T). In the solid solution with $x = 0.6$, which has the highest hole concentration, the contribution from the electron thermal conductivity is so large that the decrease in the lattice heat conductivity causes no diminution of the total thermal conductivity. The narrow band gap is also responsible for the stronger influence of intrinsic charge carriers on the properties of the material. It is this effect that primarily changes the temperature dependence of thermal conductivity, limiting the growth of the thermoelectric figure of merit and causing its subsequent decrease with increasing temperature. The maximal Z is $(1.0–1.1) \times 10^{-3}$ K^{-1} and is observed at 800–850 K.

The complex spectrum of holes affects primarily the temperature dependence of the thermoelectric power. The increase in this coefficient is significant, greatly exceeding that expected for the small effective hole mass in the main L-extremum and for high hole concentration in the samples. At high temperatures, the thermoelectric power becomes equal, or nearly equal, to the optimal values for a single-band semiconductor (200 μV/K).

The materials studied can find their specific application in the p-type leg of medium-temperature thermoelectric generators, and, in addition, can successfully replace SnTe as a high-temperature extension to GeTe. Such a replacement not only enhances the generator efficiency, but also improves the technology of thermoelement fabrication because of the higher melting point of the solid solution.

ACKNOWLEDGMENTS

The work was supported by the Program “Integration,” project no. 75.

REFERENCES

1. V. P. Vedeneev, M. I. Zaldastanishvili, S. P. Krivoruchko, *et al.*, in *Proceedings of the International Workshop on Thermoelectrics and Their Applications, St. Petersburg, 1999*, p. 231.
2. G. T. Alekseeva, E. A. Gurieva, P. P. Konstantinov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 2159 (1996) [Semiconductors **30**, 1125 (1996)].
3. M. S. Ablova, Fiz. Tverd. Tela (Leningrad) **12**, 3585 (1970) [Sov. Phys. Solid State **12**, 2910 (1970)].
4. V. L. Kuznetsov and V. P. Zlomanov, Neorg. Mater. **35**, 263 (1999).
5. B. B. Houston, R. S. Allgaier, I. Babiskun, and P. G. Siebenmann, Bull. Am. Phys. Soc. **9** (1), 60 (1964).
6. N. V. Kolomoets, S. A. Laptev, and E. I. Rogacheva, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 447 (1986) [Sov. Phys. Semicond. **20**, 283 (1986)].
7. S. A. Laptev and E. I. Rogacheva, in *Physical Electronics* (Vishcha Shkola, Lvov, 1986), Vol. 32, p. 56.
8. N. Kh. Abrikosov and L. E. Shelimova, *Semiconducting Materials Based on IV–VI Compounds* (Nauka, Moscow, 1975), Chap. 3, p. 47.
9. N. J. Parada and G. W. Pratt, Phys. Rev. Lett. **22**, 180 (1969).
10. B. A. Volkov and O. A. Pankratov, Zh. Éksp. Teor. Fiz. **88**, 280 (1985) [Sov. Phys. JETP **61**, 164 (1985)].
11. T. V. Saunina, D. B. Chesnokova, and D. A. Yas’kov, Izv. Leningr. Élektrotekh. Inst. im. V. I. Ul’yanova, No. 302, 74 (1982).
12. T. V. Saunina, D. B. Chesnokova, and D. A. Yaskov, Phys. Status Solidi A **94**, 161 (1986).
13. Yu. I. Ravich, B. A. Efimova, and I. A. Smirnov, *Semiconducting Lead Chalcogenides* (Nauka, Moscow, 1968; Plenum, New York, 1970), Chap. 1.
14. R. S. Allgaier and B. B. Houston, J. Appl. Phys. **37**, 302 (1966).

15. R. F. Brebrick, J. Phys. Chem. Solids **24**, 27 (1963).
16. S. A. Miloslavov, S. M. Tairov, and B. F. Ormont, Izv. Akad. Nauk SSSR, Neorg. Mater. **6**, 2063 (1970).
17. R. F. Brebrick, J. Phys. Chem. Solids **32**, 551 (1971).
18. R. F. Brebrick and R. S. Allgaier, J. Chem. Phys. **32**, 1826 (1960).
19. C. R. Hewes, M. S. Adler, and S. D. Senturia, J. Appl. Phys. **44**, 1327 (1973).
20. Y. G. Sha, K. T. Chen, and R. F. Brebrick, J. Electrochem. Soc. **138**, 2460 (1991).
21. A. R. Calava, T. C. Harman, M. Finn, and P. Youtz, Trans. Metall. Soc. AIME **242**, 374 (1968).
22. G. Dionne and I. C. Wooley, J. Electrochem. Soc. **119**, 784 (1972).
23. Yu. Huang and R. F. Brebrick, J. Electrochem. Soc. **135**, 1547 (1988).
24. G. T. Alekseeva, B. G. Zemskov, P. P. Konstantinov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 358 (1992) [Sov. Phys. Semicond. **26**, 202 (1992)].
25. L. V. Prokof'eva, M. N. Vinogradova, and S. V. Zarubo, Fiz. Tekh. Poluprovodn. (Leningrad) **14**, 2201 (1980) [Sov. Phys. Semicond. **14**, 1304 (1980)].
26. F. S. Nasredinov, L. V. Prokof'eva, and P. P. Seregin, Zh. Éksp. Teor. Fiz. **87**, 951 (1984) [Sov. Phys. JETP **60**, 542 (1984)].

Translated by D. Mashovets

Semiconductors, Vol. 34, No. 8, 2000, pp. 902–907. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 940–946.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Fine Structure of the Dielectric-Function Spectrum in Diamond

V. V. Sobolev*, A. P. Timonov, and V. Val. Sobolev

Udmurt State University, Krasnoarmeĭskaya ul. 71, Izhevsk, 426034 Russia

** e-mail: sobolev@uni.udm.ru*

Submitted March 17, 1999; accepted for publication February 17, 2000

Abstract—The spectra of the complete set of optical functions for three different diamond samples were calculated for the range of 0–32 eV. Calculations were based on the features of experimental reflection spectra; the Kramers–Kronig method was used. The features of the optical-function spectra and their distinctions for the three samples were analyzed. The method of the combined Argand diagrams was used to decompose for the first time the dielectric-function spectra into elementary components. The energies of these components and the oscillator strengths were determined. The structure of the components was compared to the structures of theoretical spectra of permittivity and the expected spectra of band-to-band transitions. Good agreement between the results of this study and the theory was obtained. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Diamond is a very popular crystal in science and technology due to the fact that it possesses a number of unique properties [1–3]. It is not by chance that a large group of covalent semiconducting materials has been named after diamond, i.e., "diamond-like." Diamond is a classical model crystal in the theory of the electronic structure of solids. Numerous theoretical calculations of its energy levels (bands) have been performed over the past several decades [3–7]. The simplest types of the interatomic bonds and the unit-cell structure are characteristic of diamond. Nevertheless, the results of calculating the bands differ markedly from publication to publication. For many years, the disagreement between the band theory and experiment concerning the spectrum of imaginary part ε_2 of dielectric constant has remained unclarified. This disagreement consists mainly in the fact that the principal peak in the theoretical spectrum of ε_2 is shifted to higher energies and is too intense compared to what is observed in the experimental spectrum. This problem is so fundamental that several teams of theoreticians simultaneously began to reconsider it in recent years [8–12]. The best currently available methods were used with consideration of the exchange correlation; the nonlocality; the many-particle behavior; and, sometimes, the excitons within a wide energy range of 4–20 eV. Theoreticians have succeeded in bridging the gap between the theory and experiment. However, the features of the theoretical spectra of ε_2 remain markedly different from those of experimental spectrum. Furthermore, especially large discrepancies between the theory and experiment are observed for the spectra of ε_1 and R; as a rule, these discrepancies are explained by ignorance or incomplete consideration of the role of metastable excitons in the energy range of $E > E_g$. Unfortunately, other fundamental issues related to the spectrum of ε_2 in diamond were not considered at all [8–12]; these issues include the energy and the probability of a fine complex structure of optical transitions and their nature in the models of bands or metastable excitons. In order to study these issues, much more complete experimental and experimental–calculational data on electronic structure are needed.

It is universally accepted that the most complete information about special features of the electronic structure of a crystal resides in a wide set of optical fundamental functions. The Argand combined-diagram method [13] yields some basically new and very important parameters, such as the energies and probabilities of transitions. The spectra of only some optical functions, which are noticeably different according to the data in various publications [3], are known for diamond.

The objective of this study was to determine a complete set of optical functions for three diamond samples, to find the energies and oscillator strengths for optical transitions, and to analyze the results and discuss their physical significance in the band model.

THE METHODS OF CALCULATIONS

The complete set of optical functions includes the coefficients of reflection R and absorption μ; the imaginary ε_2 and real ε_1 parts of the dielectric constant ε; the absorptive k and refractive n indices; the integral function of the density of bound states multiplied by the transition probability and equal (to within a constant multiplier) to $\varepsilon_2 E^2$; the effective number of valence electrons $n_{ef}(E)$ involved in the transitions up to the

1063-7826/00/3408-0902$20.00

given energy E; the effective dielectric constant ε_{ef}; the functions of the characteristic electron energy losses for bulk ($-\mathrm{Im}\varepsilon^{-1}$) and surface ($-\mathrm{Im}(1+\varepsilon)^{-1}$) plasmons; the phase θ of reflected light; and the electrooptical differential functions α and β. All these functions are interrelated; however, each of them has a significance in its own right. Their physical significance and interrelation follow directly from the Maxwell equations [13].

The most-used method for obtaining this set of functions consists in calculations based on special programs, the known experimental reflection spectrum in a wide energy range, and the Kramers–Kronig relations.

The problem of determining the fine-structure parameters in the spectrum of optical transitions (the energy E_i and half-width Γ_i of the component bands and the corresponding oscillator strengths f_i) is typically solved by either of two methods: (i) by reproducing the integral curve of the spectrum for ε_2 with a set of Lorentzian oscillators N with a large number of adjustable parameters $3N$ (as large as 30 for $N = 10$) or (ii) by the method of the Argand combined diagrams. The latter method is also based on the classical model of Lorentzian oscillators, but without adjustable parameters due to simultaneous analysis of the spectra for ε_2 and ε_1.

The methods we used here for calculating the complete set of optical functions including the decomposition of the integrated spectrum of ε_2 into elementary components were outlined in detail in [3, 13–16] and were discussed in [17–19].

CALCULATIONS OF THE SET OF OPTICAL FUNCTIONS

The transmission spectra were analyzed [20] to determine the energies of indirect (E_{gi}) and direct (E_{gd}) transitions at 295 K in diamond: $E_{gi} = 5.470$ eV and $E_{gd} = 7.02$ eV.

The reflection spectrum of diamond at 300 K in the range of 4–23 K features a low-intensity peak at ~7.1 eV, an intense peak at 12.85 eV, and a very wide plateau in the range of 16–20 eV [21]. It was intimated that the reflection coefficients of untreated and mechanically polished samples were not much different from each other; this distinction amounted to $\Delta R = 10\%$ for the samples studied. These experiments were extended to the polished samples of type II-a diamond in the range of 4–30 eV with an error of 5% [22]. In [22], two peaks at ~17 and 20.5 eV were observed instead of the plateau [21] and the entire reflection curve [22] ran noticeably below than that in [21]. The repeated measurements [23, 24] confirmed basically the conclusions [21] concerning the structure of the spectrum of R. Studies of the reflection spectra of the polished (type I) and cleaved (type II-a) samples also showed that the longest wavelength low-intensity peak had a doublet structure; i.e., along with the known peak at ~7.1 eV, a peak at ~7.6 eV was observed if measurements were performed at 90 K. Detailed ellipsometry studies performed at 90 K in a fine vacuum confirmed the presence of a high-energy component at ~7.6–7.8 eV [25].

Thus, the reflection spectra [21, 23, 24] of three different diamond samples are noticeably different, especially so in the value of R in a wide energy range and in the structure of the long-wavelength peak. Naturally, this brought about the distinctions in the spectra of other optical functions calculated from the reflection spectra using the Kramers–Kronig relations. Unfortunately, in [21, 23, 24], only the spectra of some of the functions were reported and an analysis of their distinctions was lacking. As a result, for such a popular crystal as diamond, we have now neither the complete set of optical functions nor an analysis of their dependence on the type of the sample, if only by the example of the three samples measured in [21, 23, 24].

We calculated the complete sets of optical functions of diamond in the range of 0–34 eV on the basis of the reflection spectra of samples studied in [23] (sample 1), the polished sample of type I (sample 2), and a cleaved sample of type II-a (sample 3) [24]. A total of 14 functions were calculated. For the sake of brevity, we report only the spectra of R, ε_1, and ε_2 in this paper (Fig. 1). The distinctions between the spectra of the three samples are clearly seen.

The most intense peak in the experimental reflection spectra is located at 12.85 (sample 1), 12.55 (sample 2), and 12.77 eV (sample 3) with $R \approx 0.65$ (sample 1), 0.64 (sample 2) and 0.55 (sample 3). The long-wavelength part of this band is almost the same for the three samples, whereas the short-wavelength part differs significantly, especially for samples 1 and 2. In the region of the plateau (16–18 eV), the data for R for the three samples differ drastically in the magnitude and shape of the reflectivity curve. The long-wavelength component of the band in the vicinity of 7.5 eV is located at ~7.25 (sample 1), 7.15 (sample 2), and 7.10 eV (sample 3); the short-wavelength component is located at ~7.62 (sample 2) and 7.67 eV (sample 3) (see the insert in Fig. 1a). Thus, the reflection spectra for two different polished samples and for the cleaved sample differ radically in the range of 12.5–21 eV. The weakest reflection observed for the cleaved sample is apparently caused by imperfections of the cleaved surface. The lower reflectivity of sample 1 compared to sample 2 is possibly related to special features of the methods used for polishing. The intense reflectivity peak in sample 2 is shifted to longer wavelengths by ~0.3 and 0.2 eV with respect to the corresponding peaks in samples 1 and 3; the origin of this shift is unclear. The intensity of the long-wavelength component of the weak-reflectivity peak in the range of 7.10–7.25 eV depends only slightly on temperature in the range of 90–300 K; the short-wavelength component is observed only at 90 K [26].

Unfortunately, the surface structure of the samples was not studied in [21–26] and even the methods used for polishing and cleaving were not described, which

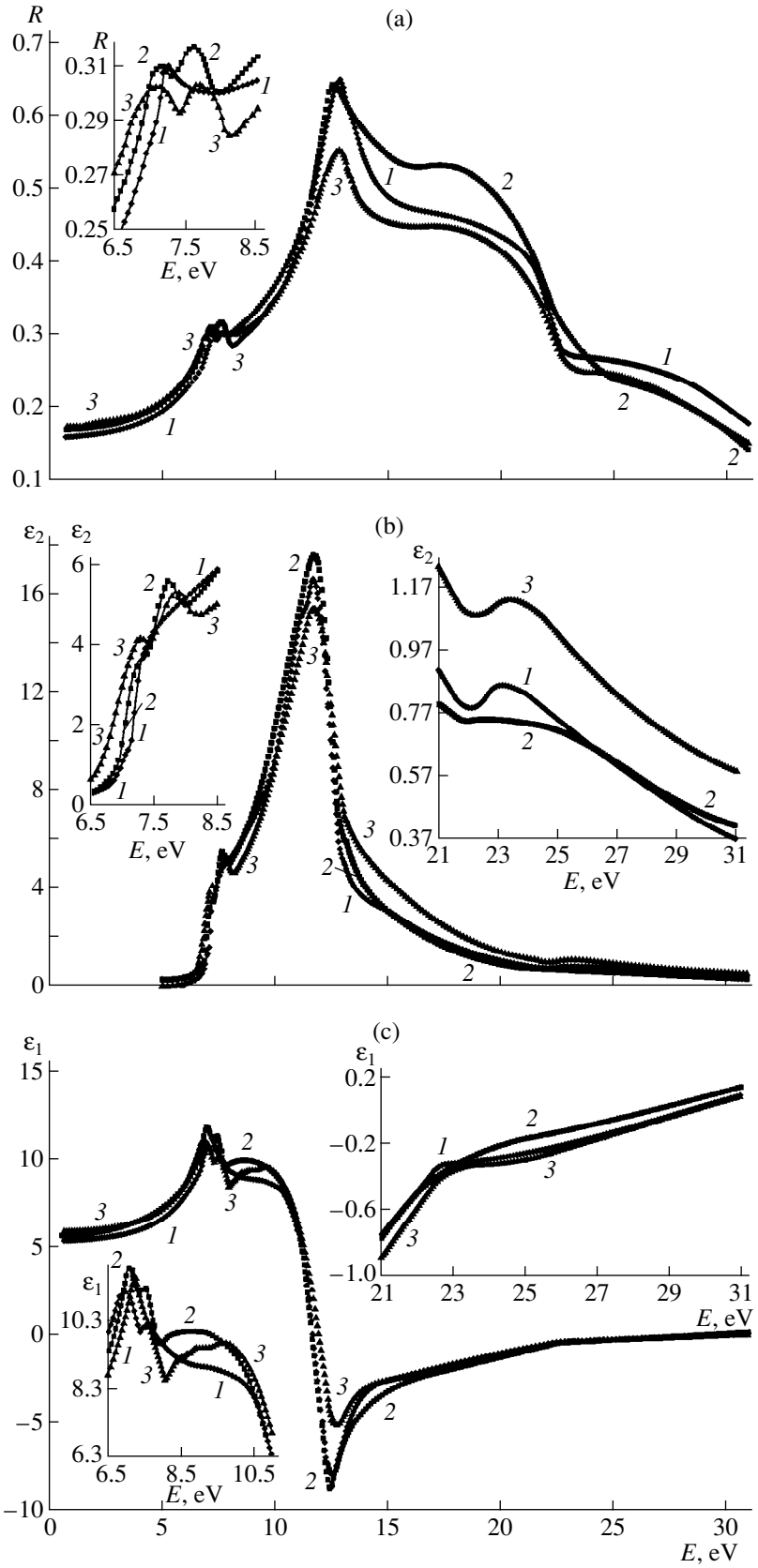


Fig. 1. The spectra of (a) reflectivity, (b) ε_2, and (c) ε_1 for three diamond samples (samples 1, 2, and 3); the spectra in the ranges of 6.5–8.5 and 21–31 eV are shown in the inserts.

makes the discussion of the origin of observed distinctions in the reflection spectra of the three diamond samples much more difficult. Therefore, in the future, it is of basic importance to study the spectra of reflection from a diamond surface thoroughly cleaned in high vacuum.

One of the most important general features of the ε_2 spectra consists in a drastic decrease in the values of ε_2 in the range of energies higher than the energy of the most intense peak. Because of this, even the intense reflectivity features located at energies higher than that of the most intense band are not observed in the spectrum of ε_2 at all. As might be expected, the calculated curves $\varepsilon_2(E)$ for the three diamond samples descend steeply beyond the peak at ~11.8 eV. As a result, the curves converge strongly in the entire wide energy range, the asymmetric intense reflectivity peak for sample 2 is transformed into a symmetric peak, and the positions of this peak for the three different diamond samples become almost identical (Fig. 1b). A large difference between the values of R at the peak of the main reflection band for samples 2 and 3 diminished appreciably in the curves for ε_2; furthermore, the almost identical R for samples 1 and 2 transformed into an appreciable difference in the intensities of the main peak of ε_2 for these samples. These nontrivial specificities are mainly caused by the distinctions between the intensities of corresponding reflection spectra in the range of 13–20 eV.

A wide reflectivity plateau in the range of 24–27 eV transformed into a poorly pronounced peak of ε_2 at ~23.1 (sample 1), 23.6 (sample 2), and 23.5 eV (sample 3) with $\varepsilon_2 \approx 0.8$–1.1.

The long-wavelength band of ε_2 involves very poorly pronounced peaks at ~7.5 (sample 1) and 7.3 eV (sample 2) and also poorly pronounced peaks at ~7.75 (sample 2), 7.3 (sample 3), and 7.9 eV (sample 3) (see the insert in Fig. 1b).

The most intense peak of ε_2 at ~11.8 eV is shifted by ~1 eV to lower energies, whereas the low-intensity band in the region of ~7.2–7.8 eV is shifted by ~0.20 eV to higher energies compared to the corresponding peaks in the reflection spectrum.

The main reflection band is represented in the spectrum of ε_1 by a poorly pronounced broad band at ~9.0 eV and a sharp negative minimum at ~12.55 (sample 1), 12.45 (sample 2), and 12.75 eV (sample 3); the position of this minimum almost coincides with that of the reflectivity peak. The long-wavelength peaks of ε_1 are observed at ~7.20 (sample 1), 7.05 and 7.50 (sample 2), and 6.95 and 7.55 eV (sample 3) (see the insert in Fig. 1c); i.e., these peaks are shifted by ~0.1–0.2 and 0.3 eV to lower energies compared to their analogues in the spectra of R and ε_2. As the energy E increases, the curve $\varepsilon_1(E)$ ascends beyond the negative minimum and almost levels off in the region of $E > 25$ eV; $\varepsilon_1 = 0$ (i.e., $n = k$) for $E = 29.5$ eV.

Thus, special features of a diamond sample profoundly affect the reflection spectrum in a wide energy range (especially in the range of 12–22 eV, including the position, intensity, and shape of the bands), clearly manifest themselves in the spectrum of ε_1 (especially in the magnitude of the negative minimum and in the structure of long-wavelength band), and affect very slightly the spectrum of ε_2. Therefore, the distinctions between the samples can be studied most effectively by analyzing the reflection spectra in the range of 12–22 eV and the spectra of ε_1 in the range of 6.5–9.5 eV.

Energies of the peaks E_i (in eV) and the half-widths Γ_i (in eV) of the bands, the oscillator strengths f_i, the values of ε_2 at the peak of the band ($\varepsilon_{2\,\max}$), and the areas under the bands S_i (in eV) for diamond sample 2; in the far right column, interpretations of the bands and their energies (in eV) are given according to the theory [6]

O_i	E_i	Γ_i	f_i	$\varepsilon_{2\max}$	S_i	$S_i/S_{\max}$	[6]
1	7.25	0.20	0.42	1.65	0.33	0.02	$\Gamma_{25'}$–$\Gamma_{2'}$, 7.0 (M_0)
2	7.75	0.60	1.05	3.69	2.21	0.13	$\Gamma_{25'}$–Γ_{15}, 8.1 (M_0); Λ_3–Λ_1, 8.0
3	8.40	0.60	0.25	1.67	1.00	0.06	Δ_5–Δ_1, 8.2; $L_{3'}$–L_2, 8.3 (M_1)
4	9.05	1.20	0.69	3.58	4.30	0.25	Transitions 4 ⟶ 5 in the volume of the Brillouin zone; 8.4–11.5; in the vicinity of K, U, X, Σ_2–Σ_3, 11.0
5	9.60	0.70	0.15	1.89	1.32	0.08	
6	10.35	1.30	0.75	7.32	9.52	0.55	
7	11.10	0.80	0.26	5.77	4.62	0.27	
8	11.85	1.20	0.70	14.39	17.27	1.00	
9	12.30	0.30	0.03	3.05	0.92	0.05	K_2–K_3, 11.7
9'	12.80	0.68	0.03	1.18	0.80	0.05	X_4–X_1, 11.8 (M_1)
10	13.80	2.10	0.14	2.16	4.53	0.26	Λ_3–Λ_3, 13.8 (M_1) (4–6)
11	16.00	2.70	0.13	1.51	4.09	0.24	Δ_5–$\Delta_{2'}$, 14.4 (M_2)
12	18.50	5.90	0.12	0.63	3.73	0.22	Σ_2–Σ_2, 18.0; $\Gamma_{25'}$–$\Gamma_{2'}$, 19.0

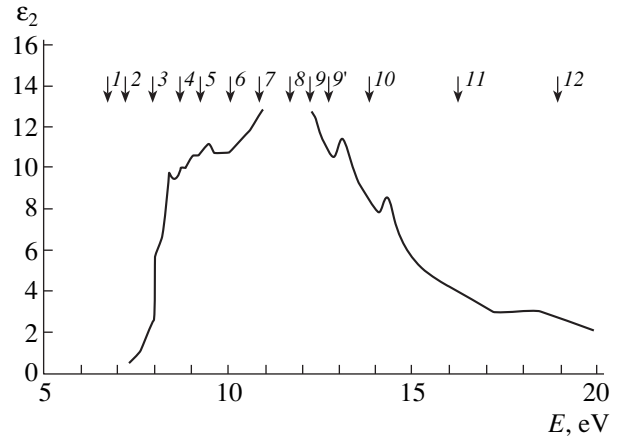


Fig. 2. Theoretical spectrum of ε_2 for diamond [6] (the solid-line curve); the arrows indicate the energies of isolated oscillators O_i.

DECOMPOSITION OF THE SPECTRUM OF ε_2 INTO COMPONENTS

Experimental reflection spectrum and the spectral curves of the other optical functions calculated from this spectrum are the result of summing all the individual numerous transitions from occupied states to unoccupied states over the entire volume of the Brillouin zone. The basic challenge consists in solving the inverse problem: recovery of the spectrum of elementary transitions and determination of parameters of each of them (the energy E_i and the half-width Γ_i of the band and the corresponding oscillator strength f_i) on the basis of the integral curve. Theoretical consideration of the dielectric constant as a result of summation of contributions of N individual Lorentzian oscillators with $3N$ parameters is generally accepted. For any crystal, the total number of these parameters is extremely large. Therefore, it may appear that the problem of decomposing the integral curve for ε_2 into elementary components is incorrect (uncertain). However, a solution was found on the basis of unique properties of the curve $\varepsilon_2(E) = F(\varepsilon_1(E))$ for a symmetric Lorentzian (the Argand diagram): this curve has the shape of an almost ideal circle whose parameters define uniquely E_i, Γ_i, and f_i. In the case of a crystal, we obtain an intricate Argand diagram owing to superposition of many elementary circles. On the basis of the known spectra of $\varepsilon_2(E)$ and $\varepsilon_1(E)$, we draw the combined Argand diagram, from which the individual bands are separated sequentially, beginning with the most intense and isolated peak of ε_2, with subsequent optimization of this decomposition.

On the basis of the spectra $\varepsilon_2(E)$ and $\varepsilon_1(E)$ obtained by us for the three diamond samples, we resolved the spectrum $\varepsilon_2(E)$ into elementary components and determined the parameters E_i, Γ_i, and f_i. A total of 13 components were found in the range of 6–27 eV, with their energies E_i differing insignificantly for the three diamond samples and the differences in the oscillator strength amounting to ±(5–25)%. For the sake of brevity, we dwell on the data for sample 2 (see table; the intensities of the bands were estimated via f_i with consideration of $n_{ef}(E)$, and also using the areas under the bands S_i and the ratio $S_i/S_{\max}$; $\varepsilon_{2\max}$ is the value of ε_2 at the peak of the band).

The energy bands in diamond have been considered theoretically in many publications [3–13]. However, only one of these [6] contains extensive data on $\varepsilon_2(E)$ and the band diagram, which are most suitable for detailed comparison with the results of experiments. In this paper, the method of nonlocal empirical pseudopotential is used to calculate the energy bands of diamond along many directions in the Brillouin zone and also the spectrum $\varepsilon_2(E)$ in the range of 6–20 eV. Theoretically, the absorption begins with the transitions $\Gamma_{25'} \longrightarrow \Gamma_{2'}$ at 7 eV (a critical point of the type M_0) followed by a large group of transitions that include $\Gamma_{25'} \longrightarrow \Gamma_{15}$ (8.2 eV, M_0), $\Lambda_3 \longrightarrow \Lambda_1$ (8 eV), $\Delta_5 \longrightarrow \Delta_1$ (8.2 eV), and $L_{3'} \longrightarrow L_{2'}$ (8.3 eV, M_1). In the range of 8.4–11.5 eV, the curve $\varepsilon_2(E)$ is formed by numerous transitions from the upper valence band to the lower conduction band (the transitions $4 \longrightarrow 5$) in the large volume of the Brillouin zone in the vicinity of the points K, U, and X; because of the latter, four peaks in the spectrum $\varepsilon_2(E)$ may be observed. It is typical of covalent tetrahedral semiconductors that the most intense peak in the $\varepsilon_2(E)$ spectrum is related to the transitions $\Sigma_2 \longrightarrow \Sigma_3$ (11 eV). For higher energies, the transitions at the points K, X, Λ, Δ, Σ, and Γ become involved. As a rule, for each of the 13 transitions [we established (O_i, i = 1–13)], two or three versions of specific origin of the transitions are possible in accordance with theoretical energy-band diagram [6]. These versions are listed in the table. The origin of these oscillators can be also discussed in terms of the model of metastable excitons. In view of the fact that theoretical calculations of metastable-exciton spectra are lacking, we have to restrict ourselves to the model of band-to-band transitions. In Fig. 2, the arrows indicate the positions of the oscillators O_i, which we calculated, with respect to the peaks of theoretical curve $\varepsilon_2(E)$. It should be emphasized that excellent agreement between the theory [6] and our experimental–calculational data is observed only in regard to the energy structure rather than the transition intensity. It is very important to determine the intensities of individual transitions in succeeding theoretical calculations.

In conclusion, we dwell very briefly on the results of other studies concerned with detailing the spectra of optical functions of diamond in the ranges of 0–20 eV for $n(E)$ and $k(E)$ [27]; 0–25 eV for $\varepsilon_1(E)$, $\varepsilon_2(E)$, $n(E)$, $k(E)$, and $R(E)$ [28]; and 5–200 eV for $R(E)$ [29]. In the above studies, the integral curves were reproduced (rather than resolved) using a large number of adjustable parameters: 16 [27], 11 [28], and 12 [29]. In these

publications, in addition to the use of the model of symmetric Lorentzian oscillators, disputable major simplifications and assumptions were made. For example, one may dispute the assumption concerning the parabolicity of all energy bands in a wide energy range, because this assumption is valid only in a very small neighborhood of the point Γ. The most intense oscillator at 11.8 eV is missing in the set of oscillators chosen in [29]. This once again confirms the advantages of the parameter-free method of combined Argand diagrams in resolving the integrated spectrum $\varepsilon_2(E)$ into elementary components.

In this study, on the basis of four peaks in the reflection spectrum, we found for the first time 13 bands of the most intense transitions in diamond in the range of 6–30 eV, determined their parameters, and suggested a scheme of their specific origin in the model of energy bands. In addition, we obtained the complete sets of optical fundamental functions for three different diamond samples and analyzed the special features of these functions. These results make it possible to analyze the electronic structure of diamond more thoroughly and in more detail and provide a new basis for further refined theoretical calculations of energy bands, excitons, and electronic properties of diamond.

ACKNOWLEDGMENTS

This work was supported by the Center for Basic Natural Science (St. Petersburg University).

REFERENCES

1. S. V. Vavilov, A. A. Gippius, and E. A. Konorova, *Electronic and Optical Properties of Diamond* (Nauka, Moscow, 1985).
2. V. K. Bazhenov, I. M. Vikulin, and A. G. Gontar', Fiz. Tekh. Poluprovodn. (Leningrad) **19**, 1345 (1985) [Sov. Phys. Semicond. **19**, 829 (1985)].
3. V. V. Sobolev, *Intrinsic Energy Levels of Solids Belonging to Group IVA* (Shtiintsa, Kishinev, 1978).
4. A. A. Levin, Fiz. Tekh. Poluprovodn. (Leningrad) **3**, 1864 (1969) [Sov. Phys. Semicond. **3**, 1579 (1969)].
5. F. German, R. L. Kortum, and C. D. Kuglin, Int. J. Quantum Chem., Symp. **1**, 533 (1967).
6. L. A. Hemstreet, C. Y. Fong, and M. L. Cohen, Phys. Rev. B **2**, 2054 (1970).
7. A. R. Lubinsky, D. E. Ellis, and G. A. Painter, Phys. Rev. B **6**, 3950 (1972).
8. H. J. Mattausch, W. Hanke, and G. Strinati, Phys. Rev. B **27**, 3735 (1983).
9. B. Adolph, V. J. Gavrilenko, K. Tenelsen, *et al.*, Phys. Rev. B **53**, 9797 (1996).
10. V. J. Gavrilenko and F. Bechstedt, Phys. Rev. B **55**, 4343 (1997).
11. F. Bechstedt, K. Tenelsen, B. Adolph, and R. Del Sole, Phys. Rev. Lett. **78**, 1528 (1998).
12. L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. B **57**, R9385 (1998).
13. V. V. Sobolev and V. V. Nemoshkalenko, *The Methods of Computational Physics in the Theory of Solid State: The Electronic Structure of Semiconductors* (Naukova Dumka, Kiev, 1988).
14. V. V. Sobolev, Zh. Prikl. Spektrosk. **63**, 143 (1993).
15. V. V. Sobolev, S. A. Alekseeva, and V. L. Gorenberg, Fiz. Tekh. Poluprovodn. (Leningrad) **12**, 2110 (1978) [Sov. Phys. Semicond. **12**, 1258 (1978)].
16. V. V. Sobolev and V. Val. Sobolev, Fiz. Tverd. Tela (St. Petersburg) **36**, 2560 (1994) [Phys. Solid State **36**, 1393 (1994)].
17. V. V. Sobolev, V. I. Donetskikh, and E. F. Zagaĭnov, in *Proceedings of the International Conference "Excitons in Semiconductors," Leningrad, 1977*, p. 46.
18. E. L. Busygina and V. V. Sobolev, in *Abstracts of the International Conference "Optics of Excitons in Condensed Matter," St. Petersburg, 1997*, p. 70.
19. V. V. Sobolev, in *Proceedings of the International Conference "Optics of Semiconductors," Ul'yanovsk, 1998*, p. 3.
20. C. D. Clark, P. J. Dean, and P. V. Harris, Proc. R. Soc. London, Ser. A **277**, 312 (1964).
21. H. R. Philipp and E. A. Taft, Phys. Rev. **127**, 159 (1962).
22. W. C. Walker and J. Osantowsky, Phys. Rev. A **134**, 153 (1964).
23. H. R. Philipp and E. A. Taft, Phys. Rev. A **136**, 1445 (1964).
24. R. A. Roberts and W. C. Walker, Phys. Rev. **161**, 730 (1967).
25. S. Logothetidis, J. Petalas, H. M. Polatoglou, and D. Fuchs, Phys. Rev. B **46**, 4483 (1992).
26. R. A. Roberts, D. M. Roessler, and W. C. Walker, Phys. Rev. Lett. **17**, 302 (1966).
27. A. R. Forouhi and J. Bloomer, Phys. Rev. B **38**, 1865 (1988).
28. A. D. Papadopoulos and E. Anastassakis, Phys. Rev. B **43**, 5090 (1991).
29. J. Nithianandam and J. C. Rife, Phys. Rev. B **47**, 3517 (1993).

Translated by A. Spitsyn

Semiconductors, Vol. 34, No. 8, 2000, pp. 908–915. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 947–954.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Mutual Electron–Phonon Drag and Low-Temperature Anomalies of Thermoelectric and Thermomagnetic Effects in HgSe:Fe Crystals

I. G. Kuleev* and I. Yu. Arapova

Institute of Metal Physics, Ural Division, Russian Academy of Sciences, ul. S. Kovalevskoĭ 18, Yekaterinburg, 620219 Russia
** e-mail: kuleev@imp.uran.ru*

Submitted January 24, 2000; accepted for publication February 17, 2000

Abstract—Various approaches to solving a set of kinetic equations describing the behavior of a nonequilibrium electron–phonon system in a magnetic field are considered. The effect of the mutual electron–phonon drag on the magnetic-field dependence of the Nernst–Ettingshausen coefficients is analyzed. A detailed analysis of the dependence of the diffusion contribution to the thermoelectric power on the Fe content in HgSe:Fe crystals is carried out taking into account the mutual electron–phonon drag. The kinetic coefficients corresponding to these solutions are calculated for conductors with degenerate carrier statistics. © *2000 MAIK "Nauka/Interperiodica".*

Experimental studies of the dependence of thermoelectric power on magnetic field [1] have revealed its unusual behavior in HgSe and HgSe:Fe crystals with a Fe concentration of $N_{Fe} < 1 \times 10^{19}$ cm^{-3} in the low temperature range $20 < T < 60$ K. The longitudinal Nernst–Ettingshausen (NE) coefficient characterizing changes in the thermoelectric power on applying a magnetic field, $\Delta\alpha(H) = \alpha(H) - \alpha(0)$, first grows quadratically with increasing H at $\Omega\tau < 1$ ($\Omega = eH/mc$ is the cyclotron frequency and τ is the electron relaxation time), passes through a maximum for some $H = H_m$, and then decreases as the field increases further. Such behavior contradicts the conventional viewpoint [2, 3] that the thermoelectric power of a degenerate electron gas levels off in classically strong magnetic fields (at $\Omega\tau \gg 1$) and is independent of the carrier-scattering mechanism. Another unusual fact is the sign reversal of the transverse NE coefficient $Q_\perp(H)$ with a magnetic field increasing in the range of $\Omega\tau > 1$ [4]. An attempt was made [5] to explain the unusual dependences of the thermomagnetic coefficients observed experimentally for HgSe and HgSe:Fe crystals by the mutual electron–phonon drag.

Experimental studies of thermoelectric power in HgSe:Fe crystals with varied Fe content have shown that the dependence $\alpha(N_{Fe})$ is nonmonotonic at $T < 10$ K [6]. The absolute value of $|\alpha(N_{Fe})|$ first decreases, then increases with growing concentration N_{Fe}, and attains a maximum at $N_{Fe} = (1–2) \times 10^{19}$ cm^{-3}. With the Fe concentration increasing further, the thermoelectric power decreases steadily. Such a dependence is qualitatively similar to the dependence of the electron mobility on the Fe impurity content $\mu(N_{Fe})$ [7]. Since in HgSe crystals at $T < 10$ K the absolute value of the diffusion component of the thermoelectric power α_{dif} is an order of magnitude smaller than the phonon component associated with the electron–phonon drag, the authors of [6] related the anomalous rise in thermoelectric power to a change in the phonon contribution. The rise in thermoelectric power in the concentration range $5 \times 10^{18} < N_{Fe} < (1–2) \times 10^{19}$ cm^{-3} was shown to result from the reduction of the Rayleigh scattering of phonons by a spatially correlated Fe ion array (SCIA), just as the anomalous mobility rise in the same concentration range is a result of the reduced phonon scattering by the Fe^{3+} SCIA [7]. The effect of the mutual electron–phonon drag was not considered in [6]. The authors of [5] believe that the anomalous rise in $|\alpha(N_{Fe})|$ is related to the diffusion component of the thermoelectric power and results from the mutual electron–phonon drag. However, no detailed analysis of formulas confirming this conclusion was presented, with only a qualitative explanation given.

Prior to presenting the results of our calculation, we note the following:

(I) The method for solving the set of kinetic equations for nonequilibrium electron and phonon distribution functions, proposed in [8] and used in [5], makes it possible to consider correctly the mutual drag effect only in the zero approximation in the electron gas degeneracy parameter (k_BT/ζ). In this approximation, the diffusion fluxes and both the NE coefficients vanish. Using this method to solve the kinetic equations provides no means of correctly taking into account the contribution of mutual electron–phonon drag to the diffusion component of the charge and heat fluxes. This circumstance leads to incorrect results for both the longitudinal and the transverse NE coefficients. We

present results that are exact in the linear approximation in the degeneracy parameter [9]. As an additional argument, we consider the solution of a set of kinetic equations for a nonequilibrium electron–phonon system in terms of the perturbation theory in a linear approximation in the mutual-drag parameter.

(II) In order to study the low-temperature anomalies of the dependence $|\alpha(N_{\mathrm{Fe}})|$, we analyze in detail the diffusion component of thermoelectric power taking into account the mutual electron–phonon drag. The diffusion contribution to the thermoelectric power $|\alpha_{\mathrm{dif}}(N_{\mathrm{Fe}})|$ will be shown to be a decreasing function of N_{Fe} in the concentration range $5 \times 10^{18} < N_{\mathrm{Fe}} < (1\text{–}2) \times 10^{19}$ cm^{-3} and, therefore, cannot account for the experimentally observed rise in thermoelectric power in this concentration range.

In Section 1, various methods for solving the set of kinetic equations for nonequilibrium electron–phonon systems in a magnetic field are considered: (i) the Gurevich–Korenblit method [8] and (ii) the method based on the perturbation theory for the mutual-drag parameter. In Section 2, the effect of mutual drag on the magnetic-field dependence of the NE coefficients is analyzed. In Section 3, the low-temperature anomaly of $|\alpha(N_{\mathrm{Fe}})|$ in HgSe:Fe crystals is considered taking into account the mutual electron–phonon drag.

1. METHODS FOR SOLVING A SET OF KINETIC EQUATIONS FOR A NONEQUILIBRIUM ELECTRON–PHONON SYSTEM IN A MAGNETIC FIELD

The initial set of kinetic equations for the nonequilibrium electron $f(\mathbf{k}, \mathbf{r})$ and phonon $N^{\lambda}(\mathbf{q}, \mathbf{r})$ distribution functions for a degenerate conductor in a magnetic field is written as [2, 8, 9]

$$\mathbf{v}_k \nabla_{\mathbf{r}} f_{\mathbf{k}} + \frac{e}{\hbar}\left(\mathbf{E}_0 + \frac{1}{c}[\mathbf{v}_{\mathbf{k}} \times \mathbf{H}]\right)\nabla_k f(\mathbf{k}, \mathbf{r}) = I_{ei}(f_{\mathbf{k}}) + I_{eph}(f_{\mathbf{k}}, N^{\lambda}_{\mathbf{q}}), \quad (1)$$

$$v^{\lambda}_{\mathbf{q}} \nabla_{\mathbf{r}} N^{\lambda}_{\mathbf{q}} = -(N^{\lambda}_{\mathbf{q}} - N^{0}_{\mathbf{q}})\nu^{(l)\lambda}_{ph} + I_{phe}(N^{\lambda}_{\mathbf{q}}, f_{\mathbf{k}}),$$

where $\mathbf{v}_{\mathbf{k}} = \partial\varepsilon_{\mathbf{k}}/\hbar\partial\mathbf{k}$, $\mathbf{v}^{\lambda}_{\mathbf{q}} = \partial\omega^{\lambda}_{\mathbf{q}}/\partial\mathbf{q}$ is the group velocity of phonons with polarization λ, $N^{0}_{\mathbf{q}\lambda}$ is the Planck function, and the relaxation frequency $\nu^{(l)\lambda}(\mathbf{q})$ accounts for all nonelectronic phonon-scattering mechanisms: phonon–phonon scattering and phonon scattering by defects and sample boundaries. The collision integrals for scattering of electrons by impurities I_{ei} and phonons I_{eph}, and phonons by electrons I_{phe}, were calculated elsewhere [2, 8, 10]. We note that initial set (1) is valid if the frequencies of electron and phonon relaxation in normal scattering, $\nu^{N}_{ee}(\varepsilon)$ and $\nu_{ph\text{–}ph}(\varepsilon)$, are low compared with the corresponding electron-transport relaxation frequencies $\nu_{eR}(\varepsilon) = \nu_{ei}(\varepsilon) + \nu_{eph}(\varepsilon)$ and $\nu_{phR}(\mathbf{q}) = \nu_{phL}(\mathbf{q}) + \nu_{phi}(\mathbf{q}) + \nu^{U}_{ph-ph}(\mathbf{q})$. Here, $\nu_{ei}(\varepsilon)$ and $\nu_{eph}(\varepsilon)$ are the frequencies of electron relaxation via scattering by impurities and phonons, respectively, and $\nu_{phL}(\mathbf{q})$, $\nu_{phi}(\mathbf{q})$, and $\nu^{U}_{ph-ph}(\mathbf{q})$ are the frequencies of phonon relaxation via scattering by boundaries and defects (the Rayleigh mechanism) and via the Umklapp processes, respectively.

Let us represent the electron and phonon distribution functions as

$$f_{\mathbf{k}} = f_0(\varepsilon_k) + \delta f_{\mathbf{k}}, \quad N^{\lambda}_{\mathbf{q}} = N^{0}_{\mathbf{q}\lambda} + g_{\lambda}(\mathbf{q}), \quad (2)$$

where $f_0(\varepsilon_{\mathbf{k}})$ and $N^{0}_{\mathbf{q}}\lambda$ are the locally equilibrium distribution functions for electrons and phonons, respectively; while $\delta f_{\mathbf{k}}$ and $g_{\lambda}(\mathbf{q})$ are the nonequilibrium additions to the distribution functions, linear in external action. We linearize the collision integrals in terms of external action. The collision integrals $I_{ei}(\delta f_{\mathbf{k}})$, $I_{eph}(f_0, g_{\lambda}(\mathbf{q}))$, and $I_{eph}(\delta f_{\mathbf{k}}, N^{0}_{\mathbf{q}\lambda})$ can be expressed in terms of the relaxation frequencies [9–11] in the elastic scattering approximation. We take into account the inelastic electron–phonon scattering to a first order in the inelasticity parameter $\hbar\omega_{q\lambda}/\zeta$ in calculating the collision integral $I_{eph}(f_0, g_{\lambda}(\mathbf{q}))$.

Let us represent the electron distribution function $\delta f_{\mathbf{k}}$ as

$$\delta f_{\mathbf{k}} = \left(-\frac{\delta f_0}{\delta\varepsilon}\right)(\mathbf{v}_k \boldsymbol{\chi}(\varepsilon)). \quad (3)$$

Then, we obtain a closed integral equation for the function $\boldsymbol{\chi}(\varepsilon)$. This equation accounts for both the effect of the nonequilibrium phonons on electrons and that of the nonequilibrium electrons on electrons via the phonon subsystem (mutual electron–phonon drag); i.e., we have

$$\boldsymbol{\chi}(\varepsilon) = \boldsymbol{\chi}^{(1)}(\varepsilon) + \gamma(\varepsilon)[\mathbf{h} \times \boldsymbol{\chi}(\varepsilon)] + \boldsymbol{\chi}^{(2)}(\varepsilon), \quad \mathbf{h} = \mathbf{H}/H, \quad (4)$$

where $\gamma(\varepsilon) = \Omega(\varepsilon)\tau(\varepsilon)$ and $\Omega(\varepsilon) = eH/m(\varepsilon)c$ is the cyclotron frequency. The function $\boldsymbol{\chi}^{(1)}(\varepsilon)$ directly accounts for the effects of the electric field and temperature gradient on the electron subsystem and the electron–phonon drag (the term proportional to $A_{ph}(\varepsilon)$):

$$\boldsymbol{\chi}_1(\varepsilon_{\mathbf{k}}) = -e\tau(\varepsilon_{\mathbf{k}})\left(\mathbf{E} + \frac{k_B}{e}\left(\frac{(\tilde{m}(\varepsilon))^2}{\tilde{k}^3}A_{ph}(\varepsilon) + \frac{\varepsilon_k - \zeta}{k_B T}\right)\nabla T\right),$$

where

$$A_{ph}(\varepsilon) = \sum_\lambda \frac{m_F s_\lambda^2}{k_B T}\left\langle \frac{\nu_{eph}^\lambda(k_F, \mathbf{q})}{\nu_{ph}^\lambda(\mathbf{q})}\right\rangle_{z_{2k}^\lambda} \equiv \sum_\lambda \frac{m_F s_\lambda^2}{k_B T}\int_0^{z_{2k}^\lambda} dz_q^\lambda \frac{\nu_{eph}^\lambda(k_F, \mathbf{q})}{\nu_{ph}^\lambda(\mathbf{q})}, \quad (5)$$

$z_q^\lambda = \frac{\hbar\omega_{q\lambda}}{k_B T} = \frac{q}{q_T^\lambda}$, $q_T^\lambda = \frac{k_B T}{\hbar s_\lambda}$, $z_{2k}^\lambda = \frac{2k}{q_T^\lambda}$, $\tilde{m}(\varepsilon) = \frac{m(\varepsilon)}{m_F}$, $m_F = m(\zeta)$ is the effective electron mass at the Fermi level, $\tilde{k} = k/k_F$, and $\hbar k_F$ is the Fermi momentum. Here, $\tau(\varepsilon)$ is the total electron-relaxation time, and the function $A_{ph}(\varepsilon)$ depends on the energy ε only via the upper integration limit $z_{2k(\varepsilon)}^\lambda$:

$$\boldsymbol{\chi}^{(2)}(\varepsilon) = \frac{\tilde{m}^2(\varepsilon)\tau(\varepsilon)}{\tilde{k}^3}\mathbf{Q}(\varepsilon) = \frac{\tilde{m}^2(\varepsilon)\tau(\varepsilon)}{\tilde{k}^3}\sum_\lambda \int_0^{z_{2k}^\lambda} dz_q^\lambda \frac{\nu_{phe}^\lambda(k_F, q)\nu_{eph}^\lambda(k_F, q)}{\nu_{ph}^\lambda(q)} \times \int_{\varepsilon_{q/2}}^\infty d\varepsilon'\left(-\frac{\partial f_0}{\partial \varepsilon'}\right)\tilde{m}(\varepsilon')\boldsymbol{\chi}(\varepsilon'). \quad (6)$$

Equation (4) can be solved for the function $\boldsymbol{\chi}(\varepsilon)$:

$$\boldsymbol{\chi}(\varepsilon) = \boldsymbol{\chi}_{1H}(\varepsilon) + \boldsymbol{\chi}_{2H}(\varepsilon),$$

$$\boldsymbol{\chi}_{1H}(\varepsilon) = \{\boldsymbol{\chi}_1(\varepsilon) + \gamma(\varepsilon)[\mathbf{h}\times\boldsymbol{\chi}_1(\varepsilon)]\}(1+\gamma^2(\varepsilon))^{-1},$$
$$\boldsymbol{\chi}_{2H}(\varepsilon) = \frac{\tilde{m}^3(\varepsilon)\tau(\varepsilon)}{\tilde{k}^3(1+\gamma^2(\varepsilon))}\mathbf{Q}_H(\varepsilon), \quad (7)$$
$$\mathbf{Q}_H(\varepsilon) = \mathbf{Q}(\varepsilon) + \gamma(\varepsilon)[\mathbf{h}\times\mathbf{Q}(\varepsilon)].$$

Integral equation (4)–(6) was solved in [12, 13] for the case of a nondegenerate electron gas, and in [5, 8, 14], for conductors with degenerate carrier statistics. The kernel of this integral equation is rather cumbersome: the sought-for function $\boldsymbol{\chi}(\varepsilon)$ is under the double integral sign, and the dependences of the relaxation frequencies on the phonon wave vector [7] are to be specified to solve the equation.

The integral equation (4)–(6) can be transformed into an inhomogeneous Volterra integral equation with the following integral term [9–11]:

$$\mathbf{Q}(\varepsilon) = \Phi(\varepsilon)\int_\varepsilon^\infty d\varepsilon'\left(-\frac{\partial f_0}{\partial\varepsilon'}\right)\tilde{m}(\varepsilon')\boldsymbol{\chi}(\varepsilon') + \int_0^\varepsilon d\varepsilon'\left(-\frac{\partial f_0}{\partial\varepsilon'}\right)\tilde{m}(\varepsilon')\Phi(\varepsilon')\boldsymbol{\chi}(\varepsilon'), \quad (8)$$

$$\Phi(\varepsilon) = \sum_\lambda\left\langle\frac{\nu_{phe}^\lambda(k_F, \mathbf{q})\nu_{eph}^\lambda(k_F, \mathbf{q})}{\nu_{ph}^\lambda(\mathbf{q})}\right\rangle_{z_{2k}^\lambda}.$$

Here, $\Phi(\varepsilon)$ depends on energy ε via the upper limit of integration $z_{2k(\varepsilon)}^\lambda$. The reciprocal quantity $\Phi(\varepsilon)^{-1}$ characterizes the time $\tau_{e\text{–}ph\text{–}e}$ during which the momentum transferred by electrons to the phonon subsystem is regained by electrons [10]. The integral equation (7) with integral term (8) is not only simpler (the sought-for function is under a single integral sign). It allows the construction of a regular solving procedure without specifying the dependence of the relaxation frequencies on a phonon wave vector, with only the strong-degeneracy condition ($k_B T/\zeta \ll 1$) used. Such a solution was found elsewhere [9].

The integral equation was solved by the Gurevich–Korenblit method (see [5, 8]) by substituting $\delta(\varepsilon - \zeta)$ for $(-\partial f_0/\partial\varepsilon)$ in (6) with the integral equation transformed into an algebraic equation for the function $\boldsymbol{\chi}(\zeta)$. The latter equation could be readily solved. As shown below, this solution is valid only in zero order in degeneracy, and the authors of [8] restricted themselves to the consideration of galvanomagnetic and thermoelectric effects in just this approximation. Since the influence of the mutual electron–phonon drag on diffusion fluxes is not taken into account in this approximation, it cannot be used to analyze the effect of the mutual electron–phonon drag on thermoelectric effects: the diffusion fluxes vanish, together with the NE coefficients, in the zero approximation in $(k_B T/\zeta)$. When analyzing thermoelectric effects, the authors of [5] went beyond the applicability limits of method [8]: first, a solution for $\boldsymbol{\chi}(\zeta)$ was found in zero approximation in degeneracy and, then, $\boldsymbol{\chi}_{2H}^{(0)}(\varepsilon)$ was determined using this solution. This enabled them to write solution (7) in the form

$$\boldsymbol{\chi}(\varepsilon) = \boldsymbol{\chi}_{1H}(\varepsilon) + \boldsymbol{\chi}_{2H}^{(0)}(\varepsilon), \quad (9)$$

where

$$\boldsymbol{\chi}_{2H}^{(0)}(\varepsilon) = \frac{\tau(\varepsilon)}{1+\gamma^2(\varepsilon)}\{\mathbf{Q}^{(0)}(\zeta) + \gamma(\varepsilon)[\mathbf{h}\times\mathbf{Q}^{(0)}(\zeta)]\},$$
$$\mathbf{Q}^{(0)}(\zeta) = -\frac{e\Gamma}{(1-\Gamma)^2+\gamma_F^2} \times\{\mathbf{E}_A(1-\Gamma) + \gamma_F[\mathbf{h}\times\mathbf{E}_A]\}, \quad (10)$$
$$\mathbf{E}_A = \mathbf{E} + \frac{k_B}{e}A_{ph}(\zeta)\nabla T.$$

Obviously, in this approach, the diffusion contributions are accounted for only in the function $\boldsymbol{\chi}_{1H}(\varepsilon)$ and are

absent in $\boldsymbol{\chi}_{2H}^{(0)}(\zeta)$. Equations (8) and (9) are written in the notation used in [9–11]: they are more convenient and shorter (cf. (8) and (9) with formulas (11)–(13) for $\boldsymbol{\chi}(\varepsilon)$ in [5]). According to [10, 12], the mutual-drag parameter $\Gamma = \tau_F\Phi(\varepsilon) = \tau_F/\tau_{e\text{–}ph\text{–}e}$ is equal to the ratio of the electron mean free time to the time during which the momentum transferred by electrons to phonons is regained by the electron subsystem.

A solution of the integral equation for the electron distribution function was found in [9–11] in an approximation linear in the degeneracy parameter. However, the obtained NE coefficients differ significantly from those in [5]. Therefore, we consider it necessary to present as an additional argument the solution to the integral equation, obtained using the perturbation theory for the mutual-drag parameter Γ, and to demonstrate in this way what can be taken into account in the Gurevich–Korenblit approximation. A solution to the integral equation (7) in an approximation linear in parameter Γ can be found as follows: the function $\boldsymbol{\chi}_{1H}(\varepsilon)$ is substituted for $\boldsymbol{\chi}(\varepsilon)$ in the integral term (10) for $Q(\varepsilon)$. Then we obtain in an approximation linear in (k_BT/ζ)

$$\boldsymbol{\chi}_{2H}^{(1)}(\zeta) = \frac{\tau_F}{1+\gamma_F^2}\{\mathbf{Q}^{(1)}(\zeta) + \gamma_F[\mathbf{h}\times\mathbf{Q}^{(1)}(\zeta)]\},$$

$$\mathbf{Q}^{(1)}(\zeta) = \Phi(\zeta)\int_\zeta^\infty d\varepsilon\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\tilde{m}(\varepsilon)\boldsymbol{\chi}_{1H}(\varepsilon) + \int_0^\zeta d\varepsilon\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\tilde{m}(\varepsilon)\Phi(\varepsilon)\boldsymbol{\chi}_{1H}(\varepsilon) = -\frac{e\Gamma}{1+\gamma_F^2}\{\mathbf{E}_{A1} + \gamma_F[\mathbf{h}\times\mathbf{E}_{A2}]\}, \quad (11)$$

where

$$\mathbf{E}_{AN} = \mathbf{E}_A(1-\ln(2)D_\Phi) + \frac{k_B\pi^2}{e}\frac{}{3}D_{QN}\nabla T,$$

$$D_\Phi = k_BT\frac{d}{d\varepsilon}[\ln(\Phi[\varepsilon])]_{\varepsilon=\zeta}, \quad (12)$$

$$D_{QN} = k_BT\frac{d}{d\varepsilon}\left[\ln\left(\frac{\gamma(\varepsilon)^{N-1}m(\varepsilon)\tau(\varepsilon)\Phi^{1/2}(\varepsilon)}{1+\gamma^2(\varepsilon)}\right)\right]_{\varepsilon=\zeta}.$$

Comparison of (11) and (9) shows that there is no diffusion contribution to (9). It is easy to verify that formulas (10) and (11) for $\boldsymbol{\chi}_{2H}(\zeta)$ coincide in zero approximation in (k_BT/ζ) and linear approximation in the mutual-drag parameter Γ. The exact solution of (7) is rather cumbersome [9], and we do not present it here; instead, we give formulas for the kinetic and NE coefficients following from this solution.

2. EFFECT OF MUTUAL ELECTRON–PHONON DRAG ON THE MAGNETIC-FIELD DEPENDENCE OF THE LONGITUDINAL AND TRANSVERSE NERNST–ETTINGSHAUSEN COEFFICIENTS

The magnetic-field dependences of the transverse $Q_\perp(H)$ and longitudinal $\Delta\alpha(H)$ NE coefficients are given by [2]:

$$Q_\perp = -\frac{E_y}{H\nabla_xT} = \frac{\sigma_{yx}\beta_{xx} - \sigma_{xx}\beta_{yx}}{H(\sigma_{xx}^2+\sigma_{yx}^2)},$$
$$\alpha(H) = \frac{\sigma_{xx}\beta_{xx} + \sigma_{yx}\beta_{yx}}{\sigma_{xx}^2+\sigma_{yx}^2}. \quad (13)$$

To analyze the NE effects under isothermal conditions ($j_x = j_y = 0$, $\nabla_yT = 0$), it suffices to calculate the conduction current

$$\mathbf{j}_\mu = -\frac{e^2n}{m_F}\int_0^\infty d\varepsilon\left(-\frac{\partial f_0}{\partial\varepsilon}\right)\frac{(\tilde{k}(\varepsilon))^3}{\tilde{m}(\varepsilon)}(\boldsymbol{\chi}_{1H}^{(\mu)}(\varepsilon) + \boldsymbol{\chi}_{2H}^{(\mu)}(\varepsilon))$$
$$= \sum_\nu(\sigma_{\mu\nu}E_\nu - \beta_{\mu\nu}\nabla_\nu T).$$

Further, we present expressions for the conductivity tensor components $\sigma_{\mu\nu}$ and for the thermoelectric tensor components in three different approximations. The rigorous solution within the linear approximation in (k_BT/ζ) is written as [9]

$$\sigma_{xx} = \tilde{\sigma}_{xx}^0\left\{1 - \frac{C_1\Gamma D_\Phi(1-\tilde{\gamma}_F^2)}{(1-\Gamma)(1+\tilde{\gamma}_F^2)}\right\},$$

$$\sigma_{yx} = -\sigma_{xy} = \tilde{\sigma}_{xx}^0\tilde{\gamma}_F\left\{1 - \frac{2C_1\Gamma D_\Phi}{(1-\Gamma)(1+\tilde{\gamma}_F^2)}\right\},$$

$$\beta_{xx} = -\frac{k_B}{e}\left\{\sigma_{xx}A_{ph}(\zeta) + \frac{\pi^2}{3}\sigma_{xx}^0 \times\left[D_{j1} + \frac{\Gamma[(1-\Gamma)D_{Q1} - \gamma_F^2D_{Q2}]}{(1-\Gamma)^2+\gamma_F^2}\right]\right\}, \quad (14a)$$

$$\beta_{yx} = -\beta_{xy} = -\frac{k_B}{e}\left\{\sigma_{yx}A_{ph}(\zeta) + \frac{\pi^2}{3}\sigma_{yx}^0 \times\left[D_{j2} + \frac{\Gamma[D_{Q1} + (1-\Gamma)D_{Q2}]}{(1-\Gamma)^2+\gamma_F^2}\right]\right\},$$

where $\tilde{\sigma}_{xx}^0 = e^2n_e\tilde{\tau}_F/m_F(1+\tilde{\gamma}_F^2)$, $\tilde{\sigma}_{yx}^0 = \tilde{\gamma}_F\tilde{\sigma}_{xx}^0$, $C_1 = J_1 + \ln(2) \sim 1$, $\tilde{\gamma}_F = \Omega\tilde{\tau}_F$, and $\tilde{\tau}_F = \tau_F/(1-\Gamma)$.

Using the Gurevich–Korenblit approach [8], we can write these coefficients in the notation adopted in [9–11]:

$$\sigma_{xx} = \tilde{\sigma}_{xx}^{(0)} = \frac{e^2 n\tilde{\tau}}{m_F(1+\tilde{\gamma}_F^2)} = \frac{\tilde{\sigma}_0}{1+\tilde{\gamma}_F^2},$$

$$\sigma_{yx} = \tilde{\sigma}_{yx}^{(0)} = \frac{e^2 n\tilde{\tau}\tilde{\gamma}_F}{m_F(1+\tilde{\gamma}_F^2)} = \tilde{\gamma}_F\tilde{\sigma}_{xx}^0,$$

$$\beta_{xx} = -\frac{k_B}{e}\left\{\tilde{\sigma}_{xx}^0 A_{ph}(\zeta) + \frac{\pi^2}{3}\sigma_{xx}^0 D_{j1}\right\}, \tag{14b}$$

$$\beta_{yx} = -\frac{k_B}{e}\left\{\tilde{\sigma}_{yx}^{(0)} A_{ph}(\zeta) + \frac{\pi^2}{3}\sigma_{yx}^0 D_{j2}\right\}.$$

The perturbation theory in parameter Γ yields

$$\sigma_{xx} = \tilde{\sigma}_{xx}^0\left\{1 + \frac{\Gamma[1-D_\Phi\ln(2)](1-\tilde{\gamma}_F^2)}{(1+\tilde{\gamma}_F^2)}\right\},$$

$$\sigma_{yx} = \tilde{\sigma}_{yx}^0\left\{1 + \frac{2\Gamma[1-D_\Phi\ln(2)]}{(1+\tilde{\gamma}_F^2)}\right\},$$

$$\beta_{xx} = -\frac{k_B}{e}\left\{\sigma_{xx}A_{ph}(\zeta) + \frac{\pi^2}{3}\sigma_{xx}^0 \times\left[D_{j1} + \frac{\Gamma}{1+\gamma_F^2}[D_{Q1} - \gamma_F^2 D_{Q2}]\right]\right\}, \tag{14c}$$

$$\beta_{yx} = -\frac{k_B}{e}\left\{\sigma_{yx}A_{ph}(\zeta) + \frac{\pi^2}{3}\sigma_{yx}^0 \times\left[D_{j2} + \frac{\Gamma}{1+\gamma_F^2}[D_{Q1} + D_{Q2}]\right]\right\}.$$

Comparison of (14a), (14b), and (14c) for the thermoelectric coefficients $\beta_{\mu\nu}$ shows that there are no diffusion terms proportional to the mutual-drag parameter Γ in (14b). As noted above, the Gurevich–Korenblit method fails to take into account correctly the influence of the mutual electron–phonon drag on the diffusion fluxes.

As seen from (14a) and (14c), the renormalization of diffusion terms in rigorously solving the set of kinetic equations (14a) for a nonequilibrium electron–phonon system has a significantly more intricate form than (14c). However, formulas (14a) and (14c) yield identical expressions for the transverse $Q_\perp(H)$ and longitudinal $\Delta\alpha(H)$ NE coefficients:

$$Q_\perp(H) = -\frac{\pi^2 k_B}{3eH}\frac{\gamma_F}{1+\gamma_F^2}\{D_{j1} - D_{j2} + \Gamma(D_{j2} - D_{Q2})\},$$

$$\alpha(H) = -\frac{k_B}{e}\left\{A_{ph}(\zeta) + \frac{\pi^2}{3(1+\gamma_F^2)} \times[D_{j1} + \gamma_F^2 D_{j2} + \Gamma(D_{Q1} - D_{j1})]\right\}. \tag{15}$$

This follows from the fact that, under isothermal conditions, the coefficients $Q_\perp(H)$ and $\Delta\alpha(H)$ are found for $\mathbf{j} = 0$. Apparently, there must exist a simpler, compared with that proposed in [9], method for calculating these coefficients. The first three terms in equation (15) for $\alpha(H)$ are given by the conventional theory [2]; the fourth term in curly brackets, proportional to the parameter Γ, is due to the mutual influence of the nonequilibrium states of electrons and phonons. The origin of the additional contribution to the diffusion component of the thermoelectric power is physically clear too. Since the mean velocity of the regular motion of electrons is zero ($\mathbf{j} = 0$), the transfer of momentum from electrons to the phonon subsystem occurs because of the dependences of the effective mass and quasimomentum of electrons and the scattering parameters on the electron energy in the vicinity of the Fermi level. Therefore, this contribution must be proportional to derivatives of corresponding parameters with respect to the electron energy. This follows from (15) too. As can be seen from (15), consideration of the mutual electron–phonon drag does not change the magnetic-field dependence of the longitudinal and transverse NE coefficients:

$$\Delta\alpha(H) = \alpha(H) - \alpha(0) = \frac{k_B}{e}\frac{\pi^2\gamma_F^2}{3(1+\gamma_F^2)}[D_0 + \Gamma D_{Qj}],$$

$$Q_\perp(H) = -\frac{\pi^2 k_B}{3eH}\frac{\gamma_F}{1+\gamma_F^2}[D_0 + \Gamma D_{Qj}], \tag{16a}$$

$$\alpha(0) = -\frac{k_B}{e}\left\{A_{ph}(\zeta) + \frac{\pi^2}{3}[D_{j1} + \Gamma D_{Qj}]\right\},$$

where

$$D_0 = k_B T\frac{d}{d\varepsilon}\left[\ln\left(\frac{\tau(\varepsilon)}{m(\varepsilon)}\right)\right]_{\varepsilon=\zeta},$$

$$D_{Qj} = k_B T\frac{d}{d\varepsilon}\left[\ln\left(\frac{(m(\varepsilon))^2\Phi^{1/2}(\varepsilon)}{k^3(\varepsilon)}\right)\right]_{\varepsilon=\zeta}.$$

As can be seen from (16a), the transverse NE coefficient decreases with increasing magnetic field and tends to zero without changing its sign for $H \longrightarrow \infty$. The absolute value of $\Delta\alpha(H)$ grows quadratically with increasing magnetic field at weak fields $\gamma_F \ll 1$ and gradually levels off for $\gamma_F \gg 1$, as it also does in the absence of mutual drag [15]. Such behavior is associated with some special features of the magnetic-field effect on the momentum relaxation in the electron–phonon system (see Fig. 1 in [9]). The expression for the thermoelectric power $\alpha(0)$ in a zero magnetic field coincides with that derived elsewhere [10, 11]. It follows from formulas (14b) that

$$Q_\perp(H) = -\frac{\pi^2 k_B}{3eH}\frac{\gamma_F}{1+\gamma_F^2}[D_{j1} - D_{j2} + \Gamma D_{j2}],$$

$$\alpha(H) = -\frac{k_B}{e}\left\{A_{ph}(\zeta) + \frac{\pi^2}{3(1+\gamma_F^2)}\right. \tag{16b}$$

$$\left.\times [D_{j1} + \gamma_F^2 D_{j2} - \Gamma D_{j1}]\right\}.$$

Expressions (18)–(20) for $\alpha(H) = \alpha_{ph} + \alpha_{\text{dif}}$ in [5] can be transformed to the form (16b) if we change the notation adopted in [5] to that of [10, 11]: $3/4\Pi \longrightarrow A_{ph}(\zeta)$, $\Pi^* \longrightarrow \Gamma$, and $S_p \longrightarrow \Gamma/((1-\Gamma)(1+\tilde{\gamma}_F^2))$. In this case, the following approximation should be used for the mutual-drag parameter Γ

$$\Gamma = \tau_F\Phi(\varepsilon) \approx \tau_F\left\langle\frac{\nu_{eph}(\mathbf{q})\nu_{phe}(\mathbf{q})}{\nu_{ph}(\mathbf{q})}\right\rangle \longrightarrow \tau_F\langle\nu_{eph}(\mathbf{q})\rangle\left\langle\frac{\nu_{phe}(\mathbf{q})}{\nu_{ph}(\mathbf{q})}\right\rangle. \tag{17}$$

Furthermore, we should use the high-temperature approximation for phonons, i.e., $N_q(N_q + 1) \longrightarrow k_BT/\hbar\omega_q$. We then use expression (18) in [5] to obtain

$$\Delta\alpha(H) = -\frac{k_B}{e}\frac{\pi^2 k_B T}{3\zeta}\frac{\gamma_F^2}{1+\gamma_F^2} \times\left\{\frac{3}{2}\Gamma - D_0\left[1 - \Gamma\frac{3+\tilde{\gamma}_F^2}{1+\gamma_F^2}\right]\right\} \tag{18}$$

for the case of a parabolic band $D_0 = \zeta\tau'(\zeta)/\tau(\zeta)$.

A comparison of (16a), (16b), and (18) shows that the NE coefficients (16a) and (18) exhibit different magnetic-field dependences because of the terms due to the mutual electron–phonon drag. Although the dependences $\Delta\alpha(H)$ derived in [5] agree qualitatively with experimental data, this result is a consequence of the fact that the terms linear in the degeneracy parameter are not taken into account correctly in solving the set of kinetic equations for nonequilibrium electron–phonon systems. As mentioned above, the inaccuracy of this method [5, 8] consists in the fact that the diffusion fluxes are calculated with the mutual electron–phonon drag ignored.

Thus, consideration of the mutual drag in the context of a single-parameter description of the electron- and phonon-momentum relaxation by introducing total relaxation times cannot account for the unconventional magnetic-field dependence of the NE coefficients [1, 4].

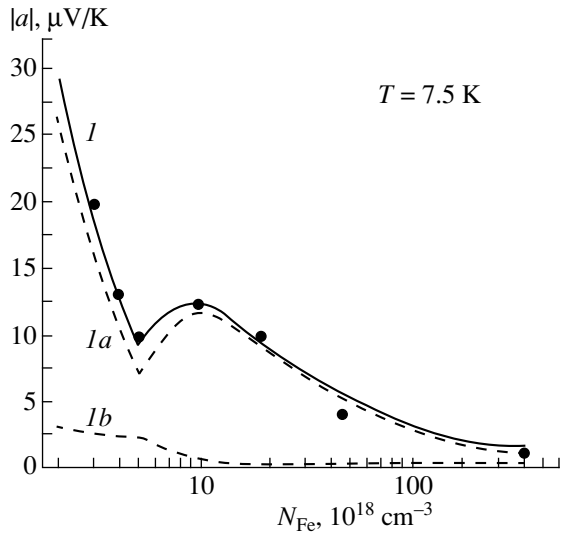
3. ANOMALOUS CONCENTRATION DEPENDENCE OF THERMOELECTRIC POWER IN HgSe:Fe CRYSTALS AT LOW TEMPERATURES

The following equation for the thermoelectric power in a zero magnetic field can be derived from formulas (18)–(20) in [5] (see (16b)):

$$\alpha(0) = \alpha_{ph} + \alpha_{\text{dif}} = -\frac{k_B}{e}\left\{A_{ph} + \frac{\pi^2}{3}(1-\Gamma)\frac{k_BT}{\zeta} \times\left[\zeta\frac{3k'}{k} + \zeta\left(\frac{\tau'}{\tau} - \frac{m'}{m}\right)\right]\right\}. \tag{19}$$

It can be seen from (19) that the diffusion contribution to the thermoelectric power is a decreasing function of the Fe impurity concentration in the range of $5\times10^{18} < N_{\text{Fe}} < 2\times10^{19}$ cm^{-3}. This is caused by the following two factors. First, in this concentration range, the electron relaxation time τ_F increases approximately fourfold with increasing degree of the spatial ordering of the Fe^{3+} SCIA [7], while τ_{e-ph-e} remains constant at a fixed temperature. Therefore, the multiplier $(1 - \Gamma) = (1 - \tau_F/\tau_{e-ph-e})$ is a decreasing function of the iron impurity concentration. Second, the expression in the square brackets also decreases with increasing level of doping: for a chaotic system of charged centers, $\zeta\tau'(\zeta)/\tau(\zeta) = 3/2$. With increasing spatial ordering of the Fe^{3+} SCIA, this quantity decreases, vanishing at $N_{\text{Fe}} = 8\times10^{18}$ cm^{-3} (see [15]) and becoming negative with the Fe impurity concentration increasing further. Thus, $\alpha_{\text{dif}}(N_{\text{Fe}})$ decreases with growing N_{Fe}, whereas the experimental data show that the total thermoelectric power $|\alpha(N_{\text{Fe}})| = |\alpha_{ph}(N_{\text{Fe}}) + \alpha_{\text{dif}}(N_{\text{Fe}})|$ increases in this concentration range. Therefore, the anomalous rise in the thermoelectric power can only be associated with the phonon contribution.

The authors of [6] attributed the anomalous rise in the thermoelectric power to the effect of the reduction of phonon scattering by the spatially correlated array of charged impurities. The dependences of the thermoelectric power on temperature and Fe impurity content were quantitatively analyzed [6] taking into account the electron–phonon drag. The analysis allowed for electron scattering by the Fe^{3+} SCIA, alloy composition



The dependences of (*1a*) phonon, (*1b*) diffusion, and (*1*) total thermoelectric power on the content of Fe impurities at $T = 7.5$ K. The dots in the plot correspond to experimental data. The phonon-scattering parameters are the same as in [6].

fluctuations, and acoustic phonons, and also all basic mechanisms of phonon scattering. The characteristic time of phonon relaxation via scattering by the Fe^{3+} SCIA was calculated. The calculated dependences of the thermoelectric power on the Fe impurity content and temperature were in good agreement with experimental data. However, the effect of the mutual drag was not taken into account [6]. The expression for the thermoelectric power, taking into account this effect, is written as

$$\alpha(0) = -\frac{k_B}{e}\left\{A_{ph} + \frac{\pi^2}{3}\frac{k_B T}{\zeta}\left((1-\Gamma)\frac{3k'}{k} + \frac{\tau'}{\tau} - (1-2\Gamma)\frac{m'}{m} + \frac{\Gamma}{2}\frac{\Phi'}{\Phi}\right)\right\}. \quad (20)$$

The calculated dependences $|\alpha(N_{Fe})|$ and the contributions $|\alpha_{ph}(N_{Fe})|$, $|\alpha_{dif}(N_{Fe})|$ at $T = 7.5$ K at the same values of the parameter as in [6] are shown in the figure. These values are consistent with those obtained in [16] on the basis of a detailed study of the thermal conductivity of HgSe crystals. As can be seen from the figure, if the mutual electron–phonon drag is taken into account, the diffusion contribution to the thermoelectric power $|\alpha_{dif}(N_{Fe})|$ is a decreasing function of the Fe impurity content within the concentration range of $5 \times 10^{18} < N_{Fe} < 2 \times 10^{19}$ cm^{-3}. Obviously, an increase in the thermoelectric power in this range is due only to the phonon contribution and is related to reduction of the phonon scattering by Fe^{3+} SCIA. This occurs in the same way as the anomalous rise in mobility $\mu(N_{Fe})$ in the same concentration range occurs due to the reduction of electron scattering by Fe^{3+} SCIA [7]. It should be noted in this connection that the nature of scattering by an array of iron ions with mixed valence is not the same for electrons and phonons. Electrons obviously readily respond to the charge state of a defect: the probability of their scattering by charged impurities is two orders of magnitude larger than that by neutral impurities. Phonons naturally do not experience the defect charge directly, but the Rayleigh-scattering probability is proportional to the squared size of a perturbed crystal domain induced by a defect [17]. The perturbing effect of a neutral defect is usually localized within a unit cell. For a charged impurity in a semiconductor, the size of the perturbed region is determined by the screening radius. For HgSe:Fe crystals with $n_c = N^* = 4.5 \times 10^{18}$ cm^{-3}, $r_s \approx 5 \times 10^{-7}$ cm $\sim 10a_0$ (a_0 is the lattice constant). Therefore, the perturbing effect of the Fe^{3+} ion (changes in force constants and in lattice distortion) involves a large number of cells rather than a single cell. It was shown [6] that the probability of phonon scattering by Fe^{2+} ions (neutral in the lattice) is an order of magnitude lower than that by Fe^{3+} ions. Being in agreement with the data on the phonon Rayleigh scattering in other semiconductor systems [18, 19], this result is not surprising. An experimental study of the thermal conductivity of doped germanium and silicon samples, reported in [18], confirms this conclusion. The presence of electrically inactive (neutral in the lattice) impurities (C, O, and Ge) moderately reduces the total thermal conductivity. At the same time, electrically active impurities (B, P, Al, Ga, and In) significantly suppress the lattice thermal conductivity of silicon and germanium at high temperatures. This suppression becomes more pronounced with decreasing temperature (the electronic contribution is small in doped samples). At temperatures on the order of 10 K, the thermal conductivity of silicon samples doped with charged impurities having a concentration of $\sim 10^{20}$ cm^{-3} falls by three orders of magnitude [18]. Thus, unlike the electrically active impurities, neutral impurities scatter phonons to a significantly lesser extent in other semiconductor systems too [18, 19].

In addition to phonon scattering by Fe ions with mixed valence Fe^{3+}–Fe^{2+}, phonon scattering by the sample boundaries and electrons and phonon–phonon relaxation processes contribute significantly to the phonon quasi-momentum relaxation. Therefore, the relative increase in thermoelectric power is about 25–30%. Scattering by Fe^{3+} and Fe^{2+} ions is the main quasi-momentum relaxation mechanism for electrons in contrast to phonons. Therefore, electron mobility increases by almost a factor of 4 with the spatial ordering of the Fe^{3+} SCIA increasing in the concentration range of $5 \times 10^{18} < N_{Fe} < 2 \times 10^{19}$ cm^{-3}.

Conclusive evidence that the effect of the reduction of phonon scattering by the spatially correlated array of Fe^{3+} ions does occur in HgSe:Fe crystals was furnished by our study of thermal conductivity [20]. We showed

that the spatial ordering of trivalent Fe ions leads not only to an increase in electron mobility [7] and thermoelectric power [6] but also to a substantial rise in thermal conductivity at fairly low temperatures within the concentration range of $5 \times 10^{18} < N_{Fe} < 2 \times 10^{19}$ cm^{-3}. The electron contribution at $T < 10$ K is about 1–2% of the total thermal conductivity and is an order of magnitude smaller than the experimentally observed rise in $\kappa(N_{Fe})$. Therefore, this rise is due to the phonon contribution only. Thus, we can state with confidence that the main reason for the low-temperature anomaly of thermal conductivity [20] is, as in the case of thermoelectric power [6], the effect of the reduction of the Rayleigh scattering of phonons by the spatially correlated system of Fe^{3+} ions.

4. CONCLUSION

Various methods for solving the set of kinetic equations for a nonequilibrium electron–phonon system in a magnetic field were considered. The kinetic coefficients for conductors with degenerate carrier statistics are calculated. It is shown that the method for the solution proposed in [8] and used in [5] to analyze the NE effects gives no means of correctly taking into account the effect of the mutual electron–phonon drag on the diffusion fluxes. The effect of the mutual electron–phonon drag on the magnetic-field dependence of the NE coefficients is analyzed. It is shown that consideration of the mutual drag in the context of a single-parameter description of the relaxation of electron and phonon momenta by introducing the total relaxation times cannot account for the unconventional field dependence of the NE coefficients.

The dependence of the thermoelectric power on the Fe impurity content in HgSe:Fe crystals was analyzed in detail, taking into account the mutual electron–phonon drag. The diffusion contribution to the thermoelectric power is a decreasing function of N_{Fe} in the concentration range of $5 \times 10^{18} < N_{Fe} < (1–2) \times 10^{19}$ cm^{-3} and cannot be used to explain the low-temperature anomaly of the thermoelectric power in HgSe:Fe crystals. The physical reason for the observed increase in the thermoelectric power in this concentration range is the reduction of the probability of the Rayleigh scattering of phonons by Fe^{3+} SCIA because of the spatial ordering of trivalent Fe ions.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 00-02-16-299.

REFERENCES

1. I. G. Kuleev, I. I. Lyapilin, A. T. Lonchakov, and I. M. Tsidil'kovskiĭ, Zh. Éksp. Teor. Fiz. **106**, 1205 (1994) [JETP **79**, 653 (1994)].
2. V. M. Askerov, *Electronic Transport Phenomena in Semiconductors* (Nauka, Moscow, 1985).
3. I. M. Tsidil'kovskii, *Thermomagnetic Effects in Semiconductors* (Fizmatgiz, Moscow, 1960; Academic, New York, 1962).
4. I. I. Lyapilin and Kh. M. Bikkin, in *Proceedings of the 4th Russia Conference on Physics of Semiconductors, Novosibirsk, 1999*, p. 52.
5. I. I. Lyapilin and Kh. M. Bikkin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33** (6), 701 (1999) [Semiconductors **33**, 648 (1999)].
6. I. G. Kuleev, A. T. Lonchakov, I. Yu. Arapova, and G. I. Kuleev, Zh. Éksp. Teor. Fiz. **114**, 191 (1998) [JETP **87**, 106 (1998)].
7. I. G. Kuleev, I. I. Lyapilin, and I. M. Tsidil'kovskiĭ, Zh. Éksp. Teor. Fiz. **102**, 1652 (1992) [Sov. Phys. JETP **75**, 893 (1992)].
8. L. É. Gurevich and I. Ya. Korenblit, Fiz. Tverd. Tela (Leningrad) **6** (3), 856 (1964) [Sov. Phys. Solid State **6**, 661 (1964)].
9. I. G. Kuleev, Fiz. Met. Metalloved. **89** (2), 29 (2000).
10. I. G. Kuleev, Fiz. Tverd. Tela (St. Petersburg) **41** (10), 1753 (1999) [Phys. Solid State **41**, 1608 (1999)].
11. I. G. Kuleev, Fiz. Met. Metalloved. **87** (6), 5 (1999).
12. J. Appel, Z. Naturforsch. A **12** (5), 410 (1957); Z. Naturforsch. **A 13** (5), 386 (1958).
13. J. E. Parrott, Proc. Phys. Soc. London, Sect. B **70** (6), 590 (1957).
14. I. G. Lang and S. T. Pavlov, Zh. Éksp. Teor. Fiz. **63** (4), 1495 (1972) [Sov. Phys. JETP **36**, 793 (1973)].
15. I. G. Kuleev, I. I. Lyapilin, A. T. Lonchakov, and I. M. Tsidil'kovskiĭ, Zh. Éksp. Teor. Fiz. **103**, 1447 (1993) [JETP **76**, 707 (1993)]; Fiz. Tekh. Poluprovodn. (St. Petersburg) **28**, 937 (1994) [Semiconductors **28**, 544 (1994)].
16. C. R. Whitset, D. A. Nelson, J. G. Broerman, and R. Paxhia, Phys. Rev. B **7**, 4625 (1973).
17. J. W. S. Rayleigh, *The Theory of Sound* (Macmillan, London, 1896; Gostekhizdat, Moscow, 1955), Vol. 2.
18. B. M. Mogilevskiĭ and A. F. Chudnovskiĭ, *Thermal Conductivity of Semiconductors* (Nauka, Moscow, 1972).
19. V. S. Oskotskiĭ and I. A. Smirnov, *Defects in Crystals and Thermal Conductivity* (Nauka, Leningrad, 1972).
20. I. G. Kuleev, A. T. Lonchakov, and I. Yu. Arapova, in *Proceedings of the 4th Russia Conference on Physics of Semiconductors, Novosibirsk, 1999*, p. 50; Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 402 (2000) [Semiconductors **34**, 389 (2000)].

Translated by S. Ktitorov

Semiconductors, Vol. 34, No. 8, 2000, pp. 916–923. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 955–961.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Ionization Energy of Copper in $Hg_{0.8}Cd_{0.2}Te$ Crystals under Conditions of Light and Intermediate Doping

V. V. Bogoboyashchiĭ

Kremenchug State Polytechnical Institute, Kremenchug, 39614 Ukraine

Submitted December 9, 1999; accepted for publication January 24, 2000

Abstract—Resistivity and the Hall effect were studied in copper-doped *p*-$Hg_{0.8}Cd_{0.2}Te$ crystals in the temperature range of 4.2–125 K and the range of Cu concentrations from 2.6×10^{15} to 2×10^{18} cm^{-3}. It is shown that the conventional method for determining the ionization energy of impurities from the slope of the dependence $R_H(T)$ is inapplicable in this case. In order to obtain the correct results, it is necessary to take into account the structure of the impurity band and the screening of the impurity charge with free charge carriers. A simplified model of the impurity band is suggested; this model makes it possible to calculate the ionization energy of acceptors under conditions of light doping and a small degree of compensation. This approach is used to find that ionization energy of copper depends only slightly on the copper concentration at $T = 0$ and is equal to E_A = 7.6 meV for an isolated acceptor, which coincides with the theoretical value. At finite temperatures, the ionization energy of acceptors decreases appreciably as a result of screening. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The behavior of shallow impurity levels in the band gap of a semiconductor under conditions of intermediate doping is of interest in connection with the problem of the metal–insulator concentration transition in doped semiconductors. Origination of metallic conduction is commonly related in this case to overlap of the wave functions of impurity states [1, 2] and to screening of the impurity charge by free charge carriers [3].

In both cases, the criterion for the transition involves the bound-state size; for the latter, the radius of an isolated impurity atom is typically used [1–3]. On the other hand, this size depends on the binding energy, which, as is customarily assumed, decreases rapidly with increasing impurity concentration. For this reason, it was questioned [2] whether it is correct to use the state parameters of isolated impurity in the criterion for transition to metallic conduction.

In this paper, we report the results of studying the electrical conductivity and the Hall effect in copper-doped *p*-$Hg_{0.8}Cd_{0.2}Te$ crystals in order to determine the concentration dependence of the binding energy E_A for holes and verify experimentally the validity of using the radius of the state of an isolated neutral acceptor in the Mott criterion.

Typically, when determining the ionization energy of impurity from the results of measuring the Hall coefficient R_H, this energy is identified with the activation energy for R_H in the region of freeze-out of impurity conductivity, if the degree of compensation K of the crystals is close to unity; if the compensation is slight, the doubled activation energy is used. For example, this method was used [4] to determine the concentration dependence of the ionization energy E_A of copper in $Hg_{0.8}Cd_{0.2}Te$ on the basis of measurements of R_H in the temperature range of 20–50 K. According to the results obtained [4], the value of E_A as a function of $N_A^{1/3}$ decreases linearly from 11.5 meV for $N_A = 0$ to zero for $N_A \approx 10^{17}$ cm^{-3}.

Such a method is undoubtedly valid for extremely low concentrations of impurity, in which case the energy levels of impurity atoms of the same type may be considered identical. At the same time, it is well known [2] that, in actual crystals (which are always partly compensated), the impurity levels are shifted by the value of potential of electric field generated by compensating impurities at the site of the defect. In moderately doped $Hg_{0.8}Cd_{0.2}Te$, the amplitude of fluctuations of this field is comparable with E_A; therefore, the activation energy of R_H is defined by both the value of E_A and the large-scale potential of charged donors. It is clear that, in such a situation, the conventional approach to calculating the value of E_A from the data obtained by measuring R_H would yield erroneous results. Correspondingly, the values of E_A in *p*-$Hg_{0.8}Cd_{0.2}Te$ obtained previously (see, e.g., [4]) for the range of intermediate doping cannot be considered reliable.

On the other hand, the efficiency of using the Hall method for measuring the quantity E_A in the narrow-gap crystals was doubted in general [3]. In this paper, we show that this method yields quite reliable results even for narrow-gap semiconductors, such as $Hg_{0.8}Cd_{0.2}Te$, if the experimental data are properly processed. To this end, it is necessary to separate the con-

1063-7826/00/3408-0916$20.00

tributions of different mechanism of charge transport and to take into account the actual structure of the acceptor band, the screening, and the capture of holes by neutral acceptors; the effect of the above factors is quite appreciable in intermediately doped crystals.

2. EXPERIMENTAL

In the experiments, we used the as-grown samples of high-quality $Hg_{1-x}Cd_xTe$ (x = 0.21–0.22) produced commercially by the method of vertical planar crystallization. Most of the samples contained only the uncontrolled impurities. Several crystals were intentionally doped with indium in the course of the growth in order to obtain the heavily compensated p-type samples.

A control group of samples (one sample from each ingot) was set aside from the selected crystals in order to determine the concentration N_D of uncontrolled donors. The remaining samples were subjected to heat treatment in Hg vapors with the aim of eliminating the Hg vacancies and the inclusions of the second phase. After the composition of the heat-treated crystals had been measured, copper was introduced into these crystals from a vacuum-deposited Cu film of specified thickness using a diffusion annealing in saturated Hg vapors at 200°C. The conditions of the annealing were determined from the previously published data [5] and ensured the complete dissolution of the film and macrouniform distribution of Cu in the crystal.

The Cu concentration in the crystals was in the range from 2.6×10^{15} to 2×10^{18} cm^{-3}. This concentration was determined from the value of R_H in a magnetic field of B = 1 T at T = 77 K, in which case copper is completely ionized [5] and the Hall factor for holes is given by r_H = 1.24 [6]. The concentration of Hg vacancies in the crystals belonging to the control group was determined optically [6] at T = 293 K, in which case these vacancies were completely (doubly) ionized [7]. The concentration of In in the doped crystals was varied from 5×10^{15} to 10^{17} cm^{-3}; it was identified with the concentration of extrinsic electrons in these crystals after a low-temperature annealing in the Hg vapors.

Contacts of In were soldered (in the atmosphere of cooled nitrogen vapors) to the samples that had been freshly etched and rinsed in a warm deionized water. Immediately after that, the samples were put in a helium-filled cryostat to measure the quantities ρ and R_H in the temperature range of 4.2–125 K.

Some of the results of measurements are shown in Figs. 1 and 2. It can be seen that, in the Cu-doped crystals with $N_A > 10^{17}$ cm^{-3}, the ε_1-conductivity caused by holes from the valence band is poorly pronounced and is comparable in magnitude with conductivity over the impurity band, so that it is practically impossible to separate the former conductivity in the uncombined form (Fig. 1).

In all crystals without exception, conductivity over impurities was prevalent at the lowest temperatures. In

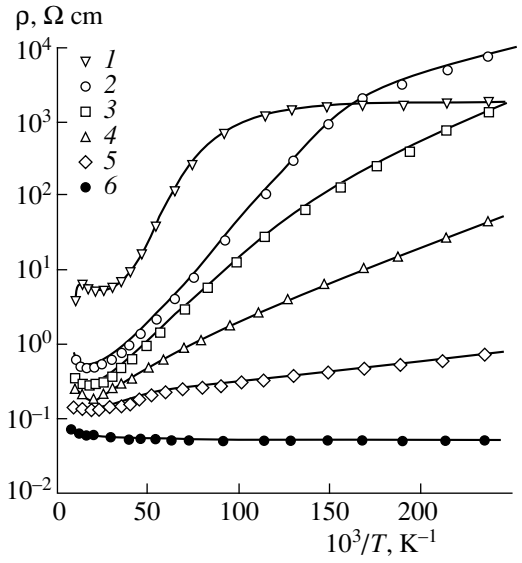


Fig. 1. Temperature dependence of resistivity of Cu-doped p-$Hg_{0.8}Cd_{0.2}Te$ crystals. N_A = (*1*) 2.6×10^{15}, (*2*) 3.2×10^{16}, (*3*) 6×10^{16}, (*4*) 1×10^{17}, (*5*) 1.8×10^{17}, and (*6*) 4.7×10^{17} cm^{-3}.

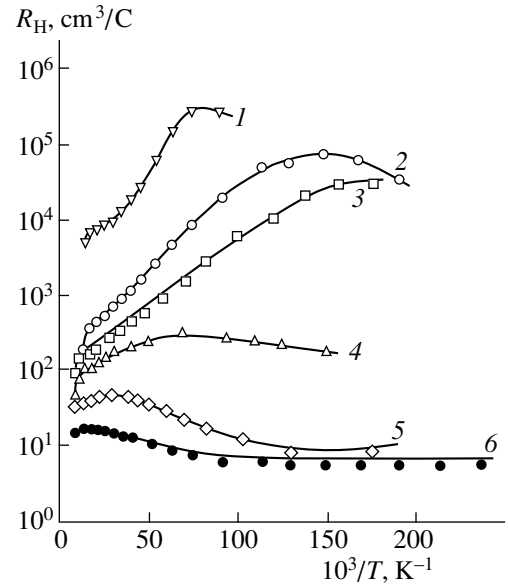


Fig. 2. Temperature dependence of the Hall coefficient for Cu-doped p-$Hg_{0.8}Cd_{0.2}Te$ crystals. N_A = (*1*) 2.6×10^{15}, (*2*) 3.2×10^{16}, (*3*) 6×10^{16}, (*4*) 1×10^{17}, (*5*) 1.8×10^{17}, and (*6*) 4.7×10^{17} cm^{-3}.

the samples from the control group, this conductivity obeyed the Mott law [8]

$$\rho = \rho_0 \exp[(T_0/T)^{1/4}], \quad (1)$$

whereas, in the Cu-doped crystals, its temperature dependence was described by the Arrhenius equation

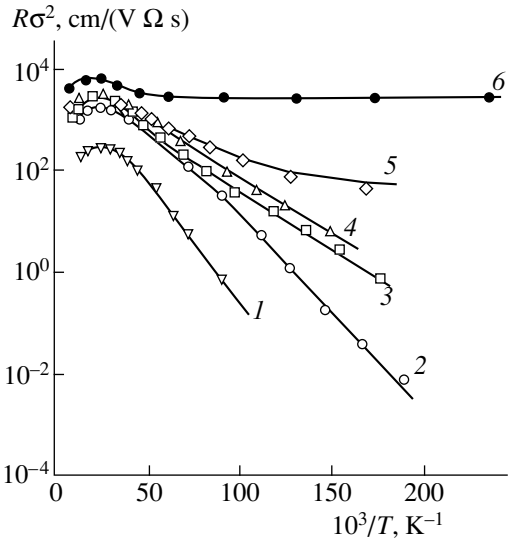


Fig. 3. Temperature dependence of the product $R_H\sigma^2$ for Cu-doped p-$Hg_{0.8}Cd_{0.2}Te$ crystals. N_A = (*1*) 2.6×10^{15}, (*2*) 3.2×10^{16}, (*3*) 6×10^{16}, (*4*) 1×10^{17}, (*5*) 1.8×10^{17}, and (*6*) 4.7×10^{17} cm^{-3}.

$$\rho = \rho_0 \exp\left(\frac{\varepsilon}{k_B T}\right). \quad (2)$$

In slightly compensated crystals with $N_A < 6 \times 10^{16}$ cm^{-3}, the activation energy of the low-temperature conductivity (2) increased with increasing N_A and the multiplier ρ_0 was proportional to $N_A^{-1/3}$. Therefore, according to [2], this low-temperature conductivity in such crystals was identified with the hopping ε_3-conductivity. In the compensated crystal with $K \approx 0.5$ and $N_A = 2.6 \times 10^{15}$ cm^{-3}, the activation energy was very low (0.14 meV, see Fig. 1), which was also characteristic of the ε_3-conductivity [1, 2].

For $N_A > 6 \times 10^{16}$ cm^{-3}, the activation energy of conductivity (2) decreased with increasing N_A according to the law

$$\varepsilon_2 = 4.4[1 - (N_A/N_{AM})^{1/3}] \text{ meV} \quad (3)$$

and vanished for $N_A > N_{AM} = 3.8 \times 10^{17}$ cm^{-3}. According to [1, 2], dependence (3) corresponds to the ε_2-conduction over the A^+ band formed by acceptors that are positively charged as a result of the attachment of an excess hole.

The sign of R_H was positive below the high-temperature inversion point (Fig. 2). A characteristic maximum of R_H was observed in the vicinity of the transition to the conduction over impurities. As is well known, such a maximum is observed if the hole mobility in the acceptor band and the magnitude of the corresponding Hall effect are low [2], i.e., if the value of the product $R_H\sigma^2$ in the transition region is controlled exclusively by the contribution of the valence-band holes in the case under consideration.

In fact, for $N_A < 10^{17}$ cm^{-3}, the temperature dependence of $R_H\sigma^2$ does not exhibit any inflections in the region where the dominant mechanism of conductivity σ changes (Fig. 3). Conversely, for $N_A > 10^{17}$ cm^{-3}, such an inflection can be clearly seen. In this concentration region, the Hall coefficient defined by charge carriers in the A^+ band is anomalously small and amounts to only 5 cm^3/C irrespective of N_A. The mobility of such carriers decreases steadily from 350 cm^2/(V s) for $N_A = 2 \times 10^{18}$ cm^{-3} to 15 cm^3/(V s) for $N_A = 1.8 \times 10^{17}$ cm^{-3}. Thus, for $N_A < 10^{17}$ cm^{-3} and $\sigma_1 > 0.1\sigma$, the charge carriers in the A^+ band do not contribute to $R_H\sigma^2$. Here, $\sigma_1 = ep\mu_p$ is the contribution of free holes to the total electrical conductivity.

3. PROCESSING OF THE RESULTS OF MEASUREMENTS

3.1. Determination of the Hole Concentration

The evidence for the conduction over impurities in p-$Hg_{0.8}Cd_{0.2}Te$ crystals is clear even at $T > 10$ K (Figs. 1, 2), and the ionization energy of acceptors is low, so that the freeze-out of the ε_1-conductivity becomes important only below 20 K. Therefore, in this case, the correct determination of E_A directly from the value of R_H is possible only in a very narrow range of T (15–20 K); at lower temperatures, it is necessary to combine the results of measurements of R_H and ρ.

Since the value of the product $R_H\sigma^2$ in the crystals with $N_A < 10^{17}$ cm^{-3} is controlled exclusively by the valence-band holes, the concentration of free holes in the valence band can be found from the expression

$$p = \frac{r_H(\rho\sigma_1)^2}{eR_H}. \quad (4)$$

In order to determine σ_1, we had to separate the contributions of different mechanisms of charge transport in the region of mixed conduction.

Since dissimilar mechanisms of charge transport are dominant in different temperature ranges and the total electrical conductivity is obtained by adding the contributions of different types of charge carriers, σ_1 could be separated by empirical extrapolation. For $N_A < 6 \times 10^{16}$ cm^{-3} (for which case the ε_3-conductivity is observed), such a procedure creates no difficulties, because the activation energy ε_3 is much lower than ε_1 and depends on temperature only slightly [2]. For $N_A > 6 \times 10^{16}$ cm^{-3}, we observe the ε_2-conductivity, whose temperature dependence is defined by the probability of occupation of the A^+ band and, thus, is appreciably nonmonotonic. Its activation energy was constant only at

$T < 10$ K; therefore, the separation of σ_1 using an extrapolation was correct in this temperature region alone. Further on, the portion of dependence $\sigma_1(T)$ that had been separated for $T < 10$ K was smoothly joined to the portion corresponding to $T > 50$ K where $\sigma \approx \sigma_1$. The accuracy of extrapolation was acceptable only if $N_A < 10^{17}$ cm^{-3}, in which case the value of σ_2 was distinctly smaller than that of the exhausted ε_1-conductivity.

3.2. The Mobility and Hall Factor of Holes

The data on the mechanism of low-temperature hole mobility in $Hg_{0.8}Cd_{0.2}Te$ are contradictory. The value of μ_p in this material is much smaller than that predicted by the Brooks–Herring theory and is typically accounted for by scattering either by neutral defects [9] or by the composition fluctuations [10]. However, experimental data show that $\mu_p \propto T^{3/2}$ at $T < 12$ K, which exactly corresponds to the scattering of holes by impurity ions and is at variance with the scattering mechanisms suggested elsewhere [9, 10]. In addition, the value of μ_p in the samples compensated with indium decreased in inverse proportion to an increase in the ionized-center concentration N_I. In particular, at $T \approx 20$ K, the Hall mobility was about 1150, 550, and 180 cm^2/(V s) in the samples with the impurity-ion concentrations of 1.5×10^{16}, 4.4×10^{16}, and 1.8×10^{17} cm^{-3}, respectively. On this basis, we assumed that holes were scattered by impurity ions at $T < 25$–30 K.

The Hall factor r_H in p-$Hg_{0.8}Cd_{0.2}Te$ may be formally represented as the product $r_{Hh}r_{Hl}(B)$, where r_{Hh} is the Hall factor of heavy holes and $r_{Hl}(B)$ is the factor describing the contribution of light holes. According to [6], $r_{Hl}(0) = 1.35 \pm 0.05$ at $T = 77$ K. Experimentally, the value of $r_{Hl}(0)$ depends on temperature only slightly at $T < 77$ K (apparently owing to the fact that the light holes are scattered by impurity ions even at $T = 77$ K). Therefore, we assumed in calculations that $r_{Hl}(0) = 1.35$ in the entire temperature range under consideration.

The Hall factor of heavy holes is equal to unity at $T = 77$ K [6], in which case they are scattered by the lattice vibrations. At $T < 70$ K, the scattering mechanism begins to change, which is evidenced by a characteristic kink in the dependence $\mu_p(T)$. We assumed that $r_{Hh} = 1.93$ at low temperatures T, in which case $\mu_p \propto T^{3/2}$. A thorough analysis shows that (for example, for $N_A \sim (2–3) \times 10^{16}$ cm^{-3}) the Hall factor of heavy holes levels off even at $T \approx 15$ K and that its value decreases gradually to unity in the temperature range from 15 to 70 K.

3.3. Determination of the Degree of Compensation in the Crystals

The donor concentration N_D was estimated both from the results of measuring the hole mobility in each of the samples and from the data on hopping conductivity in unannealed crystals.

Since, in the case of scattering of holes by ionized centers, we have $T^{3/2}/\mu_p \propto N_I$, the donor concentration could be estimated from this relation for various of N_D using the eletroneutrality condition

$$p + N_D + N_A^+ = N_A^-, \tag{5}$$

which yields $N_I = p + 2N_D + 2N_A^+$. The values of N_D obtained by this method varied from 1×10^{15} to 2×10^{15} cm^{-3}.

The method for determination of N_D from the value of the parameter T_0 in the expression for hopping conductivity (1) is based on the percolation theory [2], according to which

$$T_0 = \frac{21.2}{k_B g(F) a^3}. \tag{6}$$

It was shown [11–13] that this method can be applied to unannealed p-$Hg_{0.8}Cd_{0.2}Te$ crystals. Our studies of In-doped crystals, in which case the value of N_D could be determined by the independent method, showed that the approach based on (6) yields the correct estimate of N_D, if we assume that the density of states at the Fermi level is $g(F) = \varepsilon N_D/(e^2 N_A^{1/3})$, where ε is the permittivity and the localization radius of hole bound by a vacancy is $a = \hbar/\sqrt{m_{lh}(E_{A1} + E_{A2})}$. Here, E_{A1} and E_{A2} are the energies of the first and second ionization of vacancy ($E_{A1} \approx 18$ meV and $E_{A2} \approx 2E_{A1}$). According to these data, the concentration of uncontrolled donors in the ingots we used amounted to $(1.2 \pm 0.2) \times 10^{15}$ cm^{-3}, which is in good agreement with the results of analysis of μ_p.

3.4. Determination of the Acceptor-Level Energy

The electrical field of compensating donors affects appreciably the position of the Fermi level in the crystal and depends on K and N_A [2]. Therefore, in calculations of E_A, the condition for electroneutrality was written differently for dissimilar ranges of N_A, K, and T. Calculations were performed separately for each measurement point in the region of $T < 22$ K where the freeze-out of free holes was significant, screening was relatively slight, and the value of r_H was determined quite reliably. The data on the density of states N_V in the valence band were taken from [6].

In the sample with $N_A \approx 2.5 \times 10^{15}$ cm^{-3} and $K \approx 0.5$, the acceptor band is fairly symmetric, its width is comparable with k_BT, and the Fermi level at $T = 0$ is located

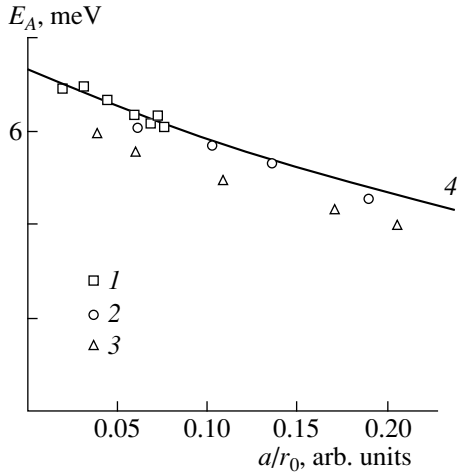


Fig. 4. Influence of screening on the copper-ionization energy in p-$Hg_{0.8}Cd_{0.2}Te$ crystals. Dots *1–3* correspond to experimental data and curve *4* represents the results of calculation for $E_A(0) = 7.4$ meV. $N_A = (1)\ 2.6 \times 10^{15}$, $(2)\ 2.1 \times 10^{16}$, and $(3)\ 3.2 \times 10^{16}$ cm^{-3}.

almost at the midgap [2], which is evidenced by an extremely low activation energy for the ε_3-conductivity (0.14 meV). On the other hand, we have in this case $N_A^+ \ll p$. As a result, in order to calculate E_A in this sample, we can use the condition for electroneutrality in its conventional form [3]; i.e.,

$$p + N_D = \frac{N_A}{1 + 4\exp[(E_A - F)/k_BT]}. \quad (7)$$

The results of calculations are shown in Fig. 4 (dots *1*). It can be seen that the value of E_A decreases with an increasing degree of screening. Furthermore, the dependence of E_A on a/r_0 follows exactly the law suggested elsewhere [3]:

$$E_A = E_A(0)[2\exp(-a/r_0) - 1], \quad (8)$$

if we assume that $E_A(0) = 7.4$ meV and the radius of the state is given by $a = \hbar/\sqrt{2m_{hh}E_A(0)}$ [14]. Here, r_0 is the Debye radius for screening the charges by free holes. The value of $m_{hh} \approx 0.39m_0$ [6] was used in the calculations.

For $N_A > 2 \times 10^{16}$ cm^{-3}, the effects of both the potential of compensating donors and the occupancy of the A^+ acceptor states were significant and, thus, were taken into account in determining the value of E_A.

For N_A^+, we performed a variational calculation of the energy of attaching a hole to a neutral acceptor. The results show that an isolated simple acceptor in $Hg_{0.8}Cd_{0.2}Te$ forms three doubly degenerate bound A^+ states with binding energies ΔE_j (j = 1, 2, and 3) equal to $0.04E_A$, $0.078E_A$, and $0.115E_A$, respectively. If we ignore the splitting and assume that copper introduces into the band gap of $Hg_{0.8}Cd_{0.2}Te$ a fourfold degenerate A^- level with the energy E_A and a sixfold degenerate A^+ level with the energy $\Delta E \approx 0.08E_A \approx 0.6$ meV, the probability for occupation of the charged states of such a multicharged acceptor would be defined by the Gibbs distribution given by

$$f_A^+(F) = \frac{6\exp[(\Delta E - F)/k_BT]}{4 + \exp[(F - E_A)/k_BT] + 6\exp[(\Delta E - F)/k_BT]}, \quad (9)$$

$$f_A^-(F) = \frac{\exp[(F - E_A)/k_BT]}{4 + \exp[(F - E_A)/k_BT] + 6\exp[(\Delta E - F)/k_BT]}. \quad (10)$$

The structure of impurity band was considered in terms of the theory outlined elsewhere [2].

The crystals with $N_A > 2 \times 10^{16}$ cm^{-3} were slightly compensated ($K < 0.1$). According to [2], in such crystals at $T = 0$, the compensating donors form complexes with the nearest acceptors and thus decrease the acceptors' energy by about $\varepsilon_A = e^2/(\varepsilon r_A)$, where $r_A = (4\pi N_A/3)^{1/3}$. Most of these pairs are entirely neutral. The remaining pairs form positively charged N_0 complexes and negatively charged N_2 complexes with concentrations of $N_0 = N_D\exp[-\varepsilon_A^3(F - E_A)^3]$ and $N_2 = 7.14 \times 10^{-4}N_D[\varepsilon_A/(F - E_A)]^6$, respectively. At $T = 0$, the concentrations of differently charged complexes are equal and the ionization energy for acceptor and Fermi energy are related as

$$E_A(0) = F(0) + 0.61\varepsilon_A \approx F(0) + 0.99\frac{e^2}{\varepsilon}N_A^{1/3}. \quad (11)$$

If temperatures are not too high, the charged complexes do not decompose. According to [2], in this case, we have $(N_0, N_2) \gg (p, N_A^+)$ and the condition for electroneutrality is given by

$$P + N_0 + \frac{3}{2}N_A\exp\left[\frac{\Delta E - F}{k_BT}\right] = N_2 + \frac{1}{4}N_A\exp\left(\frac{F - E_A}{k_BT}\right) \quad (12)$$

(here, we additionally took into account the occupation of the valence-band states and the A^+ states of acceptors).

The condition $(N_0, N_2) \gg (p, N_A^+)$ was met in the crystals with $N_A \approx (2–3) \times 10^{16}$ cm^{-3} at T = 5–6 K. Therefore, in this range of T and N_A, relationship (12) was used in calculations of E_A. As a result, we found

that $E_A(0) = 7.4$ meV for $N_A = 2.1 \times 10^{16}$ cm^{-3} and $E_A(0) = 7.1$ meV for $N_A = 3.2 \times 10^{16}$ cm^{-3}.

At higher temperatures (in this case, at $T > 8$ K), the complexes decompose [2] and the Fermi energy is controlled by the probability of thermal excitation of holes. A proper consideration of the structure of the impurity band is a difficult problem. Therefore, when determining E_A at $T > 8$ K in the crystals with $N_A \approx (2–3) \times 10^{16}$ cm^{-3}, the structure of the acceptor band was considered approximately.

We assumed that all donor–acceptor pairs are identical and their influence on the energy of acceptor levels in the main peak may be disregarded. In this approximation, the A^- band is reduced to two levels with energies of E_A and $E_A - \langle\varepsilon_A\rangle$ that correspond to the acceptor concentrations of $N_A - N_D$ and N_D, respectively. The probability of occupation of the A^+ acceptor state in the complex is low; therefore, the A^+ band was replaced with a single level with the energy of ΔE. The optimal value $\langle\varepsilon_A\rangle$ depends on T; numerical calculation demonstrated that, in the range of T, this temperature dependence might be approximated (with allowance made for screening) with the expression $\langle\varepsilon_A\rangle \approx 1.2\varepsilon_A \exp(-r_A/r_0)$. In such a model, the Fermi level is located exactly at the midposition between the levels E_A and $E_A - \langle\varepsilon_A\rangle$, and $F(0)$ almost coincides with the value defined by relationship (11).

As a result, the condition for electroneutrality takes the following form:

$$p + N_D = (N_A - N_D)[f_A^-(F) - f_A^+(F)] + N_D f_A^-(F + \langle\varepsilon_A\rangle). \quad (13)$$

The results of determination of E_A in this case are shown in Fig. 4 (dots *2*, *3*). It can be seen that E_A decreases as the degree of screening increases and follows theoretical dependence (8), as in the case of the compensated sample with $N_A = 2.6 \times 10^{15}$ cm^{-3}. Extrapolating the dependence $E_A(T)$ to $T = 0$, we obtain $E_A(0) = 7.1$ meV for $N_A = 2.1 \times 10^{16}$ cm^{-3} and $E_A(0) = 6.7$ meV for $N_A = 3.2 \times 10^{16}$ cm^{-3}. These values are somewhat smaller than those obtained above for the same samples using expression (12). Presumably, the agreement would be appreciably improved if we took into account the broadening of the main peak in the impurity band.

In the case of intermediately doped samples with $N_A > 5 \times 10^{16}$ cm^{-3}, the acceptor-band width exceeds E_A. Consequently, we should not expect here good agreement with the experiment in the context of such a simplified model of the impurity band; therefore, E_A was not calculated for finite temperatures. Instead, we used the results of measurements of the hole concentration to calculate the temperature dependence of the Fermi energy $F(T)$ (Fig. 5). At $T < 10$ K, this dependence was linear in the actual range of T. For the samples with $N_A = 3.2 \times 10^{16}$ cm^{-3}, linear extrapolation of the dependence $F(T)$ to $T = 0$ yields $F(0) = 4.5$ meV; the latter value, in view of (9), corresponds to $E_A(0) = 7.1$ meV, which is in agreement with the results of calculation with formula (12). This gave grounds to presume that the linear extrapolation of $F(T)$ for $T < 10$ K is valid for the samples with $N_A \approx (5–9) \times 10^{16}$ cm^{-3} as well. This presumption is also supported by the fact that the activation energy for the ε_2-conductivity is constant at low temperatures, which is possible only if the dependence $F(T)$ is linear.

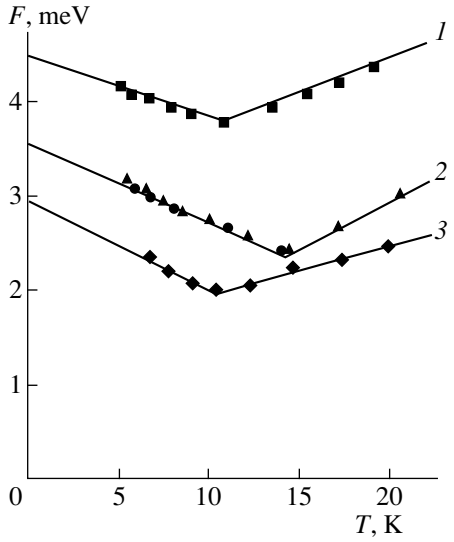


Fig. 5. The temperature dependence of the Fermi energy for electrons in Cu-doped p-$Hg_{0.8}Cd_{0.2}Te$ crystals. N_A = (*1*) 3.2×10^{16}, (*2*) 5.3×10^{16}, and (*3*) 9.1×10^{16} cm^{-3}.

The values of $E_A(0)$ obtained by this method are shown in Fig. 6 (curve *1*) along with the corresponding values for the other samples. It follows from the data obtained that the value of $E_A(0)$ is almost independent of N_A in the entire range of concentrations under consideration.

The results of calculation of E_A by the conventional empirical method [i.e., directly from the slope of the dependence $R_H(T)$] are also shown in Fig. 6 (curve *2*). It can be seen that the values of E_A obtained by the conventional method decrease with increasing N_A and vanish for $N_A \approx 2 \times 10^{17}$ cm^{-3} (i.e., before the Mott transition), similarly to what was reported in [4].

For $N_A > 10^{17}$ cm^{-3}, attempts to separate the ε_1 conductivity in uncombined form fail. Therefore, in this case, we can estimate the value of E_A only in a rather crude approximation using the activation energy for the ε_2-conductivity. To this end, we assume that the model of the impurity band suggested in [2] is valid in the vicinity of the Mott transition and the ε_2-conductivity is related to excitation of holes from the Fermi level to the nearest A^+ level with the energy of $\Delta E_3 \approx 1$ meV. We

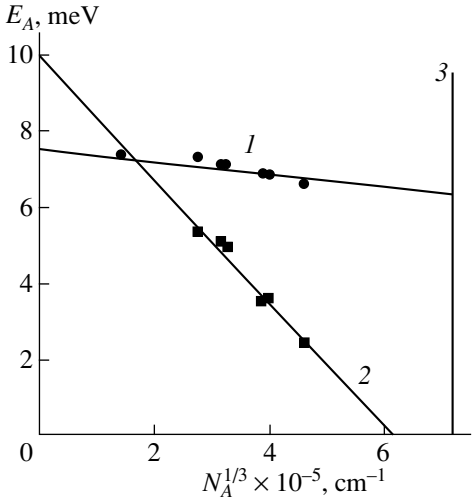


Fig. 6. The concentration dependence of ionization energy for Cu in p-$Hg_{0.8}Cd_{0.2}Te$ crystals at $T = 0$. Curve *1* corresponds to calculations with allowance made for the impurity-band structure, and curve *2* was calculated with the impurity-band structure disregarded. Vertical line *3* indicates the point of the Mott transition.

then use (3) and (11) to find that $E_A \approx 6.5 \pm 0.2$ meV for N_A varying from 10^{17} to $N_{AM} = 3.8 \times 10^{17}$ cm^{-3}.

4. DISCUSSION OF THE RESULTS

Thus, as a result of considering (i) an actual structure of the acceptor band and (ii) the screening of the bound charges with free holes, we obtained data that challenge the conventional concepts of the behavior of E_A in the vicinity of the Mott transition in p-$Hg_{0.8}Cd_{0.2}Te$. We found that the acceptor-ionization energy extrapolated to $T = 0$ depended on N_A only slightly. Since the localization radius of the impurity state is less sensitive to screening than is the case for E_A [3], the value of this radius is most likely independent of N_A altogether at $T = 0$. Consequently, it is quite justified to use the parameters of an isolated acceptor in the Mott criterion.

It is not difficult to explain the result obtained. In fact, in the case of light doping ($N_A < 3 \times 10^{16}$ cm^{-3}), the value of the overlap integral for wave functions of neighboring acceptors is negligibly small compared to E_A and there are no free charge carriers at $T = 0$; therefore, there is no physical reason for the dependence of $E_A(0)$ on the acceptor concentration. As the doping becomes heavier, the value of the overlap integral increases; however, this increase mainly manifests itself in symmetric splitting of interacting levels and affects the mean value of $E_A(0)$ only in the second approximation of the perturbation theory [2]. As a result, the binding energy for holes at acceptors in p-$Hg_{0.8}Cd_{0.2}Te$ depends only slightly on the acceptor concentration right up to the metal–insulator transition.

At finite temperatures, free charge carriers appear in the crystal, and, as a result of screening, the acceptor-ionization energy diminishes. For sufficiently high T and N_A, this can induce the metal–insulator transition of another type, i.e., such that E_A vanishes at the transition point. For example, at $T = 77$ K and in view of (8), such a transition should occur for $N_A \approx 2.2 \times 10^{17}$ cm^{-3}. This value is almost the same as for N_A at the point where the value of E_A determined by the conventional empirical method (Fig. 6, curve *2*) vanishes. Thus, the conventional method for determination of E_A from the results of measurements of R_H yields an estimate of mean effective E_A at finite temperatures, in which case E_A is subject to the effect of screening.

Extrapolating curve *1* (Fig. 6) to $N_A = 0$, we obtain the ionization energy for isolated acceptor $E_{A0} = 7.6 \pm 0.2$ meV. This value is in excellent agreement with the result of calculation based on the effective-mass method developed in [15] as applied to diamond-like semiconductors. According to this calculation, the ionization of a simple acceptor in $Hg_{0.8}Cd_{0.2}Te$, where $m_{lh}/m_{hh} \sim 0.01$, is equal to exactly 7.6 meV, if we assume that $m_{hh} \approx 0.39m_0$ [7] and $\varepsilon \approx 17.6$ [16].

5. CONCLUSION

(i) The conventional empirical approach to processing the results of the Hall measurements with the aim of determining the ionization energy for impurities becomes incorrect in the case of narrow-gap semiconductor crystals. In order to obtain reliable data, we have to consider carefully the impurity-band structure and the screening of the impurity potential by free charge carriers. In the case of light doping and slight compensation, the ionization energy for shallow-level centers in such crystals can be calculated in the context of a simplified impurity-band model that makes it possible, in the first approximation, to take into account both of the above factors.

(ii) At $T = 0$, the acceptor-ionization energy depends only slightly on the acceptor concentration and remains finite at the point of the Mott transition. Correspondingly, under these conditions, the localized-state radius for a hole in the vicinity of the Mott transition is almost equal to the localization radius of a hole at an isolated acceptor.

(iii) The Mott transition can be observed in the pure form in the impurity band of Cu-doped p-$Hg_{0.8}Cd_{0.2}Te$ crystals only in the vicinity of absolute zero of temperature. For finite temperatures, the metal–insulator transition is strongly affected by screening of the acceptor charge with free holes.

ACKNOWLEDGMENTS

I am grateful to OAO "Pure Metals" and, especially, to Dr. K.R. Kurbanov for placing at my disposal the samples of high-quality $Hg_{1-x}Cd_xTe$ crystals for the experiments.

REFERENCES

1. N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Clarendon Press, Oxford, 1979; Mir, Moscow, 1982), Vol. 1.
2. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984).
3. I. M. Tsidil'kovskiĭ, G. I. Kharus, and N. G. Shelushinina, *Impurity States and Transport Phenomena in Zero-Gap Semiconductors* (Sverdlovsk, 1987).
4. M. C. Chen and J. H. Tregilgas, J. Appl. Phys. **61** (2), 787 (1987).
5. V. V. Bogoboyashchiĭ, A. I. Elizarov, V. A. Petryakov, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **21** (8), 1469 (1987) [Sov. Phys. Semicond. **21**, 893 (1987)].
6. V. V. Bogoboyashchiy, Proc. SPIE **3486**, 325 (1997).
7. A. I. Elizarov, V. V. Bogoboyashchiĭ, and N. N. Berchenko, Fiz. Tekh. Poluprovodn. (Leningrad) **24** (3), 446 (1990) [Sov. Phys. Semicond. **24**, 278 (1990)].
8. N. F. Mott, J. Non-Cryst. Solids **1**, 1 (1968).
9. J. R. Mayer, F. J. Bartoli, and C. A. Hoffman, J. Vac. Sci. Technol. A **5** (5), 3035 (1987).
10. M. A. Berding, S. Krishnamurthy, A. Sher, and A.-B. Chen, J. Vac. Sci. Technol. A **5** (5), 3014 (1987).
11. I. S. Shlimak, A. L. Éfros, and I. Ya. Yanchev, Fiz. Tekh. Poluprovodn. (Leningrad) **11** (2), 257 (1977) [Sov. Phys. Semicond. **11**, 149 (1977)].
12. A. I. Elizarov and V. I. Ivanov-Omskiĭ, Fiz. Tekh. Poluprovodn. (Leningrad) **15** (5), 927 (1981) [Sov. Phys. Semicond. **15**, 531 (1981)].
13. A. I. Elizarov, V. V. Bogoboyashchiĭ, and N. N. Berchenko, Fiz. Tekh. Poluprovodn. (Leningrad) **18** (3), 455 (1984) [Sov. Phys. Semicond. **18**, 283 (1984)].
14. B. L. Gel'mont, A. R. Gadzhiev, B. L. Shklovskiĭ, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **8** (12), 2377 (1974) [Sov. Phys. Semicond. **8**, 1549 (1974)].
15. B. L. Gel'mont and M. S. D'yakonov, Fiz. Tekh. Poluprovodn. (Leningrad) **5** (11), 2191 (1971) [Sov. Phys. Semicond. **5**, 1905 (1971)].
16. A. V. Lyubchenko, E. A. Sal'kov, and F. F. Sizov, in *Physical Foundations of Semiconductor Quantum Photoelectronics* (Naukova Dumka, Kiev, 1984), p. 126.

Translated by A. Spitsyn

Semiconductors, Vol. 34, No. 8, 2000, pp. 924–926. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 962–964.

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Nonlinear Optical Absorption in a Heavily Doped Degenerate *n*-GaAs

V. L. Malevich* and I. A. Utkin

Division of Optical Problems of Information Science, Belarussian Academy of Sciences, Minsk, 220141 Belarus
** e-mail: vitaly@optoinform.bas-net.by*

Submitted January 13, 2000; accepted for publication January 27, 2000

Abstract—The nonlinear optical absorption in a heavily doped degenerate *n*-GaAs in the spectral region near the fundamental absorption edge was studied. The relative contribution of the effects of states' filling, band-gap narrowing, and screening of the tails of density of states by photoexcited charge carriers to the nonlinear optical absorption was analyzed on the basis of the quasi-classical model of fluctuation levels. An experimentally observed increase in optical transmission was shown to be mainly caused by the dynamic states' filling effect. *© 2000 MAIK "Nauka/Interperiodica".*

The tails of the density of states in the band gap of heavily doped semiconductors introduce some specific features in both their linear [1–5] and nonlinear [6–8] optical properties as compared with the case of pure and lightly doped crystals. These features are most pronounced in the spectral region near the fundamental absorption edge. In this paper, the nonlinear absorption in a heavily doped slightly compensated *n*-GaAs near the fundamental absorption edge was calculated on the basis of the model used in [8, 9]. The results of the calculation were compared with the experimental spectra of nonlinear optical transmission.

The spectra of nonlinear transmission were measured with a tunable Li:F_2^- color-center laser, which emits 20 ns light pulses in the spectral region of 860–1000 nm with an energy of about 100 μJ. Laser radiation was focused to a Gaussian spot with a characteristic diameter of about 1 mm on the sample surface. The semiconductor sample was a 0.32-mm-thick *n*-GaAs:Te ($N_d^+ = 1.8 \times 10^{18}$ cm^{-3}) wafer. The laser beam was directed onto the sample surface at the Brewster angle to avoid the effects of Fresnel reflection.

The interband optical absorption in heavily doped *n*-type degenerate semiconductors in the spectral region near the fundamental absorption edge is mainly due to indirect electron transitions from the valence-band tail to the vacant conduction-band states below the Fermi level [1, 2]. The corresponding expression for the absorption factor α at the frequency ω can be written as [3]

$$\alpha(\omega) = \frac{4\pi^2 e^2}{c\bar{n}m_0^2\omega}\int d\varepsilon \int d\varepsilon' g_c(\varepsilon) g_v(\varepsilon') |M_{c,v}(\varepsilon, \varepsilon')|^2 \times [f_v(\varepsilon') - f_c(\varepsilon)]\delta(\varepsilon - \varepsilon' - \hbar\omega), \quad (1)$$

where $g_{c,v}$ and $f_{c,v}(\varepsilon)$ are the densities of states and the energy-distribution functions of the electrons in the conduction (c) and valence (v) bands, c is the speed of light, $\bar{n}$ is the refractive index of the semiconductor, m_0 is the free-electron mass, and $M_{c,v}(\varepsilon, \varepsilon')$ is the matrix element of the momentum for the interband transitions.

The density of states in the valence band will be described on the basis of the quasi-classical model of fluctuation levels [1]. The conventional expression for the parabolic band ($\sim\sqrt{\varepsilon}$) will be used to describe the density of states in the conduction band. Assuming the locally equilibrium distribution of carriers in the bands and the energy-independent transition matrix element, the interband absorption factor can be written as [8]

$$\alpha(\omega) = \frac{K\gamma^{1/2}}{\omega}\int_0^\infty d\varepsilon \sqrt{\varepsilon} \exp\left[-\frac{(\varepsilon + \varepsilon_g - \hbar\omega)^2}{2\gamma^2}\right] \times D_{-3/2}\left[\frac{\sqrt{2}(\varepsilon + \varepsilon_g - \hbar\omega)}{\gamma}\right] \times \left\{\left[1 + \exp\left(\frac{\varepsilon - \hbar\omega - \mu_h}{k_B T}\right)\right]^{-1} - \left[1 + \exp\left(\frac{\varepsilon - \mu_e}{k_B T}\right)\right]^{-1}\right\}. \quad (2)$$

Here, T is the crystal temperature, k_B is the Boltzmann constant, μ_e and μ_h are the quasi-Fermi levels for electrons and holes, $D_{-3/2}(x)$ is the parabolical cylinder function, and ε_g is the band gap; the energy ε is reckoned from the conduction-band edge. The coefficient K is defined by the parameters of the semiconductor and is independent of the wavelength and energy of the incident radiation. The parameter γ characterizing the width of the tail of density of states represents the doubled rms energy of the electron in the random field of the ionized impurity. This parameter is expressed as

$\gamma = (e^2/\varepsilon_0 r_0)\sqrt{4\pi r_0^3 N}$, where r_0 is the screening length, ε_0 is the static permittivity, and N is the mean concentration of the ionized impurity.

In the spectral region under consideration and for the radiation intensities used, two-photon absorption can be neglected. Thus, expression (2) can be used to describe the nonlinear absorption with regard to the dependence of such parameters as the band gap, electron and hole quasi-Fermi levels, and the parameter of the tail γ on the carrier density. The latter in its turn is defined by the radiation intensity and the interband recombination rate.

The dependence of the band gap on the electron density is described by the expression $\varepsilon_g(n) = \varepsilon_{g0} - \lambda n^{1/3}$, where $\lambda = 2 \times 10^{-8}$ eV cm [3] and $\varepsilon_{g0} = 1.425$ eV is the band gap in pure GaAs. The variation of the parameter γ with the pumping level is defined by the dependence of the screening radius on the carrier density. The calculations were performed both with allowance made for this effect and with ε_g and γ assumed constant for the purpose of elucidation of the relative contribution of the screening effect and that of dynamic band-gap narrowing to the absorption nonlinearity. The dependences of α on the total electron density n for two wavelengths are shown in Fig. 1. These dependences were calculated by formula (2). As follows from the calculations, the absorption factor decreases almost linearly with the increase in carrier density in the case under consideration. The main contribution to the absorption saturation is due to the effect of the filling of states with photoexcited carriers (the dynamic states' filling effect).

The experimental data on the nonlinear optical transmission of the semiconductor wafer were analyzed on the basis of a numerical solution of the differential equation system

$$\frac{\partial n}{\partial t} = \frac{\alpha(n) I(z,t)}{\hbar\omega} - \frac{(n-n_0)}{\tau} - Bn(n-n_0), \qquad (3)$$

$$\frac{\partial I(z,t)}{\partial z} = -[\alpha(n) + \sigma n] I(z,t). \qquad (4)$$

Here, $I(z, t)$ is the radiation intensity (the z-axis is aligned with the inward normal to the sample surface); the parameters τ and B define the rates of linear and radiative recombination, respectively; $\alpha(n)$ is the previously calculated dependence of the interband absorption factor on the carrier density; and n_0 is the equilibrium electron density. The diffusion of nonequilibrium carriers is neglected in equation (3), since the diffusion length in this case is substantially smaller than the absorption length of radiation ($\sim\alpha^{-1}$) and the beam diameter. The second term on the right-hand side of equation (4) represents the intraband absorption (σ is the absorption cross section) whose magnitude may be comparable with the interband contribution in the spectral region under consideration.

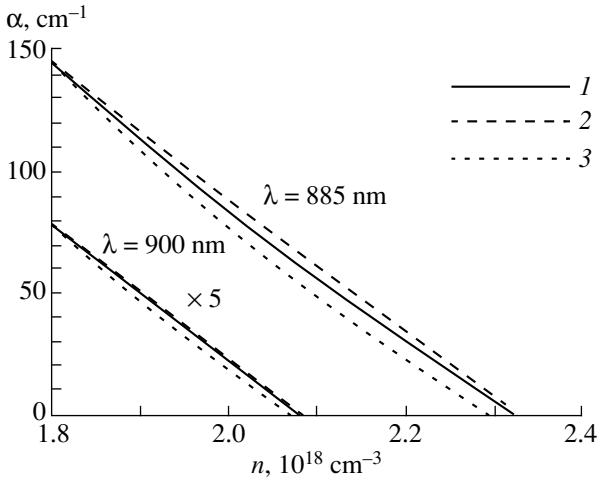


Fig. 1. Dependences of the interband optical absorption factors α of the heavily doped degenerate n-GaAs ($N_d^+ = 1.8 \times 10^{18}$ cm^{-3}) on the total electron density n calculated for two wavelengths (*1*). Curves *2* and *3* were obtained without regard for the effects of screening and dynamic narrowing of the band gap, respectively.

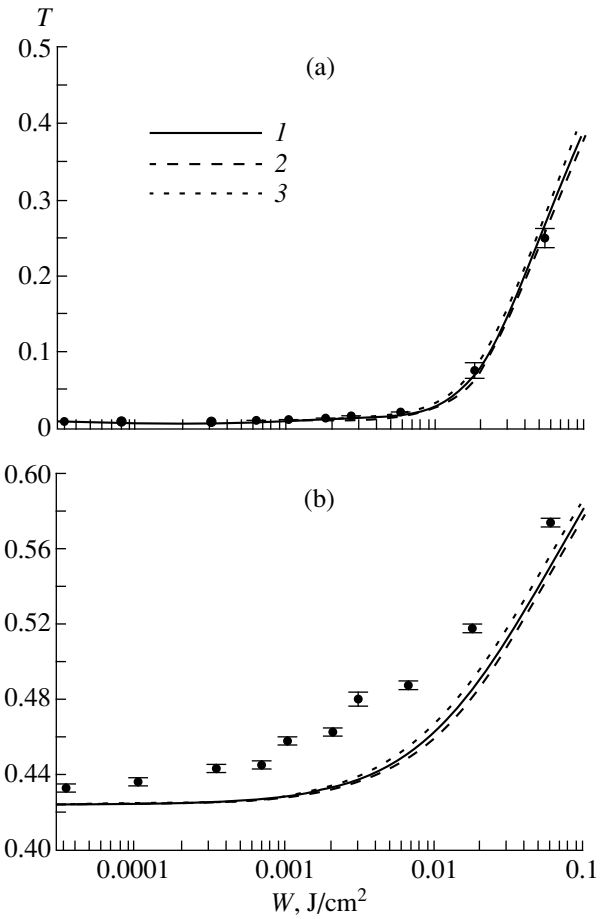


Fig. 2. Dependence of the transmission T of the semiconductor sample on the energy density of the incident radiation W at the wavelengths λ = (a) 885 and (b) 900 nm. The points are the experimental data. The designations of the calculated curves are the same as in Fig. 1.

The experimental and calculated dependences of the transmission of a heavily doped *n*-GaAs sample on the energy density in a laser pulse are shown in Fig. 2. The following values of parameters were used in the calculations: $\sigma = 2 \times 10^{-17}$ cm^2 [10] and $B = 2 \times 10^{-10}$ cm^2/s [11]. The linear-recombination time depends strongly on the technology of sample growth. The value of $\tau = 7$ ns was used in calculations, which provided the best fit of the theoretical curves to the experimental data.

It is readily seen that the effective lifetime of carriers in the case under consideration is about 2 ns, being much less than the duration of the excitation pulse. The quasi-steady conditions are realized in this case and the intensity-dependent interband absorption factor $\alpha(I) = \alpha_0/(1 + I/I_s)$ can be introduced for moderate photoexcitation levels ($(n - n_0)/n_0 \ll 1$) if we assume that the dependence of α on n is linear. Here, α_0 is the linear absorption factor, $I_s = \hbar\omega(1 + Bn_0\tau)/\beta\tau$, and $\beta = d\alpha/dn|_{n = n_0}$. The parameter I_s is equal to 0.4 and 1.9 MW/cm^2 for the wavelengths of 885 and 900 nm, respectively.

Thus, in this paper, the nonlinear transmission in a heavily doped degenerate *n*-GaAs was studied both experimentally and theoretically in the spectral region near the fundamental absorption edge. The experimentally observed levelling-off of absorption was shown to be well described in the context of the model of indirect interband transitions of electrons from the valence band tail to the conduction band states. This model takes into account the dependences of the quasi-Fermi level, band gap, and characteristic width of the tail of the density of states on the photocarrier density. The results of the numerical simulation show that the main contribution to nonlinear absorption is provided by the effect of conduction-band filling with nonequilibrium electrons.

ACKNOWLEDGMENTS

This work was supported by the International Technical-Scientific Center, grant no. B-129.

REFERENCES

1. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984).
2. A. L. Éfros, Usp. Fiz. Nauk **111**, 451 (1973) [Sov. Phys. Usp. **16**, 789 (1973)].
3. H. C. Casey, Jr. and F. Stern, J. Appl. Phys. **47**, 631 (1976).
4. W. Sritrakool, V. Sayakanit, and H. R. Glyde, Phys. Rev. B **32**, 1090 (1985).
5. G. B. Lush, M. R. Melloch, M. S. Lundstrom, *et al.*, J. Appl. Phys. **74**, 4694 (1993).
6. V. P. Gribkovskiĭ, in *The Theory of Emission and Absorption of Light in Semiconductors* (Nauka i Tekhnika, Minsk, 1976), Chap. 3, p. 245.
7. S. A. Gurevich, A. E. Fedorovich, and F. V. Fedorov, Fiz. Tekh. Poluprovodn. (Leningrad) **25**, 769 (1991) [Sov. Phys. Semicond. **25**, 465 (1991)].
8. S. A. Bystrimovich, R. G. Zaporozhchenko, V. L. Malevich, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **28**, 1020 (1994) [Semiconductors **28**, 590 (1994)].
9. S. W. Koch, S. Schmitt-Rink, and H. Haug, Phys. Status Solidi B **106**, 135 (1981).
10. E. Haga and H. Kimura, J. Phys. Soc. Jpn. **19**, 1596 (1964).
11. G. B. Lush, H. F. MacMillan, B. M. Keyes, *et al.*, J. Appl. Phys. **72**, 1436 (1992).

Translated by Yu. Aleshchenko

Semiconductors, Vol. 34, No. 8, 2000, pp. 927–930. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 965–969.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Tunnel Light-Emitting Si:(Er,O) Diodes with a Short Rise Time of Er^{3+} Electroluminescence under Breakdown Conditions

A. M. Emel'yanov*, N. A. Sobolev*, M. A. Trishenkov, and P. E. Khakuashev****

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

*** NPO Orion, Moscow, 111123 Russia*

Submitted February 9, 2000; accepted for publication February 14, 2000

Abstract—Si:(Er,O)-based tunnel light-emitting diodes were fabricated and exhibited the shortest ever recorded characteristic rise time for erbium electroluminescence. This is due to the formation of Er-related centers with an effective excitation cross section for erbium ions of ~7×10^{-16} cm^2 and an excited-state lifetime of ~17 μs. The lifetime of the first excited state of erbium ions after turning off the reverse current was measured for the first time; this lifetime is associated with Auger energy transfer to free electrons in the electrically neutral region of the diode. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Investigations of semiconductor structures doped with rare-earth elements are aimed at developing light-emitting devices for optoelectronics and fiber links. Electroluminescence (EL) associated with rare-earth elements has been observed and studied in single-crystal silicon: EL at $\lambda \approx 1.54$ μm in Si:Er, due to electron transitions from the first excited state $^4I_{13/2}$ to the ground state $^4I_{15/2}$ of Er^{3+} ions [1–4], and that in Si:Ho at $\lambda \approx 1.96$ μm, due to electron transitions from the first excited state $^5I_7$ to the ground state $^5I_8$ of Ho^{3+} ions [5]. The most intense room temperature EL from light-emitting diodes is observed under reverse bias. Two problems are to be solved to enable the practical use of erbium- or holmium-doped diodes operating in the breakdown mode—how to make the EL intensity higher and the diode faster. The former can be solved by increasing the density of excitable rare-earth ions.

The kinetics of the rise in EL associated with rare-earth ions on passing a current with density j through a light-emitting diode (LED) is characterized by the rise time τ_{on}, which can be related to the effective excitation cross section σ of the ions and their lifetime in the first excited state, τ, by the expression [2]

$$1/\tau_{on}(j) = \sigma j/q + 1/\tau, \quad (1)$$

where q is the elementary charge. In the previously studied Si:(Er,O)-based LEDs operating in the tunneling breakdown mode, $\sigma \approx 6 \times 10^{-17}$ cm^2 and $\tau \approx 100$ μs [2]. The intensity of the erbium EL in diodes of this kind practically levels off at $j \approx 200$ A/cm^2. Since the cross sections σ are small at such a current density, the τ_{on} values are relatively high (~12 μs). For $j = 50$ A/cm^2, τ_{on} is even longer (~34 μs). In Si:(Er,O) LEDs operating in the avalanche breakdown mode, $\sigma \approx 2 \times 10^{-16}$ cm^2 and $\tau \approx 380$ μs [4]. The formation of Er-related centers with large σ makes it possible to attain in this case a rise time $\tau_{on} \approx 15$ μs for $j = 50$ A/cm^2.

It was shown [1–3] that rather short decay times τ_{off} of the erbium EL can be attained owing to the Auger deexcitation of erbium ions with energy transfer to free carriers in Si:(Er, O) LEDs operating under both forward and reverse bias. However, for diodes operating in the tunneling [2] and avalanche [6] breakdown modes, the measured τ_{off} values were ~10 μs, being limited by the operating speed of the detecting system.

In this communication, the fabrication technology is described and room-temperature characteristics are reported for Si:(Er,O) LEDs operating in the tunneling breakdown mode, for which the shortest (of those recorded at fixed current density) rise times of erbium EL were observed. The room-temperature EL decay time associated with energy transfer from excited erbium ions to free electrons was measured for the first time.

2. EXPERIMENTAL

Implantation of erbium (with energies $E = 2.0$ and 1.6 MeV and dose $D = 3 \times 10^{14}$ cm^{-2}) and oxygen ions ($E = 0.28$ and 0.22 MeV and dose $D = 3 \times 10^{15}$ cm^{-2}) into polished Czochralski-grown (in the ⟨111⟩ direction) n-type Si wafers with a resistivity of 5 Ω cm was accompanied by amorphization of the near-surface layer. Ions of boron ($E = 40$ keV and $D = 5 \times 10^{15}$ cm^{-2}) and phosphorus ($E = 80$ keV and $D = 1 \times 10^{15}$ cm^{-2}) were implanted into the front and rear sides of a wafer

1063-7826/00/3408-0927$20.00 © 2000 MAIK "Nauka/Interperiodica"

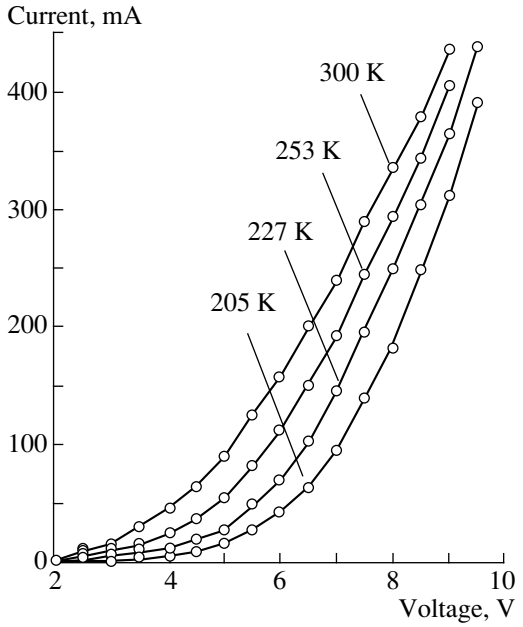


Fig. 1. Reverse *I–V* characteristics of a diode at different temperatures.

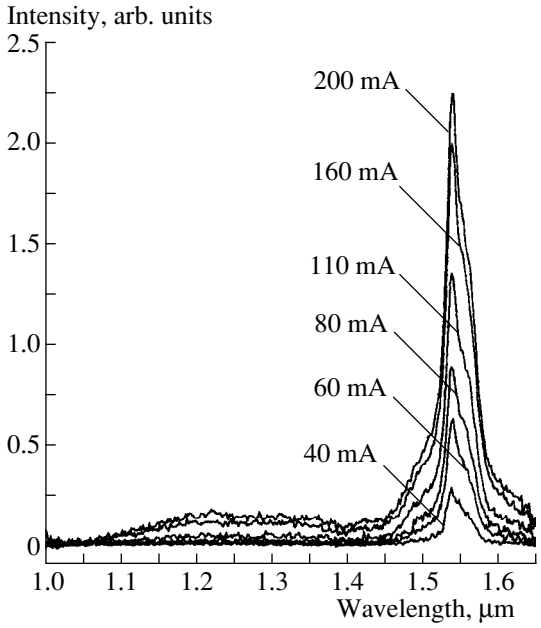


Fig. 2. EL spectra of an LED for different reverse currents at 300 K.

to form, respectively, heavily doped p^+- and n^+-layers. Annealing for 1 h at 620°C and then for 0.5 h at 900°C led to recrystallization of the amorphous layer and formation of Er-related optically and electrically active centers. According to secondary ion mass spectrometry (SIMS), the implantation and annealing gave rise to an erbium-doped silicon layer with a practically uniform erbium concentration of $\sim 1 \times 10^{19}$ cm^{-3} at distances of 0.5–0.8 μm from the semiconductor surface. The concentration of electrically active centers was determined from capacitance–voltage (*C–V*) measurements to be $\sim 1 \times 10^{18}$ cm^{-3}. According to SIMS data, the peak concentration of boron atoms was $\sim 3 \times 10^{20}$ cm^{-3}. Mesa diodes with an operating area $S = 2.3$ mm^2 were fabricated by the conventional method [4]. Current–voltage (*I–V*) characteristics of the diodes were measured at a frequency of 32 Hz and a pulse width of 0.5 ms. EL was excited with 180 μs–5 ms-long rectangular current pulses (with rise and fall times shorter than 0.1 μs) at a frequency of 30 Hz. In measuring the EL spectra, the emission of the diode structures was focused with a set of lenses onto the entrance slit of an MDR-23 monochromator and recorded at its exit with an uncooled InGaAs diode (with a resolution of 7 nm in the wavelength range $\lambda = 1.0$–1.65 μm). The kinetics of EL rise and decay were measured using a faster (but less sensitive than in the spectral measurements) diode photocurrent amplifier. To compensate for the loss of sensitivity, the radiation was directed in this case to a wide-band filter, which transmitted to the photodetector the entire spectrum of erbium ion emission associated with electron transitions from the first excited state to the ground state. The response time of the light-recording system, τ_{resp}, to a light pulse with rise and fall times $\tau_{light} \leq 0.1$ μs did not exceed 1.5 μs in this case.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The reverse *I–V* characteristics of a diode structure at different temperatures are shown in Fig. 1. An increase in the breakdown voltage, determined by extrapolating the linear portion of the *I–V* characteristic, with decreasing temperature is typical of tunnel diodes. However, the breakdown voltage at 300 K (~4.2 V) indicates that a certain contribution to reverse current may come from charge carriers generated by avalanche multiplication [7]. EL spectra measured at 300 K for reverse-biased diodes and different currents are shown in Fig. 2. In addition to the peak at $\lambda = 1.538$ μm associated with radiative transitions of electrons between the Er^{3+} ion levels $^4I_{13/2}$ and $^4I_{15/2}$ split in the crystal field, a relatively weak and nearly λ-independent emission due to transitions of "hot" electrons within the conduction band of silicon (the so-called "hot" EL) is observed within the transparency region of silicon [8].

The intensities of erbium EL at $\lambda = 1.538$ μm and hot EL at $\lambda = 1.45$ μm are shown as functions of reverse current through a diode structure at 300 K in Fig. 3. The erbium EL is observed beginning with a current of ~25 mA. The curve for hot-carrier EL can be approximated by two straight lines with threshold currents $j_{th}S \approx 25$ mA (hot-carrier EL appears) and $j_{th}S \approx 75$ mA (it starts to grow steeply in intensity).

Figure 4 illustrates the kinetics of EL rise at different amplitudes of rectangular current pulses and of EL decay after turning off a current of 0.4 A. In the presence of a single type of centers, the rise kinetics of erbium EL is typically well described by the expression [2, 4]

$$I_{Er} = I_{Er}^{m}[1 - \exp(-t/\tau_{on})], \quad (2)$$

$$1/\tau_{on} = \sigma(j - j_{th})/q + 1/\tau, \quad (3)$$

where I_{Er}^{m} is the peak EL intensity at fixed current density and t is the time elapsed after the current pulse is applied. With hot-carrier EL present, the measured EL intensity (I) is given by

$$I = I_{Si} + I_{Er}^{m}[1 - \exp(-t/\tau_{on})], \quad (4)$$

where I_{Si} is the intensity of hot-carrier EL, which is time-independent under our experimental conditions and $t \geq 1.5$ μs, since the hot-carrier EL rise time is less than 200 ns [9]. The experimental dependences shown in Fig. 4 are inadequately described by equation (4) but are well approximated by the expression

$$I(t) = I_{Si} + I_{Er}^{m1}[1 - \exp(-t/\tau_{on1})] + I_{Er}^{m2}[1 - \exp(-t/\tau_{on2})]; \quad (5)$$

i.e., the expression is written assuming that two types of independently excited Er-related centers are present in the sample, with different EL rise times (τ_{on1} and τ_{on2}), each described by formula (3), and different peak intensities I_{Er}^{m1} and I_{Er}^{m2}. It was found that the two Er-related centers are excited at different threshold current densities ($j_{th1}S \approx 25$ mA and $j_{th2}S \approx 75$ mA), while $I_{Er}^{m2} \approx 2I_{Er}^{m1}$ in the case where I_{Er}^{m1} and I_{Er}^{m2} become practically independent of j.

Figure 5 shows calculated values of $1/\tau_{on}$, the reciprocal of the EL rise time, for the two types of centers at different reverse currents. Approximation of the data in Fig. 5 by formula (3) yields $\sigma_1 \approx 8 \times 10^{-17}$ cm^2, $\tau_1 \approx 85$ μs and $\sigma_2 \approx 7 \times 10^{-16}$ cm^2, $\tau_2 \approx 17$ μs. The erbium EL for the first type of centers appears simultaneously with the hot-carrier EL, with the values of σ_1 and τ_1 close to those for the centers observed in tunnel diodes in [2] ($\sigma = 6 \times 10^{-17}$ cm^2 and $\tau = 100$ μs), for which the impact mechanism of erbium EL excitation was established [10]. This indicates that hot carriers excite centers of the first type by the impact mechanism. The fast rise in the hot-carrier EL and the appearance of erbium EL from centers of the second type at $j \geq j_{th2}$ are presumably due to the onset of avalanche multiplication of charge carriers in the p–n junction. Two mechanisms can be proposed for describing erbium ion excitation in these centers: (i) impact excitation by hot electrons appearing in the upper subband Δ_2 of the conduction band and (ii) Auger excitation via recombination of electrons and holes generated in avalanche multiplication. It should be noted that a relatively short lifetime of Er^{3+} ions in the excited state ($\tau \sim 25$ μs at ~160 K) has been observed previously for centers located within the space-charge region of a reverse-biased p–n junction [1]. Since the centers of the second type are characterized by the largest σ and shortest τ of those recorded for

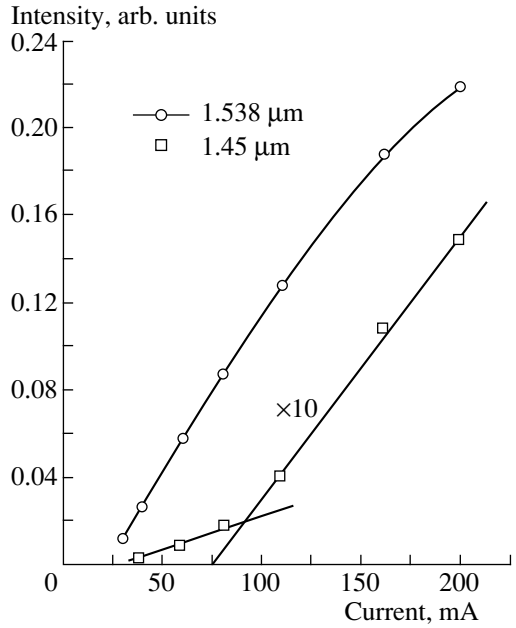


Fig. 3. EL intensity at λ = 1.538 and 1.45 μm vs. reverse current at 300 K.

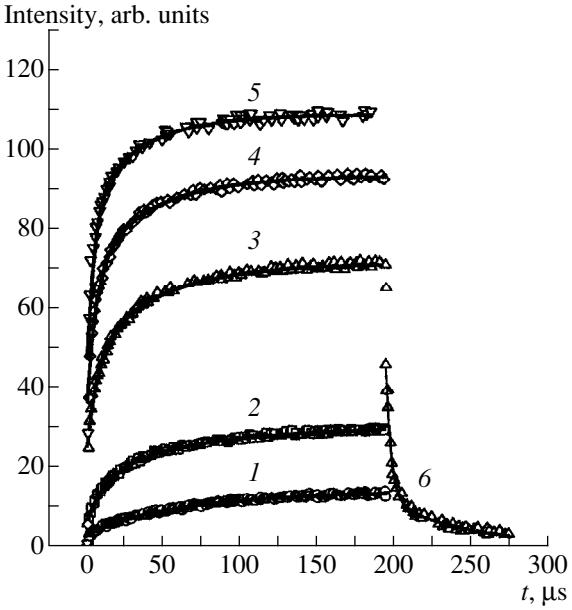


Fig. 4. Kinetics of EL rise (*1*–*5*) and decay (*6*) for the reverse currents of (*1*) 0.1; (*2*) 0.2; (*3*, *6*) 0.4; (*4*) 0.6; and (*5*) 0.8 A. Solid lines: approximation of the experimental data by formulas (5) for *1*–*5* and (6) for *6*.

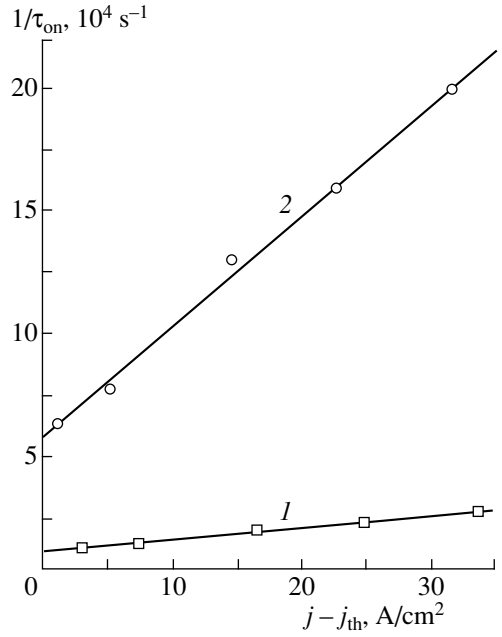


Fig. 5. Reciprocal time constant of EL intensity rise vs. the difference in the operating and threshold current densities for centers of the (*1*) first and (*2*) second types.

reverse-biased diodes, the diodes we studied exhibit the shortest rise time of the EL associated with these centers for a fixed current density. Thus, $\tau_{on} \approx 4$ μs at $j =$ 50 A/cm^2, which is much smaller than the above values for the previously studied tunnel and avalanche diodes.

The decay of the erbium EL after turning off the reverse current is adequately described by

$$I(t) = I_{Er1}\exp(-t/\tau_1) + I_{Er2}\exp(-t/\tau_2) + I_{Er3}\exp(-t/\tau_3), \quad (6)$$

where $\tau_3 \approx 3$ μs. It was found that $I_{Er3} \approx 2(I_{Er1} + I_{Er2})$. The first two terms in (6) describe the de-excitation of erbium ions that remain in the space-charge region of the *p–n* junction after the current is turned off. The appearance of the third component of erbium EL intensity decay is due to the de-excitation of erbium ions in the electrically neutral region. The most probable de-excitation mechanism is the Auger interaction of ions with free electrons [1, 2]. The rate constant of this process is defined by

$$C_A = 1/n\tau_3, \quad (7)$$

where n is the free electron concentration in the electrically neutral region. In our case, $C_A \approx 3 \times 10^{-13}$ cm^3 s^{-1} in satisfactory agreement with the Auger rate constants obtained in [1, 2], which confirms the dominant role of the Auger de-excitation mechanism.

4. CONCLUSION

Si:(Er,O) tunnel LEDs operating in the *p–n* junction breakdown mode were fabricated. Er-related centers with the largest reported effective cross sections of excitation of erbium ions and their shortest lifetimes in the excited state were observed. The characteristic rise time of erbium EL is several times shorter than the values attained previously at the same current densities. The use of a light-detecting system with a response time of $\tau_{resp} \le 1.5$ μs allowed us to measure for the first time the decay time constant of the erbium EL; this time constant is controlled by the Auger de-excitation of erbium ions with energy transfer to free electrons.

5. ACKNOWLEDGMENTS

The authors are grateful to T.M. Mel'nikova, E.O. Parshin, and A.V. Shestakov for their help in preparation of the samples.

This work was supported in part by the Russian Foundation for Basic Research, project no. 99-02-17750.

REFERENCES

1. J. Palm, F. Gan, B. Zheng, *et al.*, Phys. Rev. B **54**, 17603 (1996).
2. S. Coffa, G. Franzo, and F. Priolo, Appl. Phys. Lett. **69**, 2077 (1996).
3. F. Priolo, G. Franzo, S. Coffa, and A. Carnera, Phys. Rev. B **57**, 4443 (1998).
4. N. A. Sobolev, Yu. A. Nikolaev, A. M. Emel'yanov, *et al.*, J. Lumin. **80**, 315 (1999).
5. N. A. Sobolev, A. M. Emel'yanov, and Yu. A. Nikolaev, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 931 (1999) [Semiconductors **33**, 850 (1999)].
6. N. A. Sobolev, A. M. Emel'yanov, S. V. Gastev, *et al.*, Mater. Res. Soc. Symp. Proc. **486**, 139 (1998).
7. S. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984).
8. F. G. Chynoweth and K. G. McKay, Phys. Rev. **102**, 369 (1956).
9. W.-X. Ni, C.-X. Du, K. B. Joelsson, *et al.*, J. Lumin. **80**, 309 (1999).
10. S. Coffa, G. Franzo, F. Priolo, *et al.*, Appl. Phys. Lett. **73**, 93 (1998).

Translated by M. Tagirdzhanov

Semiconductors, Vol. 34, No. 8, 2000, pp. 931–936. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 970–975.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Distribution of Mobile Ions in Thin Insulator Films at the Insulator–Semiconductor Interface

S. G. Dmitriev* and Yu. V. Markin

Institute of Radio Engineering and Electronics (Fryazino Branch), Russian Academy of Sciences, pl. Vvedenskogo 1, Fryazino, Moscow oblast, 141120 Russia
** e-mail: sgd117@ire216.msk.su*

Submitted September 2, 1999; accepted for publication February 16, 2000

Abstract—The influence of a semiconductor on the mobile ion distribution in insulator thin films at the insulator–semiconductor interface was considered. The degree of ion localization at the interface under the field effect in the film was calculated. The threshold of the ion delocalization with a decrease in the voltage applied to the structure was determined. The relation between the delocalization thresholds and ion current peaks in dynamic current–voltage characteristics of the system is discussed. © *2000 MAIK "Nauka/Interperiodica".*

Ions that are mobile in thin insulator films ($d \sim$ 100–1000 Å) at elevated temperatures ($T \sim$ 300–600 K) are topical due to their influence on the parameters of these structures. In the Si–SiO_2 structure, these are usually the alkali metal ions, specifically sodium ions, which become mobile even at room temperature. The ion concentration in the film (per unit area) may be as high as $N_s \sim 10^{12}$–10^{13} cm^{-2}, and they can induce both fast electronic states at the interface and traps in an insulator [1].

Information about the degree of ion localization at the interface under the influence of the field, which keeps the ions close to the interface, is necessary to study these questions [1]. In addition, data on the ion capture by the surface as well as on the ion escape (emission) into insulator are also required (see, for example, [2, 3]). The ion distribution in the film for the metal–insulator–metal structures was considered using a hydrodynamic approximation in terms of ion concentration [4, 5].

In this study, the ion distribution in the metal–insulator–semiconductor (MIS) structures was analyzed using the Boltzmann ion-energy distribution. The ion delocalization thresholds with a decrease in the voltage applied to the structure, as well as the influence of the semiconductor on these thresholds, were determined. Conditions for degeneracy of the ion energy distribution at the interface are considered.

1. ION DISTRIBUTION IN THE FILM

Potential $\varphi(z)$ in the film for the single-charged positive ions is defined by the Poisson equation

$$\varepsilon_i \frac{d^2\varphi}{dz^2} = -4\pi q N_i(z), \quad (1)$$

where ε_i is the dielectric constant for insulator, q is the elementary charge, and $N_i(z)$ is the ion concentration. In order to determine the ion distribution, the requirement that the current density $j_z(z)$ should equal zero in equilibrium conditions is usually used:

$$j_z = q\mu\mathscr{E}_z N_i - qDdN_i/dz, \quad (2)$$

where μ is the mobility, D is the diffusion coefficient, and $\mathscr{E}$ is the field. Here, the Einstein relation ($\mu = qD/kT$, where k is the Boltzmann constant), which is a consequence of the Boltzmann particle-energy distribution, is taken into account.

When describing the electron subsystem, another equivalent approach is often used, in which the electron density is described directly in terms of the Boltzmann or Fermi distributions. In this case, additional (and commonly known) data on the effective density and spectrum of electron intrinsic and defect-related states are required.

For ions, this information is less readily available. However, ions in solids are usually highly localized and their motion involves hops between the neighboring sites. This situation corresponds to a narrow energy range, and we consider the model with one single-energy ion level E_i as a first step:

$$N_i(z) = N_0 \exp(-u), \quad u = (E_i + q\varphi - F_i)/kT, \quad (3)$$

where u is the normalized ion energy measured from the Fermi level F_i and N_0 is the effective density of states, whose magnitude is discussed below.

The form of the solution to equation (1) for the potential (or energy u) depends on the sign of its first integral I_0:

$$I_0(z) = \varepsilon_i\mathscr{E}^2(z)/(8\pi kT) - N_i(z) = \text{const} > 0, \quad (4)$$

where our topical case of ion localization at the inner insulator–semiconductor interface ($z = d$) corresponds to $I_0 > 0$.

For a given surface ion density in the film N_s, the form of distribution $\varphi(z)$

$$\varphi(z) - \varphi_0 = kT[u(z) - u_0]/q,$$
$$u(z) - u_0 = \ln\{\sinh^2[A(z)]/\sinh^2\alpha\} \quad (5)$$

is defined by the parameter α:

$$\alpha = \ln[(\beta^2 - 1)^{1/2} + \beta], \quad \beta = q\mathscr{E}_0 r_0/kT = \cosh\alpha;$$
$$A(z) = \alpha - q\mathscr{E}_0 z\tanh\alpha/(2kT), \quad (6)$$
$$r_0^2 = \varepsilon_i kT/[8\pi q^2 N_0 \exp(-u_0)]. \quad (7)$$

Here, the subscript 0 for φ_0, u_0, $\mathscr{E}_0$, and r_0 is related to the outer insulator boundary ($z = 0$), and r_0 is the Debye radius. From the condition $I_0 > 0$ it follows that $\beta > 1$ and $A(z) > 0$ (for $0 \le z \le d$), while the condition itself means that the field-pulling length $l_{\mathscr{E}} = kT/(q\mathscr{E})$ is invariably less than the Debye radius r. In this case, screening is substantially nonlinear.

To calculate α, we can use the equality

$$I_0(0) = I_0(d), \quad (8)$$
$$\mathscr{E}_1 = \mathscr{E}_0 + \mathscr{E}_i, \quad \mathscr{E}_i = 4\pi q N_s/\varepsilon_i, \quad (9)$$

where (9) is a consequence of the Gauss theorem for a film, subscript 1 is related to the inner insulator surface ($z = d$), and $\mathscr{E}_i$ is the field of ions. In terms of α, equation (8) takes the form

$$\exp(\Delta u) - 1 = [(1 + \mathscr{E}_i/\mathscr{E}_0)^2 - 1]\cosh^2\alpha, \quad (10)$$

where $\Delta u = u_0 - u_1$. Using (10) and (5) (the latter is considered at $z = d$), we can derive the equation for α at given N_s and $\mathscr{E}_0$:

$$\alpha = (\mathscr{E}_0/\mathscr{E}_d)\tanh\alpha + \frac{1}{2}\ln\left(\frac{1 + \mathscr{E}_i/\mathscr{E}_0 + \tanh\alpha}{1 + \mathscr{E}_i/\mathscr{E}_0 - \tanh\alpha}\right), \quad (11)$$
$$\mathscr{E}_d = 2kT/(qd), \quad (12)$$

where $\mathscr{E}_d$ is the characteristic field in the film. Other quantities ($\Delta\varphi = \varphi_0 - \varphi_1$, u_0, and u_1) can be found from equations (10) and (6), while the distribution $N_i(z)$ can be found from equations (4) and (8).

Notice that the ion concentration $N_i(z)$ and the voltage drop across the film $\Delta\varphi$ and, consequently, the voltage V_g for a system as a whole are actually independent of the effective density of states N_0. In fact, it follows additionally from (3) that the variation in N_0 always can be compensated by the shift (renormalization) of the Fermi level. We discuss the possibility of violating this statement below, after calculating $N_i(z)$.

The (degenerate) case of $\alpha = 0$ corresponds to $\mathscr{E}_0 = \mathscr{E}_{\text{th}}$. Using (11) for $\alpha \longrightarrow 0$, we obtain

$$\mathscr{E}_{\text{th}} = \frac{1}{2}\mathscr{E}_i[-1 + (1 + 4\mathscr{E}_d/\mathscr{E}_i)^{1/2}], \quad (13)$$
$$(\Delta u)_{\text{th}} = 2\ln(1 + \mathscr{E}_i/\mathscr{E}_{\text{th}}). \quad (13a)$$

In this case, the solution is degenerate; i.e.,

$$u(z) - u_0 = 2\ln[1 - (\mathscr{E}_{\text{th}}/\mathscr{E}_d)(z/d)].$$

Typically, $\mathscr{E}_{\text{th}} \approx \mathscr{E}_d$. Actually, let us carry out the estimations. For $kT = 50$ meV, $d = 1000$ Å, $N_s = 10^{12}$ cm^{-2}, and $\varepsilon_i = 3.9$ (for SiO_2 [7]), $\mathscr{E}_d \approx 0.86 \times 10^4$ V/cm and $\mathscr{E}_i \approx 4.6 \times 10^5$ V/cm, so that $\lambda = \mathscr{E}_d/\mathscr{E}_i \ll 1$. This is the most interesting case experimentally ($\lambda \ll 1$) [1], and we consider it below. In this case, $\mathscr{E}_{\text{th}} \approx \mathscr{E}_d(1 - \lambda)$, i.e., $\mathscr{E}_{\text{th}} \approx \mathscr{E}_d$. In the vicinity of $\mathscr{E}_{\text{th}}$, we have $\alpha^2 \approx 3(\mathscr{E}_0 - \mathscr{E}_{\text{th}})/\mathscr{E}_d$ (see A.1.2).

The degree of ion localization at the interface is defined by the ion–interface average distance $\delta z_i = d - \bar{z}_i$, where

$$N_s\bar{z}_i = \int_0^d zN_i(z)dz.$$

This distance can be conveniently expressed using the Poisson equation as

$$\delta z_i = (\Delta\varphi - \mathscr{E}_0 d)/\mathscr{E}_i; \quad (14)$$

i.e., the degree of ion delocalization is defined by a nonlinear component of the potential distribution.

In the case of strong fields ($\alpha \gg 1$ and $\mathscr{E}_0 \gg \mathscr{E}_d$), α is defined by formula (11) for $\tanh\alpha = 1$; thus, we have

$$\Delta u \approx q\mathscr{E}_0 d/(kT) + 2\ln[1 + \mathscr{E}_i/(2\mathscr{E}_0)], \quad (15)$$
$$\delta z_i \approx l_i\ln[1 + \mathscr{E}_i/(2\mathscr{E}_0)], \quad (16)$$
$$l_i = 2kT/(q\mathscr{E}_i). \quad (17)$$

The quantity $l_i \sim 10$ Å ($l_i \approx 20$ Å at $kT = 50$ meV, $N_s = 10^{12}$ cm^{-2}). The case of $\delta z_i \ll l_i$ is possible for $\mathscr{E}_0 \gg \mathscr{E}_i$, if $\delta z_i \approx kT/(q\mathscr{E}_0) \longrightarrow 0$. However, this case is unrealistic. In addition, the near-interface damaged layers are about 10 Å thick [6]. This also restricts the possibility for stronger localization.

It is noteworthy that the $\delta z_i(\mathscr{E}_0)$ dependence for the region $\mathscr{E}_0 \ll \mathscr{E}_i$ is also weak and $\delta z_i \sim 100$ Å only if $\mathscr{E}_i/\mathscr{E}_0 \sim 100$. In this case, the ions are kept at the boundary by the field of electrons of the semiconductor rather than the external field.

In experiments with MIS structures, the dependences of currents or capacitances on $V_g = \Delta\varphi(\mathscr{E}_0) + \varphi_s(\mathscr{E}_0)$, where φ_s is the interface potential measured from the bulk of the semiconductor, are recorded. Actually, these dependences are shifted by the magnitude of

the metal–semiconductor contact potential $\varphi_c \sim (0.1–1)$ V [7, 8]. The V_g-dependence of φ_s is considered in the following section.

2. BAND BENDING IN A SEMICONDUCTOR

A band bending in a semiconductor is defined by its surface field $\mathscr{E}_s = (\varepsilon_i\mathscr{E}_0 + 4\pi qN_s)/\varepsilon_s$, where ε_s is the dielectric constant of semiconductor and can be found using the first integral of the Poisson equation:

$$\mathscr{E}_s^2 = (8\pi kT/q\varepsilon_s)\int_0^{v_s}\rho(v)d(v). \qquad (18)$$

Here, $v_s = -q\varphi_s/kT < 0$ and $\rho(v)$ is the charge density for the semiconductor. Let us consider the *n*-type semiconductor with the concentration N_D and the ionization energy E_d of donors.

At elevated temperatures ($T \sim 600$ K), the semiconductor becomes intrinsic and it is necessary to take into account the hole charge. In the case of high fields $\mathscr{E}_s \sim 10^5–10^6$ V/cm at the surface, it is necessary to take into account the degeneracy of electrons at the donor levels and in the conduction band. In this case [7],

$$\rho(v) = qN_D[1 + (g_1/g_0)\exp(\varepsilon_d + f_0 - v)]^{-1} + qN_v\exp(v - \varepsilon_g - f_0) - qN_c\Phi_{1/2}(f_0 - v), \qquad (19)$$

where $\Phi_{1/2}$ is the Fermi–Dirac integral, $\varepsilon_d = E_d/kT$, $\varepsilon_g = E_g/kT$, $f_0 = F_e/kT$, E_g is the band gap, F_e is the electron Fermi level measured from the conduction band bottom, $N_c(N_v)$ is the density of states in the conduction (valence) band, and $g_0(g_1)$ are the degeneracy factors for the unoccupied (occupied) levels.

For silicon [8], $\varepsilon_i = 3.9$, $\varepsilon_s = 11.9$, $N_c = 2.8 \times 10^{19}(T/300)^{3/2}$ cm^{-3}, $N_v = 1.04 \times 10^{19}(T/300)^{3/2}$ cm^{-3}, $E_g = 1.12$ eV, $E_d = 0.045$ eV (for the phosphorus atom), $g_0 = 1$, and $g_1 = 2$.

For these parameters and $N_D \sim 10^{15}$ cm^{-3}, we have $|f_0| \gg 1$ (see Appendix 2): $|f_0|(500\text{ K}) \approx 11$ (0.47 eV); $|f_0|(300\text{ K}) \approx 10.2$ (0.26 eV). In this case, for the actual temperature range ($T \sim 200–600$ K), the occupancy of the donor level in the bulk semiconductor can be neglected, while the influence of holes manifests itself only at $T \sim 500–600$ K.

On determining $\varphi_s(\mathscr{E}_0)$, the potential distribution in the film $\varphi(z)$ can be found from the formula

$$\varphi(z) = \varphi_s + kT[u(z) - u_1]/q;$$

specifically, $V_g = \varphi(d)$. Notice that $kTu_1/q \neq \varphi_s + |f_0|$.

The dependence $\varphi_s(V_g)$ for $T = 500$ K, which was obtained in this manner, is shown in Fig. 1. It can be seen from comparison between φ_s and $|f_0|$ that the degeneracy at the surface manifests itself even at these temperatures.

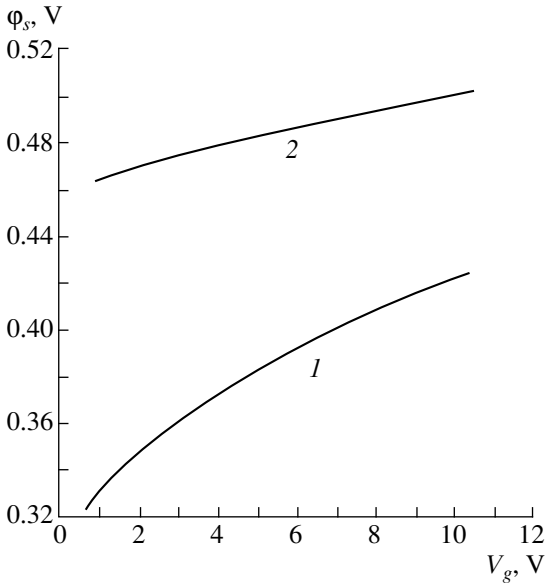


Fig. 1. Band bending φ_s in a semiconductor. V_g is the voltage; $N_s = (1)\ 10^{12}$ and $(2)\ 5 \times 10^{12}$ cm^{-2}; $T = 500$ K, $N_D = 10^{15}$ cm^{-3}, $d = 1000$ Å.

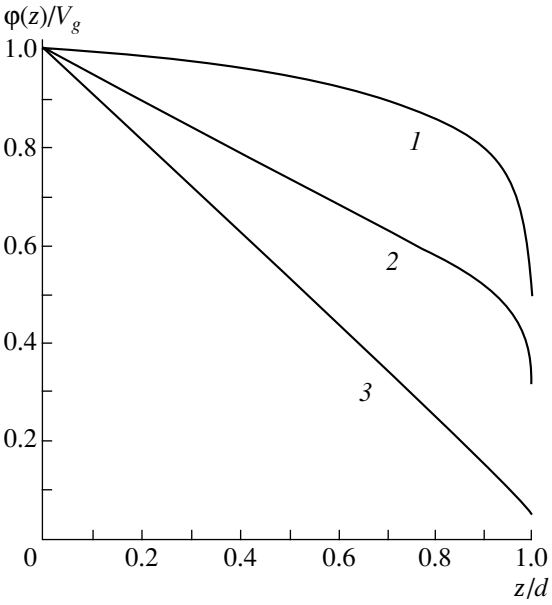


Fig. 2. Potential distribution $\varphi(z)$ in the film. V_g is the voltage; $V_g = (1)\ 0.95$, $(2)\ 1.5$, and $(3)\ 10$ V; $N_s = 5 \times 10^{12}$ cm^{-2}, $T = 500$ K, $N_D = 10^{15}$ cm^{-3}, $d = 1000$ Å.

Normalized (over V_g) potential distribution in the film for several values of V_g is shown in Fig. 2. For large V_g, $\varphi_s \sim (0.01–0.1)V_g$ is relatively small, and, for $V_g \sim 1$ V, $\varphi_s \sim \Delta\varphi$ and the influence of semiconductor is significant.

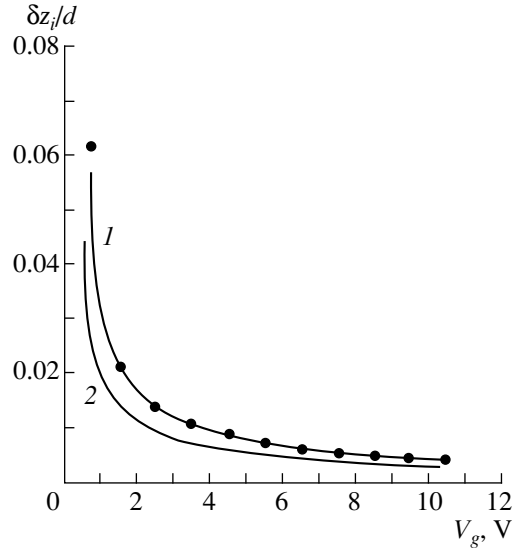


Fig. 3. Average ion distance from the insulator–semiconductor boundary δz_i. V_g is the voltage; T = (*1*) 500 and (*2*) 350 K; $N_s = 10^{12}$ cm^{-2}; $N_D = 10^{15}$ cm^{-3}; and d = 1000 Å. The points represent calculation with formula (16).

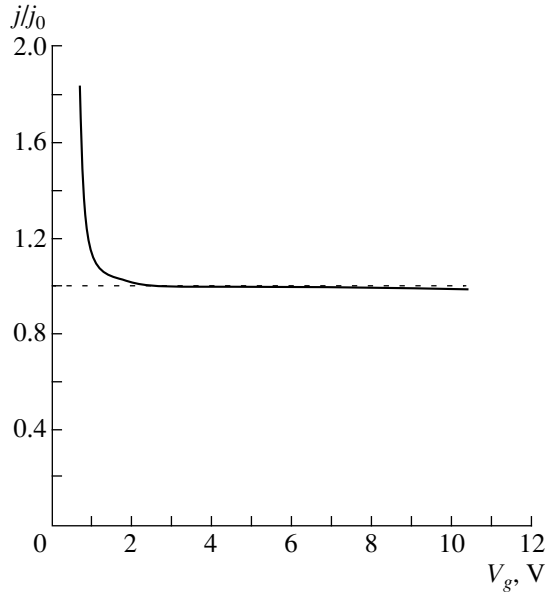


Fig. 4. Quasi-static current–voltage characteristic near the delocalization threshold. V_g is the voltage; $N_s = 10^{12}$ cm^{-2}; $N_D = 10^{15}$ cm^{-3}; d = 1000 Å; and T = 500 K. The dashed line is the capacitive current j_0 for the insulator film.

3. ION DELOCALIZATION THRESHOLDS IN THE FILM AT THE INTERFACE

The distance of the ion-distribution centroid from the interface δz_i as a function of V_g is shown in Fig. 3. A threshold increase in the ion shifts (delocalization) is observed in the voltage region of $V_g \sim 1$ V. In this case, the threshold fields $\mathscr{E}_0 \sim \mathscr{E}_d$. The points in Fig. 3 correspond to calculations with formula (16). The distinctions (~10%) are noticeable only at the threshold.

Experimentally, this phenomenon manifests itself as a threshold rise of the quasi-static ion currents in the dynamic current–voltage (*I–V*) characteristic. The current density in the MIS structure can be estimated from the formula $j = d\sigma/dt$, where $\sigma = \varepsilon_i \mathscr{E}_0/4\pi$ is the surface charge density at a metal. For linear voltage sweep ($\partial V_g/\partial t = \beta_v$),

$$j = (\varepsilon_i \beta_v/4\pi)\partial\mathscr{E}_0/\partial V_g,$$
$$\partial\mathscr{E}_0/\partial V_g = (1 - \partial\varphi_s/\partial V_g)\partial\mathscr{E}_0/\partial(\Delta\varphi). \quad (20)$$

Under enrichment conditions ($\varphi_s \gg kT$), the value $\partial\varphi_s/\partial V_g \sim 0.01$–$0.1$ (see Fig. 1), while, in strong fields ($\mathscr{E}_0 \gg \mathscr{E}_d$), we have $\partial\mathscr{E}_0/\partial(\Delta\varphi) \approx d^{-1}$. Then the current $j \approx j_0$, where

$$j_0 = C_0\beta_v, \quad C_0 = \varepsilon_i/(4\pi d), \quad (21)$$

and C_0 is the insulator capacitance per unit area (see A.1.4–5). For β_v = 20 mV/s and d = 1000 Å, we have $j_0 \approx 7 \times 10^{-10}$ A/cm^2.

The ratio j/j_0 is independent of β_v. The result of calculating this ratio is shown in Fig. 4. As is seen from Figs. 3 and 4, a threshold rise in currents corresponds to the ion delocalization threshold and $j \approx 2j_0$ for $\alpha = 0$ (see A.1.6–7). However, the curve shape is conditioned mostly by a current rise in a narrow voltage region corresponding to the field range of $\mathscr{E}_0 \sim \mathscr{E}_d \sim 0.1\mathscr{E}_d$ rather than by an increase in current itself (see Appendix 1). Thus, this phenomenon is related to the variation in the characteristics of field screening in the film for a given range of parameters, i.e., with a change in the dependence $\Delta\varphi(\mathscr{E}_0)$. A more detailed investigation of ionic peaks, which is required for additional mathematical treatment, should be considered separately.

4. THE CONCENTRATION AND ENERGY DISTRIBUTION OF IONS AT THE INTERFACE

Near the interface, ion concentration $N_i(d)$ can be determined from formula (8) taking into account formulas (3), (4) and (9):

$$N_i(d) = N_i^0(1 + 2\mathscr{E}_0/\mathscr{E}_i)[1 - \exp(-\Delta u)]^{-1}, \quad (22)$$

$$N_i^0 = 2\pi q^2 N_s^2/(\varepsilon_i kT), \quad (23)$$

where the characteristic concentration $N_i^0 \approx 0.94 \times 10^{19}$ cm^{-3} (kT = 25 meV and $N_s = 10^{12}$ cm^{-2}), and $N_i^0 \approx 2.35 \times 10^{20}$ cm^{-3} ($N_s = 5 \times 10^{12}$ cm^{-2}). As was mentioned above, $N_i(d)$ is independent of the effective density of states N_0. The voltage dependence of $N_i(d)$ is shown in

Fig. 5. The points correspond to calculations with the formula

$$N_i(d) = N_i^0(1 + 2\mathscr{E}_0/\mathscr{E}_i), \quad (22a)$$

which follows from (22), if we neglect $\exp(-\Delta u)$ (at $\Delta u \gg 1$).

When ions diffuse via the interstices, $N_0 \sim a_0^{-3}$, where a_0 is the unit cell size for a crystal ($a_0 = 3.3$ Å for SiO_2 [9]), i.e., $N_0 \sim 10^{22}$–10^{23} cm^{-3}. Since the characteristic values $N_i(d) \sim 10^{19}$–10^{21} cm^{-3}, $u_1 = -\ln[N_i(d)/N_0)] \sim$ 4–5 or more and degeneracy is absent even at the surface.

In this case, the system can be described in terms of ion concentration (hydrodynamic approximation) without direct consideration of the density of states. The ion energy distribution thereby automatically becomes a Boltzmann distribution (taking into account the Einstein relation).

However, if one more ionic state with a larger binding energy E_2 exists at the surface,

$$E_1 - E_2 > kTu_1, \quad (24)$$

the level of this state becomes lower than the Fermi level for ions. In this case, it is necessary to take into account the degeneracy effects and to use the Gibbs distribution, taking into account that only a single ion can occupy a single lattice site.

Ionic states can be induced by the interface and by lattice defects. The defect concentration near the Si–SiO_2 interface is relatively high (about the atomic one) due to the lattice mismatch [10]. Defect ionic states also exist in the bulk of an insulator [11]. Information about the spectra of ionic states is usually unavailable. However, the characteristic values of ion bonding energy in the lattice and activation energies of ionic processes are on the order of 0.1–1 eV or more. Since kTu_1 from (24) may be as large as ~0.1–0.2 eV, this condition can be satisfied. This causes accumulation of ions at the states under consideration.

In conclusion, note that the potential distribution in MIS structures with mobile ions in insulator was calculated in this study (Fig. 2). It is demonstrated that, at strong fields on the order of the ionic field $\mathscr{E}_i$ (see (9)), the ions are localized at a distance on the order of $l_i \sim$ 10 Å from the interface (see (16) and (17)). However, their drift to larger distances (~100 Å) with the voltage change occurs in weak fields $\sim\mathscr{E}_d \ll \mathscr{E}_i$ (see (12)) at voltages ~1 V and is described by a threshold function (Fig. 3). At the threshold, the ions are kept near the interface by the field of electrons in the semiconductor rather than by the external field. This process corresponds to a threshold increase in ionic currents in quasi-static *I–V* characteristics of the system (Fig. 4).

In this study, the volume ion concentration at the interface is also determined (see (22), (23), and Fig. 5). The conditions for applicability of a hydrodynamic approximation for description of spatial ion distribution and the possibility of ion energy degeneracy at the interface are discussed (see (24)).

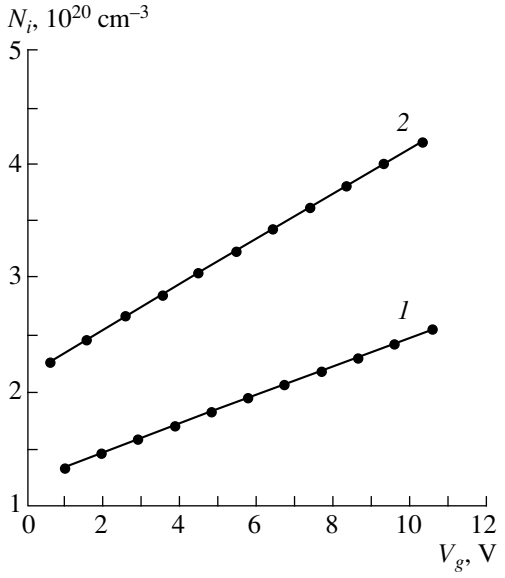


Fig. 5. Ion concentration N_i at the insulator–semiconductor boundary. V_g is the voltage; T = (*1*) 500 and (*2*) 300 K; $N_s = 5 \times 10^{12}$ cm^{-2}; $N_D = 10^{15}$ cm^{-3}; and d = 1000 Å. The points represent calculation with formula (22a).

APPENDIX 1

For small $\alpha \ll 1$, equation (11) can be represented (accurate to the terms $\sim\alpha^2\lambda^3$, where $\lambda = \mathscr{E}_d/\mathscr{E}_i$) in the form $\alpha \approx (1 + x)\tanh\alpha$, where

$$x = \mathscr{E}_0/\mathscr{E}_d - 1 + (1 + \mathscr{E}_i/\mathscr{E}_0)^{-1}. \quad (A.1.1)$$

Its solution for small x has the form

$$\alpha^2 \approx 3x\left(1 + \frac{1}{5}x\right). \quad (A.1.2)$$

Similarly, it is possible to obtain the expression for the derivative $\partial(\alpha^2)/\partial\mathscr{E}_0$ for $\alpha = 0$:

$$\partial(\alpha^2)/\partial\mathscr{E}_0 = (3/\mathscr{E}_d)(1 + \gamma_1^2\gamma_2)/(1 - \gamma_1^3), \quad (A.1.3)$$

where $\gamma_1 = \mathscr{E}_{th}/(\mathscr{E}_i + \mathscr{E}_{th})$ and $\gamma_2 = \mathscr{E}_d\mathscr{E}_i/(\mathscr{E}_{th})^2$.

For strong fields ($\mathscr{E}_0 \gg \mathscr{E}_d$), the expression for the derivative $\partial\mathscr{E}_0/\partial(\Delta\varphi)$ can be obtained from formula (15):

$$\partial\mathscr{E}_0/\partial(\Delta\varphi) \approx d^{-1}(1 - \gamma_3)^{-1}, \quad (A.1.4)$$

where $\gamma_3 = \mathscr{E}_d\mathscr{E}_i/[(2\mathscr{E}_0 + \mathscr{E}_i)\mathscr{E}_0]$, while the expression for a current, from formula (20), is

$$j \approx j_0(1 - \partial\varphi_s/\partial V_g)(1 - \gamma_3)^{-1}. \quad (A.1.5)$$

The parenthesized factors $\partial\varphi_s/\partial V_g$ and γ_3 are of the same order of magnitude (~0.01), and their effects on a

current are mutually compensated. In this case, $j \approx j_0$. For strong fields, $\Delta j = j - j_0 < 0$.

For $\alpha = 0$, using formulas (A.1.2) and (A.1.3) and neglecting the relatively small terms $\sim\lambda^2$, we derive

$$\partial\mathscr{E}_0/\partial(\Delta\varphi) \approx 2d^{-1}(1 + 3\lambda)^{-1}, \quad \text{(A.1.6)}$$

$$j \approx 2j_0(1 - \partial\varphi_s/\partial V_g)(1 + 3\lambda)^{-1}. \quad \text{(A.1.7)}$$

Here, $j \approx 2j_0$ ($j < 2j_0$). However, the distinctions between j and $2j_0$ may be as large as ~(1–10)%.

APPENDIX 2

Integral $J = \int_0^{\nu_s} \rho(\nu)d(\nu)$ from formula (18) with $\rho(\nu)$ from formula (19) can be written as

$$\begin{aligned} J = qN_D\{\nu_s + \ln[1 + (g_1/g_0)\exp(\varepsilon_d + f_0 - \nu_s)] \\ - \ln[1 + (g_1/g_0)\exp(\varepsilon_d + f_0)]\} \\ + qN_\nu\exp(-\varepsilon_g - f_0) \\ \times(-1 + \exp\nu_s) - qN_c[\Phi_{3/2}(f_0) - \Phi_{3/2}(f_0 - \nu_s)], \end{aligned} \quad \text{(A.2.1)}$$

$$\Phi_\nu(x) = [1/\Gamma(\nu + 1)]\int_0^\infty \frac{z^\nu dz}{1 + \exp(z - x)},$$

where $\Gamma(\nu)$ is the gamma function. The last term can be derived by integrating first with respect to the potential and then by integrating by parts with respect to energy.

The Fermi level position f_0 in the bulk of the semiconductor is defined by the condition of electroneutrality $\rho(0) = 0$. However, for $\lambda_1 \ll 1$, where

$$\lambda_1 = (g_1/g_0)(N_D/N_c)\exp\varepsilon_d, \quad \text{(A.2.2)}$$

we may neglect the electron degeneracy at the donor level ($\exp(\varepsilon_d + f_0) \ll 1$), and even more so in the conduction band. In this case, since all donors are charged, the equation for f_0 is simplified, and we can derive (in the context of the Boltzmann statistics for electrons and holes)

$$\begin{aligned} f_0 = \ln(\delta N_D/N_c), \\ \delta = \{1 + [1 + \lambda_2^2(1 + \lambda_1)]^{1/2}\}/[2(1 + \lambda_1)], \end{aligned} \quad \text{(A.2.3)}$$

where $\lambda_2 = 2n_i/N_D$, $n_i^2 = N_cN_\nu\exp(-\varepsilon_g)$, n_i is the intrinsic concentration, and λ_2 defines the extent of the hole influence on f_0. Thus, the donor level occupancy is negligible for low-density shallow levels ($\lambda_1 \ll 1$) up to $T \sim 200$ K. The influence of holes manifests itself ($\lambda_2 \sim 1$) only at high temperatures $T \sim 600$ K and low concentrations $N_D \sim 10^{14}$ cm^{-3}.

On determining f_0, ν_s can be found from (18) using (A.2.1).

REFERENCES

1. T. Ando, A. B. Fowler, and F. Stern, in *Electronic Properties of Two-Dimensional Systems* (Am. Phys. Soc., New York, 1982; Mir, Moscow, 1985).
2. S. G. Dmitriev and Yu. V. Markin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 1231 (1996) [Semiconductors **30**, 649 (1996)].
3. S. G. Dmitriev and Yu. V. Markin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 1439 (1998) [Semiconductors **32**, 1284 (1998)].
4. A. G. Tangena, J. Middelhoek, and N. F. de Rooij, J. Appl. Phys. **49**, 2876 (1978).
5. V. P. Romanov and Yu. A. Chaplygin, Phys. Status Solidi A **53**, 493 (1979).
6. T. Bechshtedt and R. Enderlein, *Semiconductor Surfaces and Interfaces: Their Atomic and Electronic Structures* (Academie-Verlag, Berlin, 1988; Mir, Moscow, 1990).
7. V. L. Bonch-Bruevich and S. G. Kalashnikov, *The Physics of Semiconductors* (Nauka, Moscow, 1990).
8. S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984).
9. V. A. Gritsenko, *Atomic and Electronic Structure of Amorphous Insulators in Silicon MIS Devices* (Nauka, Novosibirsk, 1993).
10. F. J. Himpsel, F. R. McFeely, A. Taled-Ibrahimi, *et al.*, Phys. Rev. B **38**, 6084 (1988).
11. D. K. Belashchenko, Usp. Fiz. Nauk **169**, 361 (1999).

Translated by N. Korovin

Semiconductors, Vol. 34, No. 8, 2000, pp. 937–938. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 976–977.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

The Dislocation Nature of a Tunneling Excess Current in GaAs–Ni Structures Modified by Laser Radiation

K. K. Dzhamanbalin and A. G. Dmitriev

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

Submitted March 17, 1999; accepted for publication February 17, 2000

Abstract—The analysis of the key factors governing the formation of point and linear defects in crystals exposed to pulsed laser radiation showed that the concentration of intrinsic defects formed in actual experimental practice is insufficient to ensure multistage tunneling of electrons through a thick space-charge layer. It is concluded that dislocations play a decisive role in the origination of a tunneling excess current. This conclusion is also supported by the fact that the behavior of a tunneling excess current under repeated irradiation of the structures studied is similar to the features of plastic deformation under repeated loading. © *2000 MAIK "Nauka/Interperiodica".*

1. It is known that the excess current in light-emitting and photosensitive devices based on *p*–*n*- and metal–semiconductor structures irreversibly increases in the course of their service, thus impairing the parameters of these devices. Numerous studies have shown that the excess current exhibits tunneling behavior even if the thickness of a space-charge layer is too large for direct tunneling. Several attempts have been made to clarify the nature of the excess tunneling, but in spite of numerous studies this problem still remains unresolved. Interest in this problem occasionally reappears whenever new materials with barrier structures become the subjects of study. Such a situation exists at present [1].

In our opinion, the difficulties in clarifying the nature of a tunneling excess current are related to the fact that experimentalists are restricted in the possibilities of controlling it with external effects.

The excess current has been observed in GaAs–Ni structures with a thick (up to 0.5 μm) space-charge layer after exposure to pulsed laser irradiation of millisecond duration [2, 3]. The excess current was also found to be irreversible. Analysis of the current–voltage characteristics and the temperature variations of their parameters made it possible to conceive the excess current as a tunneling one [3]. The current can be controllably increased by irradiating with pulses of high intensities. After the subsequent irradiation with pulses of the same or lower intensity, the excess current remains unchanged [2]. It is shown in [3] that tunneling through the thick space-charge layer is a multistage process and proceeds through the local states within a band gap. However, the nature of the centers responsible for the tunneling and the features of the behavior of the tunneling excess current after repeated irradiation have not been discussed and remain unclear.

In our opinion, the occurrence of a tunneling excess current after laser irradiation and its controllable variations, as well as the behavior of this current after repeated radiation exposures, provide obvious evidence about its origin. Let us consider the features of pulsed laser irradiation.

2. It is known that there are the following basic factors in pulsed laser irradiation producing irreversible changes in a crystal: (i) the high densities of the input energy; (ii) the high heating and cooling rates of a subsurface layer; and (iii) the existence of a temperature gradient.

The first two factors result in a large increase in the concentration of intrinsic defects in a crystal. The last factor (the temperature gradient) gives rise to stresses due to the thermal expansion of the crystal. The stresses, in turn, lead to generation and irreversible multiplication of linear defects (dislocations). Both the intrinsic defects and dislocations can be responsible for the irreversible generation and increase of a tunneling excess current.

3. First, let us consider the intrinsic defects and try to estimate the concentration necessary to ensure tunneling through a thick space-charge layer. According to [3], the number of electron hops δ occurring in the course of multistage tunneling can be estimated from the thickness of space-charge layer W and the tunneling depth l_t as

$$\delta = W/l_t. \quad (1)$$

For the structures studied [3], W = 0.3–0.4 μm, while, for the parameters of the energy barrier typical of these structures, $l_t \approx 0.01$ μm [4]. This means that the number of electron hops in the course of multistage tunneling is equal to δ = 30–40. If these hops occur over

1063-7826/00/3408-0937$20.00

the local states, it is easy to estimate their irreversible concentration n_t in a space-charge layer:

$$n_t = l_t^{-3}, \quad (2)$$

which, for the above estimated value of l_t, yields $n_t \approx 10^{18}$ cm^{-3}.

We can now estimate the concentration of intrinsic defects n^* in actual experimental conditions. This can be done from the following considerations. A crystal exposed to pulsed radiation can be considered to be locally in a state of thermodynamic equilibrium for pulses with widths up to those comparable with the period of atom vibrations in a crystal lattice. This means that the concentration of intrinsic defects at each site in a crystal is determined by its local temperature.

By the end of the irradiation pulse, the temperature T_h reaches the highest value determined by the irradiation intensity. This temperature corresponds to the equilibrium concentration n_{eq} of intrinsic defects given by

$$n_{eq} = n_i \exp(-\Delta H/kT_h), \quad (3)$$

where n_i is the concentration of the intrinsic atoms and ΔH is the enthalpy of the formation of native defects. According to the data reported [5] for GaAs, $\Delta H(V_{Ga}) = \Delta H(V_{As}) = 2.3$ eV, where V_{Ga} and V_{As} are the gallium and arsenic vacancies, respectively.

After the irradiation pulse is terminated, a heated area is cooled due to heat transport into the depth of a crystal. The characteristic recombination time of intrinsic defects of this process is found to be much longer than the time of their generation. This happens because the atom that left its crystal site (this occurs in a time comparable to the period of lattice vibrations) begins to move away from its "own" generated vacancy. This atom recombines after a period of time when it meets a "foreign" vacancy after diffusing through the crystal. As a result of such a difference between the generation and recombination characteristic times and due to the high cooling rates of the crystal, the concentration of the intrinsic defects n^* after irradiation becomes higher than the initial concentration n_0 but nevertheless remains lower than the concentration during the irradiation pulse (the effect of "freezing" of defects); i.e.,

$$n_0 < n^* < n_{eq}. \quad (4)$$

Intrinsic defects can ensure the flow of excess current through a thick space-charge layer if their concentration is no lower than $n_t = 10^{18}$ cm^{-3} (i.e., $n^* \gtrsim n_t$). Replacing n^* by smaller quantity n_t in inequalities (4) and taking into account (3), we obtain

$$n_t < n_i \exp(-\Delta H/kT_h). \quad (5)$$

This allows us to estimate the subsurface-layer temperature T_h at which the concentration of the intrinsic defects reaches the value necessary to provide multistage tunneling through a thick space-charge layer:

$$T_h > \frac{\Delta H}{k \ln(n_i/n_t)}.$$

This yields the required value $T_h > 1900$ K.

If we take into account that the integrity of the GaAs–Ni structure after irradiation is preserved and that the melting point of gallium arsenide $T_m = 1511$ K is much lower than T_h, the assumption about multistage tunneling via the states of intrinsic defects should be regarded as unrealistic.

Let us consider now the possible role of dislocations.

Since there are no other factors capable of producing irreversible changes in a crystal except those three mentioned in section 2, and the generated intrinsic defects are not able to ensure the observed excess current, it remains to be assumed that it is the dislocations that play a decisive role in the charge-carrier transport through a thick space-charge layer.

This assumption is supported by the similarity in behavior of the excess tunneling current at repeated irradiation doses with the behavior of the crystals under plastic deformation. On the one hand, this similarity consists in the fact that both the excess current and the plastic strain are increased only when the subsequent external action exceeds the preceding action. On the other hand, the excess current and the plastic strain remain unchanged when the subsequent external action is less than or equal to the preceding one. If, additionally, one takes into account that the plastic strain and its features after repeated external actions are caused only by the behavior of dislocations, then the assumption about the decisive role of dislocations in origination of the excess current becomes quite justified. Hence, the features of the behavior of a tunneling excess current also become qualitatively explicable.

REFERENCES

1. P. G. Eliseev, P. Perlin, J. Furioli, *et al.*, J. Electron. Mater. **26** (3), 311 (1997).
2. K. K. Dzhamanbalin and A. G. Dmitriev, Fiz. Tekh. Poluprovodn. (Leningrad) **24** (11), 2024 (1990) [Sov. Phys. Semicond. **24**, 1257 (1990)].
3. K. K. Dzhamanbalin, A. G. Dmitriev, V. V. Evstropov, and M. I. Shul'ga, Fiz. Tekh. Poluprovodn. (Leningrad) **25** (10), 1774 (1991) [Sov. Phys. Semicond. **25**, 1067 (1991)].
4. *Tunneling Phenomena in Solids*, Ed. by E. Burstein and S. Lundqvist (Plenum, New York, 1969; Mir, Moscow, 1973).
5. G. B. Abdullaev and T. O. Dzhafarov, *Diffusion of Atoms in Semiconductor Structures* (Atomizdat, Moscow, 1980).

Translated by A. Zalesskiĭ

Semiconductors, Vol. 34, No. 8, 2000, pp. 939–944. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 978–983.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Thermodynamic and Kinetic Aspects of Reconstruction Transitions at the GaAs(001) Surface

Yu. G. Galitsyn*, V. G. Mansurov, S. P. Moshchenko, and A. I. Toropov

Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, pr. akademika Lavrent'eva 13, Novosibirsk, 630090 Russia

** e-mail: sergem@isp.nsc.ru*

Submitted September 2, 1999; accepted for publication February 17, 2000

Abstract—A kinetic and thermodynamic analysis is carried out for reconstruction transitions on the GaAs(001) surface. It is shown that the transition from the As-stabilized $(2\times4)\beta2$ to the Ga-stabilized $(4\times2)\beta2$ structures under As_4 flux is a nonequilibrium phase transition and occurs if a certain steady concentration of arsenic adatoms is attained on the surface. The transition is continuous and can be approximated by a three-parameter isotherm. The moving force of an adsorbate-induced transition is the stabilization energy for the $(2\times4)\beta2$ phase accompanying the formation of arsenic dimers from arsenic adatoms. This energy is estimated. The features of the phase transitions occurring under the As_4 flux and under desorption conditions for an amorphous-arsenic film are discussed. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is known that a wide spectrum of reconstruction transitions from the Ga-stabilized (1×2) superstructure to the As-stabilized $c(4\times4)$ structure [1–6] is realized on the GaAs(001) surface:

$$(1\times2)\longrightarrow(4\times2)\beta2\longrightarrow(4\times2)\alpha\longrightarrow(3\times2)$$
$$\longrightarrow(3\times3)\longrightarrow(12\times6)\longrightarrow(2\times6)\longrightarrow(2\times4)\alpha$$
$$\longrightarrow(2\times4)\beta2\longrightarrow(2\times4)\gamma\longrightarrow(2\times3)\longrightarrow c(4\times4).$$

The investigation of these transitions is commonly complicated by the fact that, depending on the procedure of preparation and annealing of a surface, an arbitrary intermediate reconstruction can appear, which points to the substantial role of kinetic factors. In review [7], Yamaguchi and Horikoshi statistically considered in detail the structural transitions $(4\times2)\beta2\longrightarrow(2\times4)\beta2$ on GaAs(001) and InAs(001) surfaces in terms of the Ising two-dimensional model for lattice gas of adsorbed arsenic. In this model, a number of simplifying assumptions were used. Although the phase transitions were experimentally investigated under As_4-molecule flux, in the analysis, adsorption–desorption equilibrium was assumed between the surface and the monatomic arsenic gas. It was also assumed that the lateral interaction between arsenic dimers on the surface was attractive. The attraction energy was estimated at 0.11 eV for GaAs and 0.15 eV for InAs, which made it possible [7] to explain the difference in the phase transitions: a continuous phase transition for GaAs and an abrupt first-order transition for InAs. However, the interaction between single-type chemisorbed particles is repulsive as a rule [8]. Furthermore, only the surface atomic plane was taken into account in [7]; however, it is well known that two or three surface monatomic layers are involved in forming a particular reconstruction. The role of intermediate superstructures was also ignored.

In our study, we develop a more realistic model of structural transitions on the GaAs(001) surface in terms of the theory of adsorbate-induced phase transitions and with allowance made for various chemical processes occurring on the surface.

2. THERMODYNAMIC CONSIDERATION OF SURFACE PHASES

As usual, the stability of nonstoichiometric surfaces depends on a minimum of the Gibbs free potential Ω. This potential is a function of natural independent variables: the temperature T and the chemical potential μ. For constant T and μ, the increment $d\Omega$ is defined as $d\Omega = \sigma dS$, where σ is the sum of surface stress and surface tension (henceforth, we refer to this sum as the surface energy) and S is the surface area. It should be noted that σ is an essentially positive quantity and should be distinguished from the internal surface energy E_s. For the surface phase, we have

$$\Omega = \sigma S = E_s - TS_s + PdS - \mu_{As}N_{As} - \mu_{Ga}N_{Ga}, \quad (1)$$

where $\mu_{As,\,Ga}$ are the chemical potentials of the As and Ga surface atoms, $N_{As,\,Ga}$ is the number of As and Ga atoms on the surface, $dS = V$ is the volume of the surface phase, and d is the effective thickness of the surface layer. By virtue of the smallness of the gas-phase pressure over the surface, the third term may be neglected. We normalize all the values to the elemen-

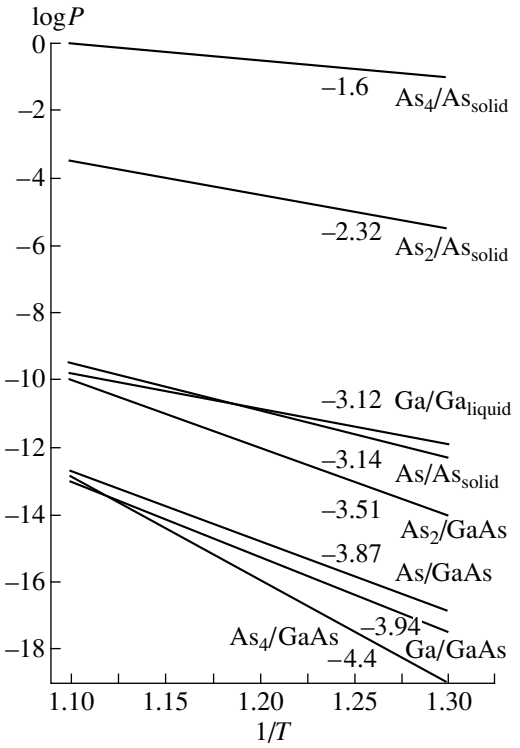


Fig. 1. Equilibrium partial pressures of various gases (in atm) over condensed media as functions of reciprocal temperature. The numbers on the straight lines correspond to the values of chemical potentials calculated from the formula $\mu_i = kT\ln[P/kT(h^2/2\pi mkT)^{3/2}]$.

tary area (1 × 1) of a perfect unreconstructed GaAs(001) surface (S = 16 Å^2). Then,

$$\sigma_{(1\times1)} = F_{s(1\times1)} - \mu_{As}n_{As} - \mu_{Ga}n_{Ga}. \quad (2)$$

Using elementary transformations and assuming that the surface is in equilibrium with the bulk ($\mu_{Ga} + \mu_{As} = \mu_{GaAs}$), the expression for σ can be written as

$$\sigma_{(1\times1)} = \text{const} - \frac{1}{2}\left[\left(n_{As} - \frac{1}{4}\right) - \left(n_{Ga} - \frac{3}{4}\right)\right] \times (\mu_{As} - \mu_{As(bulk)}), \quad (3)$$

where n_{Ga} = 3/4 and n_{As} = 1/4 correspond to the (1 × 2) phase with respected to which we counted surface energies for other reconstructions. We determine numerical values of the chemical potentials for arsenic and gallium. Using the thermodynamic data [9] for the equilibrium pressures of vapors over the condensed phases, we calculated the chemical potentials for Ga and As atoms in various media (Fig. 1). It should be noted that the values of chemical potentials $\mu_{As(bulk)} = -3.14$ eV and $\mu_{Ga(bulk)} = -3.10$ eV agree well with the data reported in [10]. It is the difference in the chemical potentials determined as $\mu_{As(bulk)} - \mu_{As(GaAs)} = 0.73$ eV and $\mu_{Ga(bulk)} - \mu_{Ga(GaAs)} = 0.84$ eV that defines the thermodynamically admissible domain of variation for the chemical potentials of As and Ga atoms on the surface for possible surface reconstructions and transitions between them.

For the reactions in which molecular gases (As_4 or As_2) are involved, it is necessary to take into account the contributions from the bonding energy of atoms in a molecule and the rotational component of the chemical potential of this molecule. Thus, for As_2 in equilibrium, we have $\mu_{As}^{GaAs} = \frac{1}{2}\mu_{As_2}^{trans} + \frac{1}{2}E_{As_2}^{bond} + \frac{1}{2}\mu_{As_2}^{rot}$, and a value $\mu_{As}^{gas} = \mu_{As}^{GaAs} = -3.87$ eV is consistent with values $\mu_{As_2}^{trans} = -3.51$ eV, $E_{As_2}^{bond} = -3.95$ eV, and $\mu_{As_2}^{rot} = -0.15$ eV. The vibrational contributions to the atomic chemical potentials are small, and their values are almost equal in a molecule and in a condensed phase.

The thermodynamic data presented testify that, if the reconstruction transitions occur under As_4 flux in the equilibrium conditions, in order to realize the complete set of reconstructions, we have to vary the pressure within a wide range (approximately by 12–14 orders of magnitude, see. Fig. 1) or, for a fixed pressure of ~10^{-6} Torr, to vary the temperature of a sample from 700 to 250°C. The translational component of the chemical potential of the As_4 molecule normalized per arsenic atom varies by $\frac{1}{4}\Delta\mu_{As_4}^{trans} = \frac{1}{4}(4.4 - 1.6) = 0.7$ eV, which agrees with the value of 0.73 eV cited above. For the transition from (4 × 2)β2 to (2 × 4)β2, the equilibrium chemical potential of an arsenic atom on the surface increases by ~0.35 eV (from –3.85 to –3.50 eV). However, the transition is experimentally realized at an As_4 pressure of 2.5×10^{-6} Torr in the temperature range from 650 to 540°C [7]; i.e., it is effected by a change from –2.77 to –2.42 eV in the translational component of the As_4 molecule chemical potential. Were the transition actually realized under the equilibrium conditions, the change in the chemical potential of an arsenic atom in the gaseous phase would amount to a mere 0.09 eV $\left(\frac{1}{4} \times 0.35 = 0.09 \text{ eV}\right)$, which is evidently insufficient for the phase transition. Thus, we may assume that the necessary concentration of adsorbed-arsenic atoms is provided by efficient nonequilibrium surface reactions of decomposition of the As_4 molecules to monatomic As. In fact, this reconstruction transition is a nonequilibrium phase transition.

3. KINETIC ASPECTS OF RECONSTRUCTION TRANSITIONS

As a rule, the reconstruction transitions are investigated in the As_4-gas flux with typical pressure values of 10^{-6} Torr, which is much lower than the equilibrium pressure over solid arsenic but higher than the equilib-

rium pressure over GaAs (see Fig. 1). Therefore, the decomposition reactions $As_4 \longrightarrow 2As_2$ and $As_2 \longrightarrow 2As$ proceed on the GaAs surface. These reactions are irreversible and considerably shift from equilibrium towards the formation of decomposition products by virtue of the absence of these products in the gaseous phase. In its turn, GaAs is also unstable with respect to the decomposition reaction into gaseous products Ga^{gas}, As_2^{gas}, and As^{gas}. However, these reactions can be neglected under routine experimental conditions of observation of the reconstruction transitions. We write the kinetic scheme of the surface reactions with participation of As_4 as

$$As_4^{gas} \underset{k_2}{\overset{k_1}{\rightleftarrows}} As_4^*, \quad 2As_4^* + 2Ga_s \xrightarrow{k_3} 2As_2^{chem} + As_4^{gas},$$

$$2As_2^{chem} + 2Ga_s \xrightarrow{k_4} 2Ga_s + 2As_s,$$

$$As_2^{chem} \xrightarrow{k_5} As_2^{gas} + Ga_s, \quad As_s \xrightarrow{k_6} As^{gas},$$

where k_1 and k_2 are the rate constants of the adsorption–desorption process of interaction between As_4 and the surface with formation of an As_4 molecule in the "precursor state" (As_4^*) and k_3 is the constant of the pairwise interaction between adsorbed molecules resulting in the dissociation of As_4^* by the Foxon and Joyce mechanism [11]. The constant k_4 describes the dissociation process in As_2^{chem} molecules leading to the formation of two-coordinated arsenic atoms on the surface according to the model [12]. The constants k_5 and k_6 specify the desorption of an As_2^{chem} molecule and an arsenic atom, respectively. Previously, we performed a kinetic analysis [13, 14] when investigating the epitaxial-growth processes for GaAs(001). For a steady-state concentration $\Theta_{As} = [As_s]/n_0$ (n_0 is specified below) of two-coordinated arsenic atoms on the surface at an As_4 pressure P and a temperature T, we have

$$\Theta_{As} = L(P, T)F(\xi),$$

where

$$L(P, T) = \frac{P}{P + P_0}, \quad F(\xi) = (\sqrt{1+\xi} - 1)^2 \frac{1}{\xi},$$

in its turn,

$$\xi = \frac{4k_1k_3P}{k_2^2}(1 + P/P_0) \quad \text{and} \quad P_0 = \frac{k_6(k_4 + k_5)}{k_1k_4}.$$

The function $F(\xi)$ defines the efficiency of conversion of an arsenic tetramer to monatomic arsenic. For $\xi \longrightarrow \infty$, $F(\xi) = 1$, and, for $\xi \ll 1$, $F(\xi) = \xi/4$. For $P > 10^{-6}$ Torr, ξ is large, and we may assume that $F(\xi) = 1$ [4]; then, $\Theta_{As} = L(P, T)$ is of the Langmuir form as if the adsorption–desorption equilibrium with the arsenic gas were established at a pressure P. Here, it should be taken into consideration that P_0 has a basically different physical meaning, which is specified by the kinetic constants k_i, than that in the true adsorption–desorption equilibrium for which $P_0 = [2\pi mkT/h^2]^{3/2}\exp(-E_{ads}/kT)kT$. In [7], the adsorption–desorption model of equilibrium between the surface and a monatomic gas was adopted without the proper kinetic analysis. In our kinetic model, the parameter P_0 specifies an excess steady concentration of adsorbed two-coordinated arsenic atoms necessary for the reconstruction transition. By virtue of the steadiness of the adsorbed-arsenic concentration, an effective chemical potential of arsenic atoms in the gaseous phase can be found from the expression $\mu_{As}^{gas} = kT\ln(P/P_0')$ or $\frac{\mu_{As}^{gas} + E_a}{kT} = \ln(P/P_0)$, where E_a is the effective energy of the arsenic-atom adsorption and P_0' is determined from the approximation $P_0 = P_0' \exp(-E_a/kT)$.

Table

Phase	n_{As}	n_{Ga}	D_{As}	D_{Ga}
(1×1)	1/4	3/4	0	0
$(4 \times 2)\beta 2$	1/2	3/4	0	3/8
(3×2)	2/3	2/3	0	3/9
$(4 \times 2)\alpha$	1/2	1/2	0	1/4
(3×3)	2/3	2/3	1/9	2/9
(2×6)	2/3	2/3	1/6	1/6
(12×6)	27/72	38/72	3/16	1/9
$(2 \times 4)\alpha$	1/2	1/2	1/4	0
$(2 \times 4)\beta 2$	3/4	1/2	3/8	0
(2×3)	1	1/4	1/6	0
$c(4 \times 4)$	5/4	0	3/8	0

4. STATISTICAL ASPECTS OF RECONSTRUCTION TRANSITIONS

We discuss the dependence of surface energy (3) for various reconstructions on the chemical potential μ_{As} and the concentration of surface arsenic and gallium atoms in detail. During the transition from the Ga-stabilized surface to the As-stabilized surface, the values of σ should be expected to be within a range bounded from above by the surface tension of liquid gallium and from below by the surface tension and stress of solid arsenic, i.e., from 0.72 to 0.17 eV (these values correspond to the GaAs (1×1)-cell area). Actually, the theoretically calculated changes in surface energies for various reconstructions in this transition are within this range [1, 4, 15–17]. The calculated dependences $\sigma(\mu_{As})$ obtained using the constants from [1, 4, 15] are shown in Fig. 2. In the table, we list the concentrations n_{As} and n_{Ga} for As and Ga two-coordinated atoms, respectively,

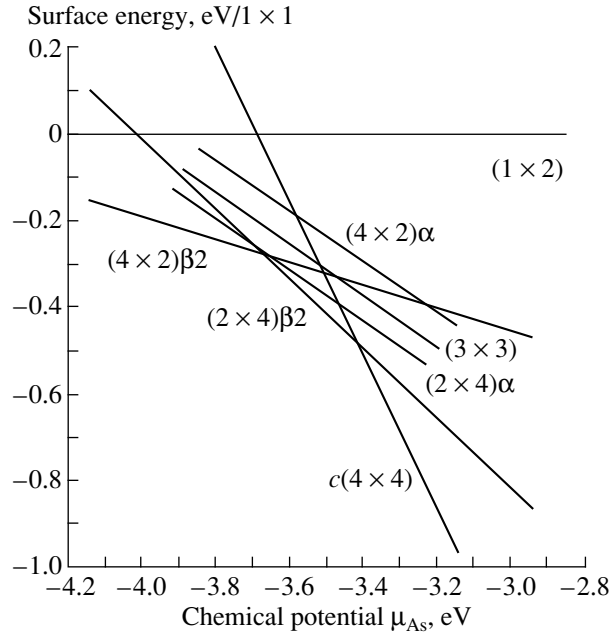


Fig. 2. Dependence of the surface energy $\sigma_{(1\times1)}$ for various reconstructions on the chemical potential for arsenic atoms on the surface.

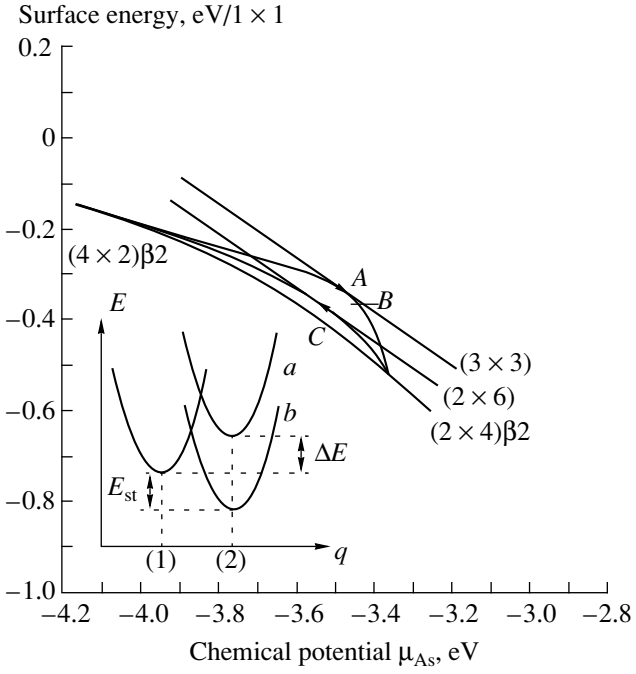


Fig. 3. The surface energy $\sigma_{(1\times1)}$ as a function of μ_{As} for the phase transition $(4\times2)\beta2 \longrightarrow (2\times4)\beta2$. (*A*) The phase transition under As_4 flux, (*B*) the phase transition during the desorption of an arsenic amorphous film, and (*C*) the equilibrium phase transition. In the insert, the variations of the potential energy during the phase transition are shown. Numbers (1) and (2) indicate the positions of gallium atoms in the (4×2) and (2×4) phases, respectively; curves *a* and *b* correspond to the states before and after the adsorption of arsenic.

and the concentrations D_{As} and D_{Ga} for arsenic and gallium dimers, respectively, which are also normalized to the (1×1) cell. The data were obtained from the experimental studies in which scanning tunneling microscopy was used to study various reconstructions [1–6]. In the case of the transition from the Ga-stabilized surface to the As-stabilized surface, the arsenic-dimer density increases, and the gallium-dimer density decreases. All the superstructures in the sequence from the $(4\times2)\beta2$ to $(2\times4)\beta2$ structures should be considered as transition metastable reconstructions, because even the most stable $(2\times4)\alpha$ structure with the largest number of As dimers is unstable according to the theoretical study [15]. It is possible to classify the relative stability of transition reconstructions according to the arsenic-dimer concentration.

We consider the phase transition between the stable reconstructions $(4\times2)\beta2$ and $(2\times4)\beta2$ in more detail. Figure 2 shows that surface energies of phases become equal for $\mu_{As} = -3.65$ eV. However, it is evident that, even in the domain of the $(4\times2)\beta2$-phase existence, arsenic adsorption on the surface takes place because any clean surface is unstable with respect to gas adsorption, and the surface energy decreases. In our case, taking into account that $d\sigma = -\Theta_{As}(\mu_{As})d\mu_{As} = -kT\Theta_{As}(P)dP/P$, a decrease in the surface energy as a function of an As_4 pressure and sample temperature is described by the expression

$$\sigma_{(1\times1)}(P) = \sigma_{(1\times1)}(n_{As}, n_{Ga}) - n_0 kT \ln(1 + P/P_0), \quad (4)$$

where $n_0 = 2D_{As}^{(2\times4)} - 2D_{As}^{(4\times2)} = 3/4$ and P_0 is the kinetic parameter specified above. When a critical concentration Θ_{As} of the two-coordinated arsenic atoms is attained, the reconstruction phase transition occurs with the formation of three-coordinated arsenic atoms in the dimer form ($2\Theta_{As} \longrightarrow D_{As}$). This transition process can be described kinetically; however, we also use the statistical approach, which clarifies the physical essence of the model proposed by us and the thermodynamic driving forces of the phase transition.

First of all, an energy gain must arise in the phase transition process. This gain, in our opinion, is proportional to the concentration of arsenic dimers formed in the new phase; i.e., we define the stabilization energy of the new phase as $E_{st} = \varepsilon_1 D_{As}$. On the other hand, in order to form a two-coordinate arsenic atom, the corresponding configuration of Ga atoms is necessary in the initial reconstruction [12], i.e., the reconstruction of a certain fraction of Ga atoms (Θ_{Ga}) in the upper plane. This process is accompanied by the decomposition of the Ga dimers. In this case, the energy expenditures are proportional to a variation $\Delta E = \varepsilon_2 D_{Ga}$ in the gallium-dimer concentration in the initial reconstruction. A change in the potential energy during the phase transition is shown in the insert in Fig. 3. The transition of gallium atoms ensuring the reconstruction of the initial surface occurs from energy position (*1*) to position (*2*). The two

parameters (E_{st} and ΔE) and also the energy (E_i) of repulsion between single-type adsorbed particles specify the three-parameter isotherm of the phase transition.

Applying a routine statistical analysis to the system composed of an adsorbed arsenic layer and a reconstructed surface of gallium atoms in the average-field approximation for the chemical potential of adsorbed arsenic atoms, we obtain

$$\mu_{As} = \frac{\partial F_{(1\times1)}}{\partial \Theta_{As}} = -E_a + E_i\Theta_{As} + kT\ln\left(\frac{\Theta_{As}}{1-\Theta_{As}}\right) - E_{st}\Theta_{Ga}. \quad (5)$$

An additional contribution to the surface free energy during the reconstruction of an initial surface layer of gallium atoms is equal to

$$\Delta F_{s(1\times1)} = \Delta E\Theta_{Ga} + kT[\Theta_{Ga}\ln\Theta_{Ga} + (1-\Theta_{Ga})\ln(1-\Theta_{Ga})] - E_{st}\Theta_{Ga}\Theta_{As}.$$

Minimizing this contribution with respect to the parameter Θ_{Ga} and substituting this value in (5), we obtain

$$\exp\left(\frac{\mu_{As}+E_a}{kT}\right) = \frac{\Theta_{As}}{1-\Theta_{As}} \times \exp\left[\frac{E_i\Theta_{As}}{kT} - \frac{E_{st}}{kT\left(1+\exp\left(\frac{\Delta E - E_{st}\Theta_{As}}{kT}\right)\right)}\right]. \quad (6)$$

The dependence of the phase transition on As_4 pressure and temperature follows from the equality between the effective chemical potential of arsenic atoms in the gaseous phase and their chemical potential in the adsorbed phase:

$$P/P_0 = \frac{\Theta_{As}}{1-\Theta_{As}} \times \exp\left[\frac{E_i\Theta_{As}}{kT} - \frac{E_{st}}{kT\left(1+\exp\left(\frac{\Delta E - E_{st}\Theta_{As}}{kT}\right)\right)}\right]. \quad (7)$$

As a result, we obtained a three-parameter isotherm which differs considerably from both the Langmuir isotherm and the Fowler–Hugenheim isotherm taken in [7]. Similar three-parameter isotherms were used for describing the reconstruction transition on surfaces of metals for the case of adsorption of various gases [18].

If the phase transitions take place in the temperature range of 400–800°C, the values of the parameters are typically within ΔE = 0.1–0.4 eV, E_{st} = 0.2–0.8 eV, and E_i = 0.05–0.1 eV. In Fig. 4, we show the isotherms for the variation of parameters ΔE within 0.1–0.6 (Fig. 4a), E_{st} within 0.1–0.6 (Fig. 4b), and a temperature T within

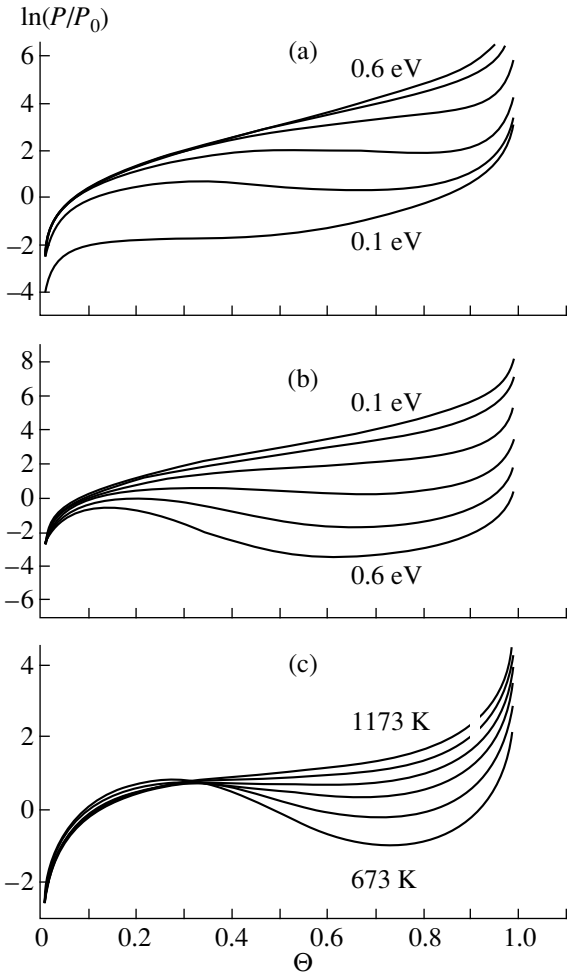


Fig. 4. Isotherms of the adsorption of arsenic for the phase transition $(4\times2)\beta2 \longrightarrow (2\times4)\beta2$ for various parameters: (a) T = 873 K, ΔE = 0.1–0.6 eV, E_i = 0.1 eV, and E_{st} = 0.4 eV; (b) T = 873 K, ΔE = 0.2 eV, E_i = 0.1 eV, and E_{st} = 0.1–0.6 eV; and (c) T = 673–1173 K, ΔE = 0.2 eV, E_i = 0.1 eV, and E_{st} = 0.4 eV.

400–900°C (Fig. 4c). It can be seen that such isotherms describe both the first-order phase transitions (appearance of the instability branch) and the continuous phase transitions. For the parameters F_i = 0.1, $\Delta E \le 0.2$ eV, and $E_{st} \le 0.4$ eV, we observe continuous phase transitions; in this case, the critical temperature of the transition is ~600°C. This fact is consistent with both our experimental data and the data reported elsewhere [7].

We estimate the stabilization energy for GaAs. Taking the energy of the arsenic dimer coupling (normalized to the (1×1) cell) $\varepsilon_1 \sim 1$ eV [19] and D_{As} = 3/8, we obtain E_{st} = 0.375 eV. On the other hand, ΔE must be lower than E_{st}, because the dimer bonding of gallium is weaker than that of arsenic. This means that, even for the direct reconstruction transition $(4\times2)\beta2 \longrightarrow (2\times4)\beta2$, the phase transition is continuous. Since the tran-

sition proceeds in an actual experiment through intermediate reconstructions for which the dimer number $D_{As} < 3/8$ (see table), every phase transition from one intermediate reconstruction to another is also continuous.

However, the authors of [7] considered the lateral attraction of arsenic dimers as the physical cause of the phase transition, whereas, in our case, the driving force of transition is the chemical conversion of the two-coordinated arsenic atom to the three-coordinated atom during the dimer-bonding formation. It should be noted that, in this process, no lateral attraction is assumed between atoms; on the contrary, we included the lateral repulsion that is natural for single-type adatoms, which does not hamper the realization of the phase transition.

The phase transition under As_4 flux proceeds via the intermediate reconstruction (3×1) that represents a version of the (3×2) and (3×3) phases [2, 7]. On the other hand, the intermediate (2×6), (12×6), and $(2 \times 4)\alpha$ phases are observed [1, 2] for the desorption of the amorphous arsenic layers. This is not surprising, because the kinetic parameter P_0 for these two cases is quite different. It is convenient to characterize the surface arsenic amorphous film, which is in equilibrium with this surface, by a lattice gas with a concentration C. Then, $\Theta_{As} = C/(C + C_0)$; however, C_0 (an analogue of P_0) is characterized by different kinetic constants. The phase transition under As_4 flux evidently proceeds under conditions of relative deficit in adsorbed arsenic on the surface, and, in the case of amorphous-film desorption, under an excess of arsenic. Therefore, in the former case, we can expect the realization of the Ga-stabilized intermediate phases of the $(4 \times 2)\alpha$, (3×2), and (3×3) types; in the latter case, the appearance of the As-stabilized phases of the $(2 \times 4)\alpha$, (2×6), and (12×6) types is more probable. In Fig. 3, we show the surface-energy change in the phase transition for both cases. Curve *A* corresponds to the phase transition under As_4 flux, and curve *B* refers to the transition during the desorption of a single-layer film of amorphous arsenic. Curve *A* is seen to intersect the domain of stability for the gallium intermediate phases, and curve *B* passes only across the domain of stability of arsenic phases. Curve *C* corresponds to the transition under conditions of adsorption–desorption equilibrium of the monatomic arsenic gas with the surface.

5. CONCLUSION

Thus, the thermodynamic and kinetic analysis of the phase transition $(4 \times 2)\beta 2 \longrightarrow (2 \times 4)\beta 2$ on the GaAs(001) surface carried out in this study shows that this transition is realized under conditions of a steady-state but nonequilibrium concentration of arsenic adatoms on the surface. We proposed a new mechanism of the reconstruction transition including a quasi-chemical reaction of conversion of the two-coordinated arsenic atom to the three-coordinated atom in the dimer form. The transition is continuous and is described by the three-parameter isotherm. The driving force of this transition is the energy of stabilization of the $(2 \times 4)\beta 2$ phase during the adsorption of arsenic. We estimated the value of this energy. The features of these phase transitions proceeding under the action of the As_4 flux on the surface and under conditions of desorption of the amorphous arsenic film were analyzed. The key parameter defining the relative stability of intermediate reconstructions is the arsenic-dimer density.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 99-02-16676) and by the Ministry of Science (grant no. 5.7.99) under the program “Surface Atomic Structures.”

REFERENCES

1. Q. Xue, T. Hashizume, A. Ichimiya, *et al.*, Sci. Rep. Res. Inst. Tohoku Univ. A **44**, 113 (1997).
2. I. Chizhov, G. Lee, R. F. Willis, *et al.*, Surf. Sci. **419**, 1 (1998).
3. A. R. Avery, D. M. Holmes, T. S. Jones, *et al.*, Phys. Rev. B **50**, 8098 (1994).
4. Q. Xue, T. Hashizume, and T. Sakuraj, Prog. Surf. Sci. **56**, 1 (1997).
5. T. Hashizume, Q. Xue, J. Zhou, *et al.*, Phys. Rev. Lett. **73**, 2208 (1994).
6. J. Behrend, M. Wassermeier, L. Daweritz, and K. H. Ploog, Surf. Sci. **342**, 63 (1995); **372**, 307 (1997).
7. H. Yamaguchi and Y. Horikoshi, Phys. Rev. B **51**, 9836 (1995).
8. V. P. Zhdanov and B. Kasemo, Surf. Sci. Rep. **20**, 113 (1994).
9. J. B. Arthur, J. Phys. Chem. Solids **28**, 2257 (1967).
10. G. X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1989); J. Vac. Sci. Technol. B **5**, 933 (1987).
11. Foxon and B. A. Joyce, Surf. Sci. **50**, 434 (1975).
12. Y. Fukunishi and H. Nakatsuji, Surf. Sci. **291**, 271 (1993).
13. Yu. G. Galitsyn, S. P. Moshchenko, and A. S. Suranov, Phys. Low-Dimens. Struct. **7/8**, 81 (1998).
14. Yu. G. Galitsyn, V. G. Mansurov, and I. I. Marahovka, Phys. Low-Dimens. Struct. **5/6**, 75 (1997); **7**, 55 (1997).
15. Northrup and S. Froyen, Phys. Rev. B **50**, 2015 (1994).
16. D. J. Chadi, J. Vac. Sci. Technol. A **5**, 834 (1987).
17. T. Ohno, Phys. Rev. Lett. **70**, 631 (1993).
18. V. P. Zhdanov, *Elementary Physicochemical Processes on Solid Surfaces* (Plenum, New York, 1991).
19. W. Monch, in *Semiconductor Surfaces and Interfaces* (Springer-Verlag, Berlin, 1993), p. 366.

Translated by V. Bukhanov

Semiconductors, Vol. 34, No. 8, 2000, pp. 945–954. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 984–992.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Electron–Ion Exchange at the Insulator–Semiconductor Interfaces and Its Influence on Ion Transport in the Insulating Layer

E. I. Gol'dman

Institute of Radio Engineering and Electronics (Fryazino Branch), Russian Academy of Sciences, pl. Vvedenskogo 1, Fryazino, Moscow oblast, 141120 Russia

Submitted October 11, 1999; accepted for publication February 24, 2000

Abstract—A theory of ion transport in insulating layers at the surface of semiconductors is developed. Neutralization of ions, disintegration of neutral ion–electron associations, the ion drift, and diffusion of neutral associations are taken into account. The theory makes it possible to eliminate the inconsistency in the interpretation of experimental data on ionic depolarization of SiO_2 layers that exhibit mutually exclusive properties; these layers can be conceived as both a system of ionic traps with a wide distribution of lifetimes and a medium with free ions. Neutral associations act as traps; the wide distribution of lifetimes is caused by a spread in the tunneling distances for electrons in the course of disintegration of the associations. An increase in the degree of ionization as a result of a quasi-steady decrease in the electric field ensures a gradual transition from a collection of low-mobility neutral associations to an ensemble of free ions. © *2000 MAIK "Nauka/Interperiodica"*.

Investigations of generation and migration of ions in insulating layers at the surface of semiconductors have been actively pursued for several decades [1–4]; however, current concepts of the mechanism of the ion transport (including that in well-developed technologically structures like SiO_2/Si) are rather contradictory. Depending on the type of experiment, the insulator manifests itself in a mutually exclusive way: either as a system of ionic traps with a wide distribution of lifetimes (isothermal or thermally stimulated polarization and depolarization of insulator [5–9]) or as a medium with free ions (the transit-time effect [10, 11] and dynamic current–voltage (*I–V*) characteristics [12–16]). In this study, we used the features of electron–ion exchange at the insulator–semiconductor interface [17] to develop a phenomenological theory of ion transport in the insulating layer; this theory resolves the above contradiction.

The basic concepts of the model [17] consist in the following. Mobile ions are assumed to be present in the insulating layer of a metal–insulator–semiconductor (MIS) structure based on an *n*-type semiconductor. For the sake of definiteness, these ions are assumed to be singly positively charged; they can migrate only within the insulating layer. Each ion in combination with the surrounding insulator matrix generates a localized electronic state. Once this state is occupied, a neutral association of the ion and electron is formed. The ion drift, diffusion of neutral associations, and also the processes of their formation (due to neutralization of ions as a result of occupation of localized states with electrons from the semiconductor) and disintegration (due to ionization of neutral associations as a result of transition of electrons from the localized states to the semiconductor or the conduction band of the insulator) define the mechanism of ion transport in the insulator. In the case of fairly high steady-state polarizing fields ($V_i > 0$, where $V_i = V_g - V_s - V_c$, V_i is the voltage drop across the insulator, V_s is the voltage drop across the semiconductor, and V_c is the contact potential difference at the gate–semiconductor interface), the ions concentrate at the insulator–semiconductor interface. Simultaneously, partial neutralization of these ions occurs as a result of tunneling transitions of electrons from the enhancement surface layer of the semiconductor to the localized states and back. The formed neutral associations diffuse to the bulk of the insulator where they decompose into ions and electrons. The latter almost immediately (compared to characteristic times of the ion transport) drift to and reach the corresponding electrode. The newly formed ions are subject to an electric field and, depending on its direction, would either return to the insulator–semiconductor interface (the polarizing fields, see Fig. 1a) or drift to the contact with the field electrode (the depolarizing fields, see Fig. 1b). The mass transport in the insulator is described by the equations

$$\frac{\partial M}{\partial t} - D\frac{\partial^2 M}{\partial z^2} + \frac{M}{\tau} + \frac{M}{\tau_i} - \frac{N}{\tau_n} = 0, \quad (1)$$

$$\frac{\partial N}{\partial t} + \frac{1}{q}\frac{\partial j}{\partial z} + \frac{N}{\tau_n} - \frac{M}{\tau_i} - \frac{M}{\tau} = 0. \quad (2)$$

Here, N and M are the volume concentrations of ions and neutral associations in the insulator; $z = 0$ is the

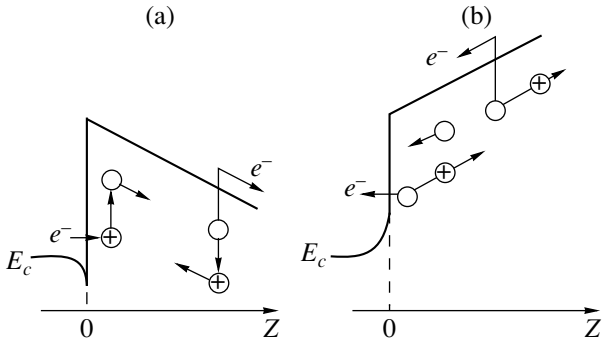


Fig. 1. Schematic representation of the ion transport at the insulator–semiconductor interface for (a) polarizing and (b) depolarizing voltages. The designations used include ○ for a neutral association composed of an ion and an electron, ⊕ for a positive ion, e^- for an electron, Z for a coordinate (the region of $Z < 0$ is occupied by a semiconductor and that of $Z > 0$, by an insulator), and E_c for the conduction-band bottom.

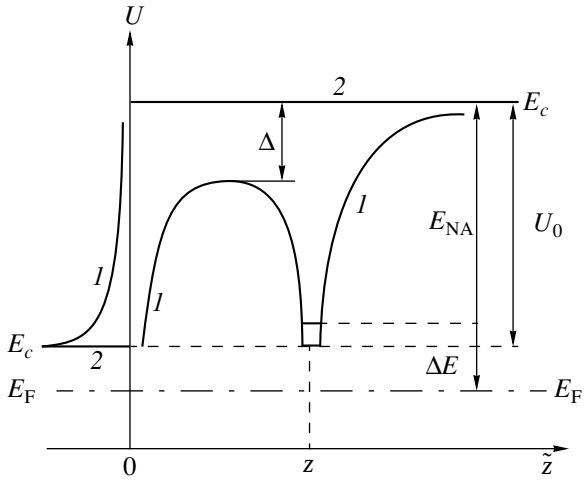


Fig. 2. Potential relief for an electron bound to an ion to form a neutral association in the vicinity of the insulator–semiconductor interface. Here, (*1*) is the potential with the image forces taken into account, (*2*) is rectangular barrier that has the height U_0 and arises from the discontinuity in the conduction bands at the semiconductor–insulator contact, $\tilde{z}$ is the coordinate of the electron ($\tilde{z} < 0$ is for the semiconductor and $\tilde{z} > 0$ is for the insulator), z is the coordinate of the ion, E_c is the conduction-band bottom, E_F is the Fermi level, Δ is a decrease in the barrier height due to the image forces, E_{NA} is the electron energy in a neutral association (this energy is measured from the conduction-band bottom in the insulator), and ΔE is the energy distance from the electron level in neutral association to the Fermi level.

coordinate of the interface (the semiconductor is in the region of $z < 0$ and the insulator is in the region of $z > 0$, see Fig. 1); D is the diffusion coefficient of neutral associations; τ is the characteristic time of the neutral-association disintegration related to the electron transition to the conduction band of the insulator; $\tau_i(z)$ is the characteristic time of ionization of an association located at a distance z from the interface (this ionization is caused by the electron leaving for the semiconductor); $\tau_n(z)$ is the characteristic time for neutralization of ion located at a distance z from the interface (this process is due to the arrival of an electron from the semiconductor); q is the elementary charge; $j = q\mu NF - \mu kT\partial N/\partial z$ is the ion current density; and F is the electric field in the insulator.

The probabilities for ionization $\tau_i^{-1}(z)$ and neutralization $\tau_n^{-1}(z)$ per unit time decrease drastically with increasing z as a result of an increase in both the tunneling distance and the tunneling-barrier height. The major contribution to the potential $U(\tilde{z})$ (see Fig. 2; here, $\tilde{z}$ is the electron's coordinate), which limits the rate of transitions of electrons from the semiconductor (and back) to the localized states generated by ions, is given by the conduction-band discontinuity U_0 at the semiconductor–insulator interface and the image forces induced by the ion and electron charges. For $\tilde{z} > 0$ and ignoring the external electric field, we have

$$U(\tilde{z}) = U_0 - (q^2/\kappa_{i\mathrm{LF}})\{1/|z-\tilde{z}| - [1/(z+\tilde{z})] \times (\kappa_{s\mathrm{LF}} - \kappa_{i\mathrm{LF}})/(\kappa_{s\mathrm{LF}} + \kappa_{i\mathrm{LF}})\} - (q^2/\tilde{z}\kappa_{i\mathrm{HF}})(\kappa_{s\mathrm{HF}} - \kappa_{i\mathrm{HF}})/(\kappa_{s\mathrm{HF}} + \kappa_{i\mathrm{HF}}),$$

where z is the ion's coordinate, κ_i and κ_s are the permittivities of the insulator and semiconductor, and the subscripts LF and HF stand for low and high frequencies [18]. The image forces reduce appreciably the barrier height for fairly small z. Thus, for the Si–SiO_2 interface ($\kappa_{s\mathrm{LF}} = \kappa_{s\mathrm{HF}} = 11.5$, $\kappa_{i\mathrm{LF}} = 3.9$, and $\kappa_{i\mathrm{HF}} = 2.13$ [18]), the actual barrier (Fig. 2, curve *1*) is lower than the rectangular barrier (Fig. 2, curve *2*) by $\Delta(z) = 1.07q^2/z = 1.54$ (10 Å/z) eV. For large z, the function $\tau_i(z)$ can be closely approximated with the expression $\tau_i = \bar{\tau}_i \exp(z/\delta_\infty)$, where $\bar{\tau}_i$ is a multiplier that is independent of the coordinate; $\delta_\infty = (\hbar^2/8mE_{\mathrm{NA}})^{1/2}$ is the characteristic length for attenuation of the electron wave function in the insulator as the distance from the interface increases; m is the effective mass of the electron in the insulator; and E_{NA} is the electron energy in a neutral association (this energy is measured from the conduction-band bottom in the insulator). According to the data reported for SiO_2 [19], $\delta_\infty \approx (U_0/E_{\mathrm{NA}})^{1/2}$. The ratio of the ionization and neutralization rates is defined by the distance ΔE from the electron level in the neutral association to the Fermi level; i.e., we have $\tau_n/\tau_i = \exp(\Delta E/kT)$. It follows from this that the degree of ion neutralization can be controlled by varying the Fermi level position at the interface. It is self-evident that, if the electron level in the neutral association is located considerably higher than the conduction-band bottom in the semiconductor,

the neutralization is not feasible. The electric field in the insulator is defined by the Poisson equation

$$\frac{\partial F}{\partial z} = \frac{4\pi qN}{\kappa_i}, \tag{3}$$

with $F|_{z\to\infty} = F_h$ (to be more precise, $F|_{z=h} = F_h$), where F_h is the field at the contact of insulator with the gate.

We now consider the steady-state distribution of ions and neutral associations at the interface in the polarizing field $F < 0$ (Fig. 1a). The boundary conditions to the system of equations (1) and (2) entail that $j = 0$ and $\partial M/\partial z = 0$ for $z = 0$ (the conditions for the absence of the ion and neutral-association transport through the interface) and $M = 0$ and $N = 0$ for $z \longrightarrow \infty$. It follows from (1) and (2) that

$$\frac{d^2M}{dz^2} - \frac{M}{\lambda^2} - \frac{M}{D\tau_i} + \frac{N}{D\tau_n} = 0, \tag{4}$$

$$j - qD\frac{dM}{dz} \equiv q\mu NF - \mu kT\frac{dN}{dz} - qD\frac{dM}{dz} = 0. \tag{5}$$

Here, $\lambda = (D\tau)^{1/2}$ is the diffusion length of neutral associations. Relationship (5) (the absence of the overall flux of ions and neutral associations) describes mathematically the self-sustained circulation of particles at the insulator–semiconductor interface. Neutral associations formed after the ions were neutralized diffuse in the bulk of the insulator where they decompose thermally into ions and electrons. The newly formed ions drift to the interface under the effect of the electric field, are again neutralized at the interface, and the process continues. Equation (4) supplemented with the conditions $(dM/dz)|_{z=0} = 0$ and $M|_{z=\infty} = 0$ can be represented in the integral form as

$$M = \frac{M_0}{2}\exp\left(-\frac{z}{\lambda}\right)$$

$$+ \frac{\tau}{2\lambda}\int_0^\infty \exp\left(-\frac{|z_1 - z|}{\lambda}\right)\left(\frac{N}{\tau_n} - \frac{M}{\tau_i}\right)dz_1, \tag{6}$$

$$M_0 = M|_{z=0}.$$

We use equations (3) and (5) to obtain

$$N_0 + (qD/\mu kT)M_0$$

$$= (2\pi q^2 N_s^2/\kappa_i kT) + (q|F_h|N_s/kT),$$

where $N_0 = N|_{z=0}$ and $N_s = \int_0^\infty N\,dz$ (in the case under consideration, $F_h < 0$). We assume that $(\lambda/\tau_n)d\tau_n/dz \gg 1$ and $(kT/q|F|\tau_n)d\tau_n/dz \gg 1$; i.e., the functions $\tau_i(z)$ and $\tau_n(z)$ vary with z much more rapidly than do $M(z)$ and $N(z)$. It then follows from (6) that

$$M \cong M_0\exp\left(-\frac{z}{\lambda}\right),$$

$$N_0 = M_0\left[\frac{\tau_{n0}\lambda}{\tau\delta_0} + \exp\left(\frac{\Delta E_0}{kT}\right)\right], \tag{7}$$

where $\delta_0 = \tau_{n0}\int_0^\infty \tau_n^{-1}\,dz$ is the characteristic distance defining the dependence of τ_n on z, $\tau_{n0} = \tau_n|_{z=0}$, and $\Delta E_0 = \Delta E|_{z=0}$ [here, the equality $\tau_n/\tau_i = \exp(\Delta E/kT)$ is used]. Formulas (7) differ from exact relationship (6) in small terms on the order of $(\delta_0/\lambda) \ll 1$; however, it is these terms that ensure the vanishing of the flux of neutral associations at $z = 0$. The degree of ionization of particles at the interface γ [$\gamma \equiv N_s/(N_s + M_s)$, where $M_s = \int_0^\infty M dz$, $N_s + M_s = N_{s0}$, and N_{s0} is the particle concentration at the interface], is given by

$$\gamma = \frac{2\gamma_H\gamma_L}{\gamma_L + \sqrt{\gamma_L^2 + 4\gamma_H^2(1-\gamma_L)}}, \tag{8}$$

$$\gamma_H^{-1} = 1 + \frac{q|F_h|\lambda}{kT\left[\frac{\tau_{n0}\lambda}{\tau\delta_0} + \frac{qD}{\mu kT} + \exp\left(\frac{\Delta E_0}{kT}\right)\right]},$$

$$\gamma_L^{-1} = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2\pi q^2\lambda N_{s0}}{\kappa_i kT\left[\frac{\tau_{n0}\lambda}{\tau\delta_0} + \frac{qD}{\mu kT} + \exp\left(\frac{\Delta E_0}{kT}\right)\right]}}. \tag{9}$$

The physical significance of the introduced parameters is evident from formulas (8) and (9): γ_H and γ_L are the degrees of ionization in the high- and low-intensity fields, respectively.[1] The quantity ΔE_0 decreases as the magnitude of the electric field at the gate and the ion concentration increase. We have $\exp(\Delta E_0/kT) = \exp(\Delta E_{FB}/kT)(n_d/n_0)$, where $n_0 \approx n_d + (2\pi q^2/\kappa_s kT)[N_s + (\kappa_i|F_h|/4\pi q)]^2$; ΔE_{FB} is the energy distance between the level electron attached to the neutral association and the Fermi level in the semiconductor under the conditions of flat bands; and n_0 and n_d are the electron concentrations at the interface and in the bulk of the semiconductor, respectively. The steady-state current through the insulator is controlled by the rate of volume recombination of electrons when the neutral association is decomposed; i.e.,

$$I = -\frac{qSM_s}{\tau}. \tag{10}$$

Here, S is the structure area and the negative sign indicates the current flows from the gate to the semiconduc-

[1] The expression for γ_H was derived in [17] in the limit of large V_i, in which case it may be assumed that the field F is uniform.

tor. As the electric-field strength increases, the magnitude of current increases sublinearly and ultimately attains the constant value of $|I| = qSN_{s0}/\tau$ in the case where the ions are completely neutralized.

We now consider the transient process related to the transport of ions and neutral associations after the polarizing field ($F < 0$) was reversed to obtain the depolarizing field ($F > 0$). Let the semiconductor surface be depleted of carriers to such an extent after the field-direction reversal that the neutralization of ions may be ignored (ΔE_0 is fairly large); we also assume that the transit time $\tau_d \sim h/\mu F$, where h is the thickness of the insulating layer, is much shorter than the lifetime of neutral associations τ in the bulk of the insulator. The relaxation of polarization is controlled by four effects: (i) ionization of neutral associations as a result of tunneling transition of electrons to the semiconductor; (ii) transit of the formed ions; and (iii) diffusion of neutral associations to the interface with subsequent (iv) disintegration of them at the boundary or in the bulk of the insulator. Correspondingly, we can recognize three characteristic time intervals. In the first of these, the depolarization current is caused by the tunneling electron transitions and the transit of ions formed as a result of the neutral-association disintegration at the interface in a thin layer, within which $\tau_i(z) \leq \tau_d$. In the second time interval, ionization of neutral associations occurs in a layer with thickness z_* defined by

$$\tau_d \ll \tau_i(z_*) = \min(z_*^2/D, \tau);$$

i.e., the transit of ions does not limit the depolarization, and the diffusion of neutral associations is insignificant. In the third (and final) time interval, the relaxation is controlled by disintegration of neutral associations in the bulk of the insulator and by diffusion of these associations to the near-surface region $z < z_*$, where their ionization rate is higher. The process occurring during the first time interval (the transit-time effect) is described by the mass-transport equations

$$\frac{\partial M}{\partial t} + \frac{M}{\tau_i} = 0, \tag{11}$$

$$\frac{\partial N}{\partial t} + \mu\frac{\partial (NF)}{\partial z} - \frac{M}{\tau_i} = 0. \tag{12}$$

Neglect of diffusion in (12) is justified if $V_i \gg kT/q$ and $F_0 > 0$ ($F_0 = F|_{z=0}$). The initial condition

$$M|_{t=0} = M_i(z) = (M_{si}/\lambda_i)\exp(-z/\lambda_i)$$

corresponds to the stationary distribution of particles (7); here, M_{si} is the initial concentration of neutral associations. The boundary condition $N|_{z=0} = 0$ ensures the absence of mass transport through the interface. Equations (3) and (12) can be used to derive the following relationship for the current through the insulator:

$$I = qS\left(\frac{\kappa_i}{4\pi q}\frac{\partial F}{\partial t} + \mu F N + \int_z^h \frac{M(z_1)}{\tau_i(z_1)}dz_1\right). \tag{13}$$

The transit time τ_d can be determined from the equality $\int_0^{\tau_d}\mu F_h dt = h$. For $\tau < \tau_d$, we use (13) to obtain

$$I = \frac{\kappa_i S}{4\pi}\frac{dF_h}{dt}, \quad I = qS\left(\frac{\kappa_i}{4\pi q}\frac{\partial F_0}{\partial t} + \int_0^h \frac{M}{\tau_i}dz\right).$$

Hence, in view of these formulas and (11), we have

$$F_h - F_0 = \bar{F} \equiv (4\pi q/\kappa_i)\int_0^h (M_i - M)dz + (4\pi q N_{si}/\kappa_i),$$

where N_{si} is the initial ion concentration. We derive another relationship for current by integrating equality (13) with respect to the coordinate; i.e.,

$$I = \frac{\kappa_i S}{8\pi h}\left[-2\frac{dV_i}{dt} + \mu(F_h^2 - F_0^2) + \frac{8\pi q}{\kappa_i}\int_0^h z\frac{M}{\tau_i}dz\right].$$

Taking into account that $V_s = 2\pi q n_d w^2/\kappa_s$, $4\pi q n_d w = \kappa_i F_0$, where w is the thickness of the depletion layer at the semiconductor surface, we obtain the following formula for the current:

$$I = \frac{\kappa_i\kappa_s\mu S\bar{F}(2F_0 + \bar{F})}{8\pi(\kappa_s h + \kappa_i w)} + \frac{qS}{(\kappa_s h + \kappa_i w)} \times \int_0^h (\kappa_i w + \kappa_s z)\frac{M}{\tau_i}dz - \frac{\kappa_i\kappa_s S}{4\pi(\kappa_s h + \kappa_i w)}\frac{dV_c}{dt}. \tag{14}$$

The three terms in expression (14) correspond to three different mechanisms of transport. The first term describes the transit of free ions that either existed at the interface for $t = 0$ or were formed as a result of disintegration of neutral associations in the time t. The second term accounts for the electron tunneling transitions that cause the neutral associations to disintegrate. The third term describes the capacitive current related to a change in the contact potential difference (this term is nonvanishing only for a thermally stimulated process). In the case of isothermal depolarization, the first component of current increases with time, whereas the second component decreases. It is noteworthy that, according to the theory of emptying the surface traps at the insulator–metal contact [20], the current increases with time ($I|_{t=0} = 0$), because the dependence $I(t)$ does not involve a term similar to the second term in the right-hand side of (14). This distinction is caused by the fact that spatial separation of electronic and ionic charges in the course of emptying the traps was disre-

garded in [20] (in the case under consideration here, this separation occurs at a distance of $w + z$).

A solution to equation (11) has the form

$$M = M_i(z)\exp[-t_{\text{eff}}/\tau_i(z, t)],$$

$$t_{\text{eff}} = \tau_i(z, t)\int_0^t dt_1/\tau_i(z, t_1).$$

The ionization time τ_i in relation to coordinate and temperature can be expressed as $\tau_i = \bar{\tau}\exp(\varepsilon)$, where $\bar{\tau}$ = const and $\varepsilon = \varepsilon(z, T) \gg 1$; ε increases with increasing z and decreases with increasing T. In addition,

$$\delta \equiv (\partial\varepsilon/\partial z)^{-1}, \quad \delta|_{z=0} = \delta_0, \quad \delta|_{z\to\infty} = \delta_\infty.$$

If T = const, $t_{\text{eff}} = t$; if $T = T_0 + \beta_T t$ (the thermally stimulated process), $t_{\text{eff}} \simeq t$ for $|(T - T_0)\partial\varepsilon/\partial T| \ll 1$ and $t_{\text{eff}} \simeq (\beta_T|\partial\varepsilon/\partial T|)^{-1}$ for $|(T - T_0)\partial\varepsilon/\partial T| \gg 1$. In the latter case, t_{eff} is the time it takes for τ_i to decrease by a factor of e. A unified description of relaxation signals in the cases of isothermal and thermally stimulated processes can be accomplished by introducing a generalized time t_{eff} and is characteristic of monomolecular processes [21]. A wide distribution of the tunneling-ionization times τ_i makes it possible to draw a parallel between neutral associations and localized electronic states with a large variance in the lifetimes [21, 22] and represent the function $M(z)$ in the form of a step function

$$M(z) = \begin{cases} 0, & z < z_m \\ M_i(z), & z > z_m, \end{cases} \qquad \varepsilon(z_m, T) = \ln\left(\frac{t_{\text{eff}}}{\bar{\tau}}\right), \quad (15)$$

where z_m is the coordinate separating the regions of already-disintegrated ($z < z_m$) and still-existing ($z > z_m$) neutral associations. Within this approach, the quantity δ should be treated as the characteristic scale of erosion of the step. Therefore, approximation (15) is valid for $\delta \ll (z_m, \lambda_i)$. For a limited time range, the explicit dependence of the coordinate z_m on t can be derived by expanding the function $\varepsilon(z_m, T)$ into a series; as a result, we have

$$z_m = \bar{z}_m + \delta\ln\left(\frac{t_{\text{eff}}}{\bar{t}}\right), \quad \varepsilon(\bar{z}_m) = \ln\left(\frac{\bar{t}}{\bar{\tau}}\right). \qquad (16)$$

Here, $\bar{z}_m$ and δ correspond to the initial point of the chosen time range $\bar{t} < t_{\text{eff}} < \bar{\bar{t}}$; i.e., they are calculated for $t_{\text{eff}} = \bar{t}$. Since $\varepsilon \gg 1$, the maximal $t_{\text{eff}} = \bar{\bar{t}}$ belonging to this range may exceed $\bar{t}$ by several orders of magnitude. Thus, we obtain

$$\bar{\bar{F}} = \frac{4\pi q M_{si}}{\kappa_i}\left[1 - \exp\left(-\frac{\bar{z}_m}{\lambda_i}\right)\left(\frac{\bar{t}}{t_{\text{eff}}}\right)^{\frac{\delta}{\lambda_i}}\right] + \frac{4\pi q N_{si}}{\kappa_i}.$$

For $F \gg \bar{\bar{F}}$ (i.e., on the assumption that $F \approx F_0 \approx F_h$) and $z_m \ll \kappa_i w/\kappa_s$, expression (14) transforms into

$$I = \frac{qS\kappa_s\mu F N_{si}}{(\kappa_s h + \kappa_i w)} - \frac{\kappa_i\kappa_s}{4\pi(\kappa_s h + \kappa_i w)}\frac{dV_c}{dt}$$

$$+ \frac{qSM_{si}}{(\kappa_s h + \kappa_i w)}\left\{\mu\kappa_s F\left[1 - \exp\left(-\frac{\bar{z}_m}{\lambda_i}\right)\left(\frac{\bar{t}}{t_{\text{eff}}}\right)^{\frac{\delta}{\lambda_i}}\right]\right. \qquad (17)$$

$$\left. + \frac{\delta\kappa_i w}{\lambda_i t_{\text{eff}}}\exp\left(-\frac{\bar{z}_m}{\lambda_i}\right)\left(\frac{\bar{t}}{t_{\text{eff}}}\right)^{\frac{\delta}{\lambda_i}}\right\}.$$

It is evident from (17) that, in the case of T = const, the time dependence of current has a minimum at $t = t_{\min} \simeq \kappa_i^2/4\pi q n_d \mu\kappa_s$. The component of I that increases with t is due to an increase (as a result of disintegration of neutral associations) in the number of ions involved in the transit. The current component decreasing with time is related to tunneling transitions of electrons when neutral associations are ionized. In the case of thermal stimulation, the temperature dependence of mobility plays a decisive role. The transit current of ions increases exponentially with temperature against the background of the components that depend on T only slightly and correspond to the changes in the contact potential difference and to the tunneling ionization of neutral associations.

In the low-intensity fields $F_h < 4\pi q N_s/\kappa_i$, a reservoir of free ions forms in the insulator within the interval $(0, z_0)$, whereas an enhancement layer is formed in the semiconductor (since $F_0 < 0$). The boundary of this reservoir $z = z_0$ is defined by the condition $F(z_0) = 0$; physically, it has the same significance as a virtual cathode in the theory of currents limited by the space charge [23]. We use equations (3) and (13) to obtain

$$I = \frac{\kappa_i S}{4\pi}\frac{dF_h}{dt},$$

$$I = -\frac{\kappa_i S}{4\pi h}\frac{dV_c}{dt} + \frac{\mu\kappa_i S}{8\pi h}F_h^2 + \frac{qS}{h}\int_{z_0}^{h}(z - z_0)\frac{M}{\tau_i}dz. \qquad (18)$$

When deriving (18), we assumed that $z_0 \ll h$ and ignored the changes in the potential drops across the region of the ion-reservoir localization and across the enhancement layer in the semiconductor. If the field is weak, i.e.,

$$F_h \ll \left[(8\pi q/\mu\kappa_i)\int_{z_0}^{h}(z - z_0)(M/\tau_i)dz\right]^{1/2},$$

the current is controlled by electron-tunneling transitions in the course of the neutral-association disintegration. Using approximations (15) and (16) and the condition $\bar{z}_m \gg z_0$, we obtain

$$I = -\frac{\kappa_i S}{4\pi h}\frac{dV_c}{dt} + \frac{qSM_{si}\bar{z}_m}{h\lambda_i \bar{t}}\exp\left(-\frac{\bar{z}_m}{\lambda_i}\right)\left(\frac{\bar{t}}{t_{\text{eff}}}\right)^{1+\frac{\delta}{\lambda_i}}. \quad (19)$$

For $t > t_i$, where t_i is the point in time beginning when the inequalities

$$F_h \gg \left[(8\pi q/\mu\kappa_i)\int_{z_0}^{h}(z-z_0)(M/\tau_i)dz\right]^{1/2},$$

$$\mu F_h^2 \gg |dV_c/dt|$$

hold, the current is caused by the transit of the free-ion front; i.e., we have

$$I = \frac{\mu}{2h}\frac{F_i^2}{\left(1-\frac{F_i}{2h}\int_{t_i}^{t}\mu(t_1)dt_1\right)^2}, \quad F_i \equiv F_h(t_i). \quad (20)$$

The general pattern of the time dependence of depolarization current for $t < \tau_d$ and for a high initial degree of neutralization is as follows. Under isothermal conditions, the current I is initially related to electron tunneling transitions in the case of ionization of neutral associations and decreases with time. As the number of free ions increases, the current component caused by the transit of these ions becomes dominant and the current I increases with t. As V_g increases, the transit time decreases and the ascending current–time curve ceases to exist. In the case of thermal stimulation, the magnitude of the current component that increases with time is much larger than that under isothermal conditions; this is due to an exponential increase in the ion mobility. A fall in the dependence $I(t)$ may not be observable against the background of the above component.

In the time range of $\tau_i(z_*) \gg t \gg \tau_d$, the delay related to the transit of free ions through the insulating layer becomes insignificant, whereas the diffusion and disintegration of neutral associations in the bulk of the insulator (with the characteristic time τ) have no time to evolve. The depolarization current is controlled by electron tunneling transitions on ionization of the neutral associations located in the region of $z < z_*$. In the approximations given by (15) and (16), we have[2]

$$I = qS\int_0^h \frac{M}{\tau_i}dz \simeq \frac{qSM_{si}\delta}{\bar{\tau}\lambda_i}\exp\left(-\frac{\bar{z}_m}{\lambda_i}\right)\left(\frac{\bar{\tau}}{t_{\text{eff}}}\right)^{1+\frac{\delta}{\lambda_i}}. \quad (21)$$

[2] For T = const, this result was obtained in [17].

Expressions (17), (19), and (21) correspond to the different time ranges. Therefore, $\bar{z}_m$ and $\bar{t}$ may be considered as steadily increasing functions of time; however, these quantities increase more slowly than do $\delta\ln(t_{\text{eff}}/\tau)$ and t_{eff}, respectively. The quantity $\delta = \delta(\bar{z}_m, T)$ decreases steadily with t for T = const and exhibits a maximum in the case of thermal stimulation; $\delta|_{t\to 0} = \delta_0$, $\delta|_{t\to\infty} = \delta_\infty$.

The final stage of depolarization for $t \gg \tau_i(z_*) \equiv \min(z_*^2/D, \tau)$ is controlled by disintegration of neutral associations in the bulk of the insulator and by diffusion of these associations to the near-surface region $z < z_*$ where their ionization rate is higher. In the region of $z > z_*$, this stage is described by the mass transport equation

$$\frac{\partial M}{\partial t} - D\frac{\partial^2 M}{\partial z^2} + \frac{M}{\tau} = 0,$$

with the boundary and initial conditions given by

$$M(z_*, t) = 0, \quad M|_{z\to\infty} = 0, \quad M[z, \tau_i(z_*)] = M_*(z).$$

For $z - z_* \gg \delta(z_*)$, the initial neutral-association distribution $M_*(z)$ coincides with the stationary distribution $M_i(z)$. In the vicinity of $z = z_*$, the function $M_*(z)$ descends to zero within a certain length $\bar{\delta} \simeq \delta(z_*)$. If we ignore this circumstance and assume that $M_*(z) = M_i(z)$ (in contradiction with the boundary condition $M(z_*) = 0$), the expression for the current would diverge as $t \longrightarrow \tau_i(z_*)$ for $[t - \tau_i(z_*)]^{-1/2}$. Using the approximations $M_*(z) = (z - z_*)M_i(z_*)/\bar{\delta}$ for $z - z_* < \bar{\delta}$ and $M_*(z) = M_i(z)$ for $z - z_* > \bar{\delta}$, ignoring the time dependence of z_*, introducing the new variable $\theta = \int_{\tau_i(z_*)}^{t} D(t_1)\lambda_i^{-2}dt_1$, and taking into account that $I = -qSd/dt(\int_{z_*}^{h} Mdz)$, we obtain

$$I = qSM_{si}\exp\left[-\frac{z_*}{\lambda_i} - \int_0^t \frac{dt_1}{\tau(t_1)}\right]\left(\frac{1}{\tau} + \frac{D}{\lambda_i\bar{\delta}}\right), \quad \theta \ll \frac{\bar{\delta}^2}{\lambda_i^2}, \quad (22a)$$

$$I = qSM_{si}\exp\left[-\frac{z_*}{\lambda_i} - \int_0^t \frac{dt_1}{\tau(t_1)}\right]\left(\frac{1}{\tau} + \frac{D}{\lambda_i^2\sqrt{\pi\theta}}\right), \quad \frac{\bar{\delta}^2}{\lambda_i^2} \ll \theta \ll 1, \quad (22b)$$

$$I = \frac{qSM_{si}}{\sqrt{\pi\theta}}\exp\left[-\frac{z_*}{\lambda_i} - \int_0^t \frac{dt_1}{\tau(t_1)}\right]\left(\frac{1}{\tau} + \frac{D}{2\lambda_i^2\theta}\right), \quad (22c)$$
$$\theta \gg 1.$$

If $D\tau \gg \lambda_i^2$, the depolarization is controlled by the diffusion of neutral associations to the insulator surface. For T = const, the current is initially time-independent [see (22a)] and then begins to decrease as $[t - \tau_i(z_*)]^{-1/2}$ [see (22b)]; in the final stage, the current I falls as $[t - \tau_i(z_*)]^{-3/2}$ [see (22c)]. In the case of thermal stimulation, the curve $I(t)$ exhibits a peak. In the ascending portion of this curve, the current [see (22a) and (22b)] increases initially proportionally to $D(T)$ and, then, proportionally to $D^{1/2}(T)$. In the descending portion of the curve $I(t)$, we have, according to (22c), $I(t) \propto D^{-1/2}(T)$. For $\lambda_i\bar{\delta} \ll D\tau \ll \lambda_i^2$, the depolarization is initially related to diffusive broadening of the edge of the starting neutral-association distribution (the region $z - z_* < \bar{\delta}$) and is later controlled by disintegration of neutral associations in the bulk of the insulator. For T = const, the current decreases with time as $I \propto \exp(-t/\tau)$ [see (22a)] in the region of

$$t - \tau_i(z_*) < \bar{\delta}^2/D,$$

as $I \propto [t - \tau_i(z_*)]^{-1/2}$ in the range of

$$\bar{\delta}^2/D < t - \tau_i(z_*) < \tau^2 D/\pi\lambda_i^2,$$

and again as $I \propto \exp(-t/\tau)$ [see (22b)] in the region of

$$t > \tau^2 D/\pi\lambda_i^2.$$

In the case of thermal stimulation, as the temperature T increases, the ascending portion of the curve $I(t)$ is sequentially determined by the temperature dependences $D(T)$, $D^{1/2}(T)$, and $\tau^{-1}(T)$. For the descending portion of this curve, we have

$$I \propto \tau^{-1}\exp\left(-\int_0^t \tau^{-1}(t_1)dt_1\right).$$

If $D\tau \ll \lambda_i\bar{\delta}$, the decisive role is played by disintegration of neutral associations in the bulk of the insulator. The depolarization current follows the law

$$I \propto \tau^{-1}\exp\left(-\int_0^t \tau^{-1}(t_1)dt_1\right)$$

that is typical of monomolecular processes [21]. Thus, the temperature dependence of the depolarization current in the case of thermal stimulation exhibits two maximums. The first of these (the low-temperature maximum) is related to the transit of ions (for the ascending portion) and the tunneling disintegration of neutral associations concentrated in the layer $z < z_*$ (the descending portion). The second, high-temperature maximum is caused by diffusion of neutral associations (initially located at $z > z_*$) to the insulator–semiconductor interface and also by thermal disintegration of neutral associations in the bulk of the insulator. It is noteworthy that, if $z_* \gg \lambda_i$, the second maximum is not observed against the background of the first maximum.

We now consider the isothermal depolarization under the conditions of a linear voltage sweep with the constant speed

$$\beta_v = |dV_g/dt| = \text{const},$$

with V_g varying from positive to negative values. This mode is important for identification of details of the transport mechanism, because the I–V characteristic features a peak whose position is related to the properties of ions and to external factors (temperature and the sweep speed). Let us assume that, under steady-state conditions (V_g = const), the MIS structure is either polarized for $V_i > V_t$ or depolarized for $V_i < V_t$, where $V_t \gg kT/q$ is a certain threshold voltage that drops at the insulating layer. An asymmetry in the states of polarization and depolarization (i.e., the inequality $V_t \neq 0$) may be related, for example, to the higher ability of the ions to be neutralized at the gate than at the semiconductor surface. For large voltage-sweep speeds β_v, the charge transport takes place for the fields $F_h > 0$ ($V_i < 0$) and is controlled by the above-analyzed processes of disintegration of neutral associations and the transit of free ions through the insulator. Henceforth, when calculating the depolarization current, we assume that the sweep speed β_v is so low that the ion transport in the insulating layer occurs in the fields $F_h < 0$ ($V_i > 0$). In the insulator (in the vicinity of the semiconductor surface), there is ample opportunity for quasi-stationary spatial distributions of ions (for which the inequality $\beta_v \ll \mu F_h^2$ should hold) and neutral associations to be established. The latter assumption is justified if the rate of decrease of N_{st} with time ($N_{st} \equiv N_s + M_s$ is the total concentration of particles at the interface) is lower than the rate of disintegration of neutral associations in the bulk of the insulator; i.e., the condition $\tau_R \gg \tau$, where $\tau_R = N_{st}|\partial N_{st}/\partial t|^{-1}$, should be satisfied. On the other hand, we assume the value of β_v to be sufficiently large, so that the ion transport occurs under nonsteady conditions; i.e., we assume that $V_t - V_{im} \gg kT/q$, where V_{im} is the voltage drop across the insulator, which corresponds to the maximum of the depolarization current. In this case, we may ignore the reverse flux of particles that accumulate at the gate. The current through the sample consists of three components: the displacement current (related to the capacitance of the insulating layer), the current of electrons generated as a result of

disintegration of neutral associations (this current is negative), and the current of the ion transport above the barrier that has the height qV_i and is formed by the polarizing voltage; thus, we have

$$I = qS[(\kappa_i\beta_v/4\pi hq) - (M_s/\tau) + i_h],$$

$$i_h = [\mu NF - (\mu kT/q)dN/dz]|_{z=h}.$$

Under the quasi-steady-state conditions, the flux density of particles

$$\bar{i} = \mu NF - (\mu kT/q)dN/dz - DdM/dz$$

is independent of z. For sufficiently low-strength fields

$$(h/\lambda) - (q|F_h|h/kT) \gg 1$$

in a "thick" insulating layer ($h/\lambda \gg 1$), we have $i_h \simeq \bar{i}$. Assuming that $N|_{z=h} = 0$ and $qV_i \gg kT$ and using expressions (7), we obtain

$$\bar{i} = \frac{\mu|F_h|M_s}{\lambda}\left[\frac{\tau_{n0}\lambda}{\tau\delta_0} + \frac{qD}{\mu kT} + \exp\left(\frac{\Delta E_0}{kT}\right)\right]\exp\left(-\frac{qV_i}{kT}\right). \quad (23)$$

On integrating the sum of equations (1) and (2) with respect to the coordinate, we have

$$\frac{dN_{st}}{dt} = -\bar{i}. \quad (24)$$

Since the barrier is "high" ($qV_i \gg kT$), the drift and diffusion fluxes are approximately equal to one another ($\mu N|F| \gg \bar{i}$) and the dependences $N(z)$ and $M(z)$ are identical to those in the steady-state case, except for a thin layer in the vicinity of $z = h$, where $N(z) \ll N_0$ and $M(z) \ll M_0$. Thus, $M_s = (1 - \gamma)N_{st}$ and expressions (8) and (9), in which N_{s0} should be replaced with N_{st}, are valid for the degree of ionization γ. In the case of slight neutralization ($1 - \gamma \ll 1$), the semiconductor surface is in the state of enhancement:

$$V_s \ll V_i \quad \text{and} \quad dV_i/dt \simeq dV_g/dt = -\beta_v.$$

Using relationships (8), (9), (23), and (24) for

$$N_{st} \gg \kappa_i|F_h|/2\pi q,$$

we obtain[3]

$$I = qS\left\{\frac{\kappa_i\beta_v}{4\pi hq} + \frac{q\beta_v N_{s0}}{kT}\left[\sqrt{\left|\frac{F_{hm}}{F_h}\right|}\exp\left(\frac{qV_i - qV_{im}}{2kT}\right) + \sqrt{\left|\frac{F_h}{F_{hm}}\right|}\exp\left(\frac{qV_{im} - qV_i}{2kT}\right)\right]^{-2}\right\},$$

$$V_{im} = \frac{kT}{q}\ln\left(\frac{2\pi q\mu|F_{hm}|N_{s0}}{\kappa_i\beta_v}\right), \quad (25)$$

$$\tau_R = \frac{kT}{2q\beta_v}\left[1 + \left|\frac{F_{hm}}{F_h}\right|\exp\left(\frac{qV_i - qV_{im}}{kT}\right)\right].$$

In the case of $N_{st} \ll \kappa_i|F_h|/2\pi q$, we have

$$I = qS\left\{\frac{\kappa_i\beta_v}{4\pi hq} + \frac{q\beta_v N_{s0}}{kT}\left(\frac{F_h}{F_{hm}}\right)^2 \times \exp\left[\frac{qV_{im} - qV_i}{kT} - \left(\frac{F_h}{F_{hm}}\right)^2\exp\left(\frac{qV_{im} - qV_i}{2kT}\right)\right]\right\},$$

$$V_{im} = \frac{kT}{q}\ln\left(\frac{\mu F_{hm}^2}{\beta_v}\right), \quad (26)$$

$$\tau_R = \frac{kT}{q\beta_v}\left(\frac{F_{hm}}{F_h}\right)^2\exp\left(\frac{qV_i - qV_{im}}{kT}\right),$$

where F_{hm} is the field at the gate, which corresponds to the maximum of the current ($V_{im} \approx -hF_{hm}$). In expressions (25) and (26), the terms accounting for the electron current that accompanies the disintegration of neutral complexes are dropped. This is permissible in the vicinity of the peak in $I(t)$ for $q\beta_v\tau \geq kT$. The condition $\tau_R \gg \tau$ should be satisfied for the voltages V_i that correspond to the transition from the profound neutralization to the slight one. For lower voltages, a violation of this condition is unimportant; i.e., the delays in disintegration and diffusion of neutral associations affect the current only slightly because most of the particles are already ionized.

In the case of a large degree of neutralization, $M_s \simeq N_{st}$, and we use equations (23) and (24) to obtain

$$I = qS\left\{\frac{\kappa_i\beta_v}{4\pi hq} + \frac{N_{s0}}{\tau}\left[\left(1 + \frac{q\beta_v\tau}{kT}\right)\left|\frac{F_h}{F_{hm}}\right| \times \exp\left(\frac{qV_{im} - qV_i}{kT}\right) - 1\right]\exp\left[-\left(1 + \frac{kT}{q\beta_v\tau}\right) \times \left|\frac{F_h}{F_{hm}}\right|\exp\left(\frac{qV_{im} - qV_i}{kT}\right)\right]\right\},$$

$$V_{im} = \frac{kT}{q}\ln\left\{\frac{\mu|F_{hm}|\tau}{\lambda}\left(1 + \frac{q\beta_v\tau}{kT}\right)^{-1} \times \left[\frac{\tau_{n0}\lambda}{\tau\delta_0} + \frac{qD}{\mu T} + \exp\left(\frac{\Delta E_0}{kT}\right)\right]\right\}, \quad (27)$$

$$\tau_R = \tau\left(1 + \frac{q\beta_v\tau}{kT}\right)^{-1}\left|\frac{F_{hm}}{F_h}\right|\exp\left(\frac{qV_i - qV_{im}}{kT}\right).$$

[3] This result for an ensemble of free ions was obtained in [24].

The condition $\tau_R \gg \tau$ is not satisfied at the maximum of current (for $V_i = V_{im}$). This means that the ions leave the interface so rapidly that the neutral associations have no time to disintegrate; formula (27) is valid only for the ascending portion of the *I–V* curve in the region of voltages

$$V_i - V_{im} \gg (kT/q)[1 + (q\beta_v\tau/kT)].$$

In the case of slight neutralization, the dynamic *I–V* characteristic is represented by a narrow (with a width of $\sim kT/q$) peak of ion current against the background of two components: the constant displacement current and the electron current, which is caused by the disintegration of neutral associations and is small compared with the magnitude of the ion current. The ascending portion of the *I–V* curve away from its peak is independent of β_v; i.e., we have

$$(I - I_\infty) \propto \exp(-qV_i/kT),$$

where $I_\infty = I(V_i \longrightarrow \infty)$. For low concentrations of particles ($N_{st} \ll \kappa_i|F_h|/2\pi q$), the half-width of the peak on the ascending-curve side is larger than that on the descending-curve side. For large N_{st} ($N_{st} \gg \kappa_i|F_h|/2\pi q$), the peak is symmetric. The position of the peak shifts to depolarizing voltages as β_v increases and as N_{s0} decreases. In the case of profound neutralization, the peak of depolarization current is appreciably asymmetric; i.e., its ascending side is much steeper than the descending one. At the peak, the ionic component is of the same order of magnitude as the electronic component related to disintegration of neutral associations. In the ascending curve away from the peak, the current is independent of β_v and the difference $I - I_\infty$ increases with decreasing V_i in proportion to $\exp(-qV_i/kT)$. In the case of the descending curve, the current is controlled by non-steady-state disintegration of neutral associations and depends on real time rather than on voltage.

The theory developed here resolves the basic contradiction in the interpretation of experimental data on ionic depolarization of the Si-based MOS structures, the insulating layer of which appears as both a system of ionic traps with a wide distribution of lifetimes and as a medium with mobile (free) ions. Neutral associations fulfill the role of the traps; a wide distribution of lifetimes is caused by a spread of the tunneling distances of electrons generated as a result of disintegration of neutral associations; and an increase in the degree of ionization in the course of a quasi-steady decrease of the electric field ensures a gradual transition from a collection of low-mobility neutral associations to an ensemble of free ions. In this regard, the lower the concentration of free electrons in the bulk of the semiconductor, the easier it is to use an external field to vary the degree of neutralization of ions by shifting the Fermi level at the interface.

Transformation of the potential $U(\tilde{z})$ (Fig. 2) with increasing z and hybridization of the electron wave function in neutral association with the wave functions of the conduction band in the semiconductor results in modification of the electronic state (possibly, a virtual one) localized at the ion. The energy level of electron in neutral association rises as the distance from the semiconductor–insulator interface increases. Therefore, neutral associations escaping from the surface to the bulk of the insulator must overcome a certain potential barrier. Consequently, in the expression for the neutral-association flux (correspondingly, in equation (1) as well), a drift component should be involved along with the diffusion component; the former component is proportional to the mobility and the concentration of neutral associations and also to the force keeping these associations close to the surface. This force acts in a thin layer with a thickness of $\sim\delta_0$; the corresponding correction would not affect the results if the limiting factors in the steady-state mass transport are diffusion and disintegration of neutral associations in the bulk of the insulator rather than the rate of detachment of these associations from the interface. It is noteworthy that formation of an association may be accompanied by rearrangement of elastic stresses surrounding the ion in the insulator. In this case, the dependence $E_{NA}(z)$ may also be modified owing to the distinction between the elastic moduli of the semiconductor and the insulator.

A number of implications of the theory under consideration, including the persistent circulation of particles in the vicinity of the interface under the conditions of steady-state polarization, the quasi-hyperbolic law of decay of current in the cases of isothermal and thermally stimulated depolarization, and the shifts of the current peaks in dynamic *I–V* characteristics with changes in the voltage-sweep speeds and in the initial particle concentration, have been observed experimentally [17, 24, 25]. Two maximums in the temperature dependence of the current of thermally stimulated depolarization have also been observed. On account of the ambiguity concerning the interpretation of these observations, these maximums have been typically related to the presence of two types of ions in the insulator [6, 7]. In the context of the theory developed here, a criterion for the identification of such experimental data can be the dependence of the shape of the characteristic $I(T)$ on the heating rate β_T. Since z_* decreases with increasing T, an increase in β_T should be accompanied by a decrease in the area under the low-temperature peak in the curve $I(T)$ and by an increase in the area under the high-temperature peak.

ACKNOWLEDGMENTS

I thank A.G. Zhdan and G.V. Chucheva for their helpful participation in discussions of the problems studied in this paper.

REFERENCES

1. M. Yamin, IEEE Trans. Electron Devices **12**, 88 (1965).

2. S. Sze, in *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984), Subsection 7.3.2.
3. E. H. Nicollian and J. R. Brews, in *MOS (Metal Oxide Semiconductor) Physics and Technology* (New York, 1982), Chap. 15.
4. J. V. Verwey, E. A. Amerasekera, and J. Bisschop, Rep. Prog. Phys. **53**, 1297 (1990).
5. T. W. Hickmott, J. Appl. Phys. **46**, 2583 (1975).
6. M. R. Boundry and J. P. Stagg, J. Appl. Phys. **50**, 942 (1979).
7. J. P. Stagg and M. R. Boundry, J. Appl. Phys. **52**, 885 (1981).
8. M. W. Hillen and D. G. Hemmes, Solid-State Electron. **24**, 773 (1981).
9. C. Choquet, C. Plossu, M. Berenguer, and B. Balland, Thin Solid Films **167**, 45 (1988).
10. S. R. Hofstein, Appl. Phys. Lett. **10**, 291 (1967).
11. R. J. Kriegler and T. F. Devenyi, Thin Solid Films **36**, 435 (1976).
12. M. Kuhn and D. J. Silversmith, J. Electrochem. Soc. **118**, 966 (1971).
13. A. G. Tangena, J. Middelhock, and N. F. De Rooij, J. Appl. Phys. **49**, 2876 (1978).
14. A. G. Tangena, N. F. de Rooij, and J. Middelhock, J. Appl. Phys. **49**, 5576 (1978).
15. N. J. Chou, J. Electrochem. Soc. **118**, 601 (1971).
16. W. Marciniak and H. M. Przewlocki, J. Electrochem. Soc. **123**, 1207 (1976).
17. E. I. Gol'dman, A. G. Zhdan, and G. V. Chucheva, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33** (8), 933 (1999) [Semiconductors **33**, 852 (1999)].
18. T. Ando, A. B. Fowler, and F. Stern, in *Electronic Properties of Two-Dimensional Systems* (Am. Phys. Soc., New York, 1982; Mir, Moscow, 1985), Chap. 3.
19. I. Lundström and C. Svensson, J. Appl. Phys. **43**, 5045 (1972).
20. G. Greeuw and B. J. Hoenders, J. Appl. Phys. **55**, 3371 (1984).
21. E. I. Goldman and A. G. Zhdan, Semicond. Sci. Technol. **5**, 675 (1990).
22. J. G. Simmons and L. S. Wei, Solid-State Electron. **17**, 117 (1974).
23. M. A. Lampert and P. Mark, *Current Injection in Solids* (Academic, New York, 1970; Mir, Moscow, 1973).
24. E. I. Gol'dman, A. G. Zhdan, and G. V. Chucheva, Fiz. Tekh. Poluprovodn. (St. Petersburg) **31**, 1468 (1997) [Semiconductors **31**, 1268 (1997)].
25. E. I. Gol'dman, A. G. Zhdan, and G. V. Chucheva, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33** (8), 962 (1999) [Semiconductors **33**, 877 (1999)].

Translated by A. Spitsyn

Semiconductors, Vol. 34, No. 8, 2000, pp. 955–959. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 993–997.

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Properties of Diode Heterostructures Based on Nanocrystalline *n*-SnO_2 on *p*-Si under the Conditions of Gas Adsorption

R. B. Vasil'ev, A. M. Gas'kov, M. N. Rumyantseva, A. S. Ryzhikov, L. I. Ryabova*, and B. A. Akimov

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia
** e-mail: mila@mig.phys.msu.su*

Submitted February 14, 2000; accepted for publication February 24, 2000

Abstract—Heterostructures *n*-SnO_2(Me)/*p*-Si (Me = Cu, Pd, or Ni) with a SnO_2 layer thickness 0.8–1 μm were synthesized. The average size of tin dioxide crystallite was 6–8 nm. The current–voltage characteristics of the structures were measured in dry air under the conditions of NO_2 and C_2H_5OH adsorption. The current–voltage characteristics measured in ambient air were satisfactorily described within the diode theory with large nonideality factors: $\beta \sim 8$–18. Gas molecule adsorption substantially modifies the current–voltage characteristics in the range of both forward and reverse biases. The shape of the current–voltage characteristics was explained taking into account activation transitions through a barrier whose height can either decrease or increase depending on the adsorbed molecule type and changes in tunnel processes at the interface. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Nanocrystalline films of semiconductor oxides are widely used as sensitive units of gas sensors. Recently, along with resistive sensors, increasing attention has been paid to the development of various gas-sensitive heterostructures. In this case, the electrical properties of the system are mainly controlled by the behavior of the heteroboundary during gas adsorption. Progress was made by using metal–insulator–semiconductor (MIS) structures [1–3], Schottky diodes [4–6], and semiconductor–semiconductor heterojunctions [7–9]. Most of the efforts were dedicated to hydrogen adsorption studies.

According to conventional concepts, the basic mechanism responsible for gas sensitivity in MIS structures is the formation of a dipole layer from adsorbed gas molecules near the heteroboundary, which shifts the capacitance–voltage characteristics along the axis of external biases *U*. The sensitivity mechanism in Schottky diodes and heterojunctions is related to changing the barrier height at the heteroboundary, which changes the saturation current in the range of reverse biases. However, the sensitivity mechanism is not well understood in most cases.

Radically new properties were detected in *n*-SnO_2/SiO_2/*p*-Si heterostructures [10]: gas molecule adsorption caused a significant (orders of magnitude) and fast (seconds) increase in capacitance. The system could be reverted to its initial state by an electric field pulse.

To develop concepts on processes induced by gas adsorption in heterostructures, in this work we studied the impact of the gas-phase composition on the current–voltage characteristics of heterostructures based on nanocrystalline *n*-SnO_2 on *p*-Si. The porous structure of the nanocrystalline oxide layer allows gas molecules to diffuse immediately to the heteroboundary. Thus, one can expect an increased sensitivity of the structures to the gaseous-medium composition.

It is noteworthy that the possibility of diode-type heterojunction development using nanocrystalline materials on single-crystal silicon substrates is of independent interest.

SAMPLES

Layers of SnO_2 and SnO_2:Me (Me = Pd, Cu, or Ni) were synthesized by aerosol pyrolysis [11]. Tin dibutyl diacetate, palladium acetyl acetonate, copper acetyl acetonate trifluoride, and nickel ethyl hexanoate were used as source materials. Layers were deposited at a temperature of 520°C up to a thickness of 0.8–1.0 μm. Single-crystal *p*-Si with a resistivity of 10 Ω cm oriented along the crystallographic direction ⟨100⟩ was used as a substrate.

1063-7826/00/3408-0955$20.00 © 2000 MAIK "Nauka/Interperiodica"

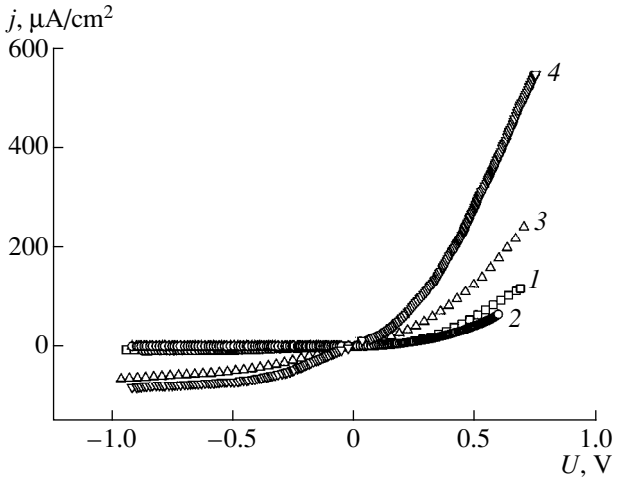


Fig. 1. Current–voltage characteristics of *n*-SnO_2(Me)/*p*-Si heterostructures with various dopants (Me): (*1*) *n*-SnO_2/*p*-Si, (*2*) *n*-SnO_2(Ni)/*p*-Si, (*3*) *n*-SnO_2(Pd)/*p*-Si, and (*4*) *n*-SnO_2(Cu)/*p*-Si.

Dopants were chosen to modify the electrical properties of the tin dioxide film in the widest possible range. As is known, copper, being an acceptor, can drastically increase the film resistance (R) due to the compensation of oxygen donor vacancies [12]. Doping with nickel also causes a significant resistance increase and gives rise to activation portions in the temperature dependences of R. As distinct from films doped with copper, in which case the activation energy E_A reaches 130 meV corresponding to the oxygen vacancy level [13] and does not increase further as the dopant concentration increases, a nickel content increase leads to a gradual E_A growth up to 500 meV, which is most likely caused by intercrystallite barrier growth [14]. Palladium is chosen as a catalytic dopant that activates gas molecule adsorption.

The microstructure of SnO_2 and SnO_2(Me) layers was studied by atomic-force microscopy and scanning electron microscopy (JSM-35 by Jeol). The layer structure was porous with an average agglomerate size of 0.05–0.10 μm. X-ray phase analysis was carried out with CuK_α emission using a Siemens diffractometer. The average size of SnO_2 crystallites, calculated from X-ray diffraction spectra by the Debye–Scherrer formula, was 6–8 nm. The dopant concentration in films was determined by X-ray electron probe analysis (Cameca-SX50). Calculated dopant concentrations are listed in the table. The uniform distribution of dopants over the film thickness was confirmed by mass spectrometry of secondary neutral particles (INA-3 by Leybold).

Parameters of samples and current–voltage characteristics (Fig. 3)

Sample no.	Dopant in SnO_2 (Me)	Dopant concentration in SnO_2, at. %	j_s, μA/cm²	β
1	–	–	7.3	8.5
2	Ni	0.8	9	10.7
3	Pd	1.4	75	18
4	Cu	0.6	88	13.5

Ohmic contacts of Au and Al to SnO_2 and Si layers, respectively, were formed by thermal evaporation in vacuum. The ohmic behavior of the contacts used was confirmed by special studies [15]. The SnO_2 active layer area was about 1 cm²; the gold contact area was smaller (about 0.07 cm²) in order to maintain free gas access to the film surface.

Current–voltage characteristics were measured at room temperature for a controlled gaseous-medium composition. Ethanol vapors and nitrogen dioxide NO_2 were used as adsorbed gases. As is known, an ethanol molecule adsorbed at the SnO_2 surface exhibits donor properties [16]. Nitrogen dioxide molecules are acceptors. Being adsorbed, these form a surface level in the SnO_2 forbidden band; this level lies below the chemisorbed oxygen acceptor level [17].

RESULTS

Current–voltage (I–V) characteristics $j(U)$ of the studied heterostructures measured in dry air at room temperature are plotted in Fig. 1. These I–V characteristics feature a distinct nonlinearity and pronounced portions of current density j_s saturation in the range of reverse biases and can be roughly approximated by the formula

$$j = j_s[\exp(eU/\beta kT) - 1] \quad (1)$$

with a large nonideality factor β. The I–V-characteristic shape substantially depends on the dopant. The highest density of saturation reverse current j_s was observed in *n*-SnO_2(Cu)/*p*-Si and *n*-SnO_2(Pd)/*p*-Si heterostructures. The values of j_s virtually coincide in the heterostructure with a tin dioxide layer undoped and doped with nickel. According to the data on tin dioxide films on insulating substrates, all the dopants used increase the film resistance. Thus, the saturation-current increase in heterostructures doped with copper and palladium correlates quite well with the resistance growth in tin dioxide films as a result of doping if the resistance change is assumed to be accompanied by a Fermi level lowering in a heterostructure component with electron conduction. The height of the barrier caused by band bending at the heterocontact decreases, while the reverse current grows. The absence of radical I–V characteristic changes in the case of nickel doping shows

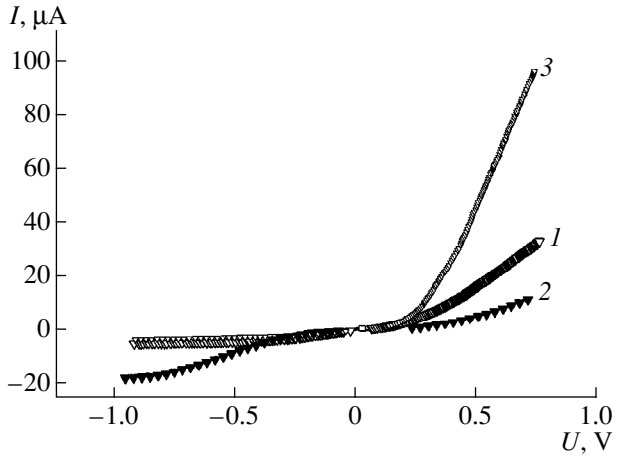


Fig. 2. Current–voltage characteristics of the n-SnO_2(Cu)/p-Si heterostructure measured in (*1*) dry air, (*2*) air with 1000 ppm of NO_2, and (*3*) air with 1000 ppm of C_2H_5OH.

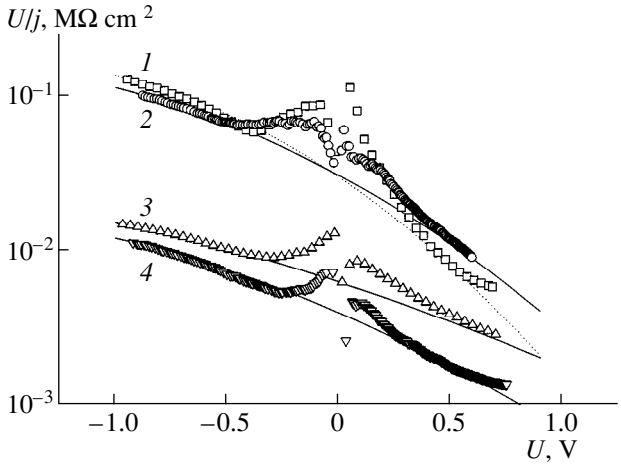


Fig. 3. Total resistance U/j versus the bias U for n-SnO_2(Me)/p-Si heterostructures with various dopants: (*1*) n-SnO_2/p-Si, (*2*) n-SnO_2(Ni)/p-Si, (*3*) n-SnO_2(Pd)/p-Si, and (*4*) n-SnO_2(Cu)/p-Si. Solid lines were calculated for Schottky diodes with parameters given in the table.

that the resistance growth in the nanocrystalline structure does not necessarily change the Fermi level position. One can assume this doping to change the current-flow behavior due to band relief modulation at nanocrystallite boundaries.

A gaseous-medium variation substantially changes the I–V characteristic both in the range of forward and reverse biases. As an example, Fig. 2 shows I–V characteristics of n-SnO_2(Cu)/p-Si heterostructures in ambient air (curve *1*), in air with 1000 ppm of NO_2 (*2*), and in air with the same ethanol concentration (*3*).

The impact of ethanol molecules appears similar for all heterostructures: the current I substantially grows and decreases in the range of forward and reverse biases, respectively. Nitrogen dioxide initiates the opposite processes in all structures except for those doped with nickel: the current decreases and grows for the forward and reverse biases, respectively. The current decreases in the entire range of applied biases in structures doped with nickel.

DISCUSSION

The shape of the heterostructures' I–V characteristics under the conditions of gas adsorption shows that the above processes cannot be described by considering a simplified the band relief in heterostructures and neglecting transition layers and charged states at the interface.

To distinguish the contributions of structure components to conduction, we describe the structure properties using the simplest equivalent circuits. To do this, it is convenient to consider the I–V characteristics in coordinates $U/j = f(U)$, where U/j is the total heterostructure resistance. Figure 3 displays experimental dependences $U/j = f(U)$ measured in ambient air and curves calculated by formula (1) (parameters j_s and β are listed in the table) for all studied heterostructures. As is evident from Fig. 3, calculated curves describing the Schottky diode with a large nonideality factor almost coincide with the experimental ones in the range of sufficiently large forward and reverse biases. In the range of small biases, dependences $U/j = f(U)$ exhibit

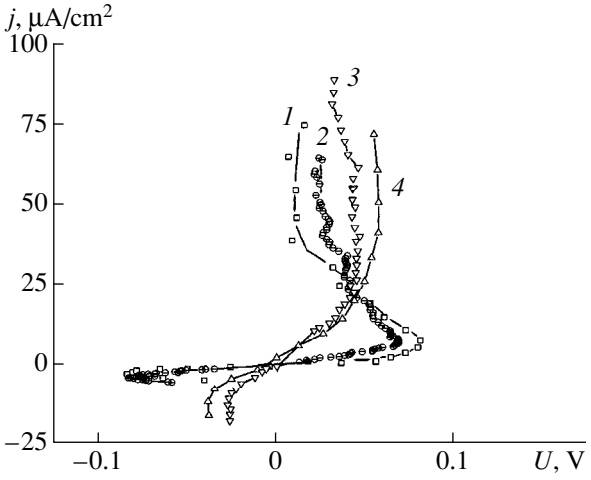


Fig. 4. Current–voltage characteristics calculated from experimental data for an element connected in series with the Schottky diode based on n-SnO_2(Me)/p-Si heterostructures with various dopants: (*1*) n-SnO_2/p-Si, (*2*) n-SnO_2(Ni)/p-Si, (*3*) n-SnO_2(Pd)/p-Si, and (*4*) n-SnO_2(Cu)/p-Si. Calculation parameters are given in the table. (See text for details.)

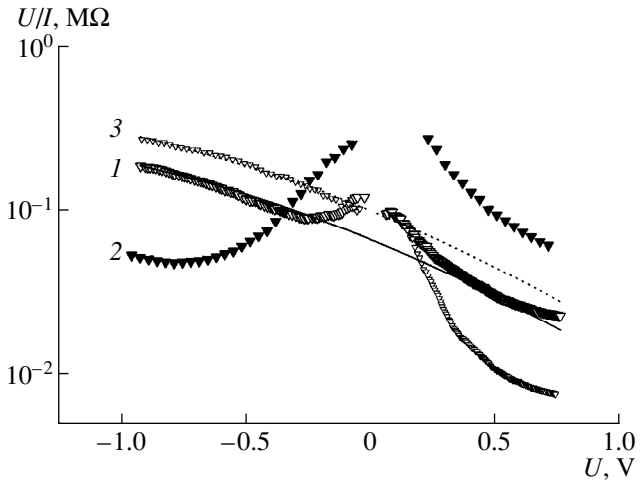


Fig. 5. Total resistance U/I of the n-SnO_2(Cu)/p-Si heterostructure versus the bias U in (*1*) dry air, (*2*) air with 1000 ppm of NO_2, and (*3*) air with 1000 ppm of C_2H_5OH. Solid lines were calculated for Schottky diodes with parameters: $\beta = 13.5$, I_s = (*1*) 5.3, and (*3*) 3.55 μA.

resistance growth that can be attributed to the resistance R_s connected in series. Assuming that the Schottky diode resistance caused by the barrier at the heterojunction is described by the calculated curve in weak fields, one can calculate the resistance R_s and the potential drop $U_s = R_s I$ across an element which can be conditionally considered as connected in series with the diode. The corresponding I–V characteristics plotted in Fig. 4 appear to be *S*-shaped. Taking into account that I–V characteristics of SnO_2(Me) films on insulating substrates are linear, a nonlinear element connected in series can be related to transition layers at the SnO_2–Si interface. These layers also account for large values of the structure nonideality factor.

The basic feature of I–V characteristics under the conditions of NO_2 and C_2H_5OH adsorption is that their forward and reverse portions cannot be described by formula (1) with fixed parameters j_s and β.

Actually, ethanol molecule adsorption decreases the saturation reverse current, which can be caused by the increased barrier height at the heteroboundary; i.e., this enhances bending of the energy bands. However, the current drastically grows in the range of forward biases. The observed forward current growth is most likely caused by tunneling through heteroboundary surface states related to adsorbed molecules rather than by the barrier change. Furthermore, the possible impact of molecules adsorbed near the heteroboundary on the lattice vibration spectrum should also be taken into account. The initiation of additional phonon modes can cause tunneling through the barrier.

The opposite processes were observed in the case of the NO_2-containing atmosphere: the current grows and decreases, respectively, in the range of reverse and forward biases. The barrier height can be assumed to decrease in this case; however, the probability of tunneling transitions through the barrier becomes lower.

Figure 5 illustrates the variation in heterostructure total resistance under the conditions of gas adsorption. The calculated dependences (solid lines) correspond to the constant nonideality factor $\beta = 13.5$ and different saturation currents, $I_s = 5.3$ and 3.55 μA for curves *1* and *3*, respectively. As is evident, the calculated curves describe the I–V characteristics reasonably well in the range of reverse biases when measurements are performed in ambient air and ethanol vapors and in the range of forward biases when measurements are performed in ambient air.

Thus, the experimental data indicate that the impact of gas adsorption on heterostructure I–V characteristics is controlled by at least two processes: barrier-height variation at the heteroboundary and tunneling through surface states at the interface. Superposition of these effects depends on the applied bias. The behavior of a current through a heterostructure under the conditions of gas adsorption depends both on the adsorbed molecule type and on the chosen external bias. The fact that the gas sensitivity of the considered heterostructures is actually controlled by processes at the heteroboundary also indicates that gas molecules diffuse freely enough through the nanocrystalline oxide layer.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project no. 98-03-32843.

REFERENCES

1. I. Lundstrom, S. Shivaraman, C. Svensson, and L. A. Lundkvist, Appl. Phys. Lett. **26**, 55 (1975).
2. L. G. Ekedahl, M. Eriksson, and I. Lundstrom, Acc. Chem. Res. **31**, 249 (1998).
3. W. P. Kang and C. K. Kim, Appl. Phys. Lett. **63**, 421 (1993).
4. A. D'Amico, G. Fortunato, G. Petrocco, and C. Coluzza, Appl. Phys. Lett. **42**, 964 (1983).
5. L. M. Lechuga, A. Calle, D. Golmayo, and F. Briones, Sens. Actuators B **4**, 515 (1991).
6. U. Hoefer, H. Bottner, E. Wagner, and C. D. Kohl, Sens. Actuators B **47**, 213 (1998).
7. Y. Ushio, M. Miyayama, and H. Yanagida, Sens. Actuators B **12**, 135 (1993).
8. Y. Ushio, M. Miyayama, and H. Yanagida, Sens. Actuators B **17**, 221 (1994).
9. W. Zhang, H. Uchida, T. Katsube, *et al.*, Sens. Actuators B **49**, 58 (1998).

10. R. B. Vasil'ev, M. N. Rumyantseva, L. I. Ryabova, *et al.*, Pis'ma Zh. Tekh. Fiz. **25** (12), 22 (1999) [Tech. Phys. Lett. **25**, 471 (1999)].
11. M. N. Rumyantseva, A. M. Gaskov, L. I. Ryabova, *et al.*, Mater. Sci. Eng. **41**, 333 (1996).
12. M. Rumyantseva, M. Labeau, G. Delabouglise, *et al.*, J. Mater. Chem. **7**, 1785 (1997).
13. C. G. Fonstad and R. H. Rediker, J. Appl. Phys. **42**, 2911 (1971).
14. B. A. Akimov, A. M. Gas'kov, M. Labo, *et al.*, Vestn. Mosk. Univ., Ser. 3: Fiz. Astron., No. 5, 60 (1996).
15. B. A. Akimov, A. M. Gas'kov, M. Labo, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 205 (1999) [Semiconductors **33**, 175 (1999)].
16. V. E. Henrich and P. A. Cox, *The Surface Science of Metal Oxides* (Cambridge Univ. Press, Cambridge, 1996).
17. J. A. Agapito and J. P. Santos, Sens. Actuators B **31**, 93 (1996).

Translated by A. Kazantsev

Semiconductors, Vol. 34, No. 8, 2000, pp. 960–964. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 998–1003.

LOW-DIMENSIONAL SYSTEMS

Band Offsets in $Zn_{1-x}Cd_xTe/ZnTe$ Single-Quantum-Well Structures Grown by Molecular-Beam Epitaxy on GaAs(001) Substrates

V. I. Kozlovskiĭ*, V. G. Litvinov, and Yu. G. Sadof'ev***

** Lebedev Physics Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia*

*** Ryazan State Radio Engineering Academy, ul. Gagarina 59/1, Ryazan, 391000 Russia*

Submitted February 2, 2000; accepted for publication February 3, 2000

Abstract—ZnTe heteroepitaxial layers and $ZnTe/Zn_{1-x}Cd_xTe/ZnTe$ strained quantum-confinement structures grown by molecular-beam epitaxy on GaAs(001) were studied by low-temperature cathodoluminescence spectroscopy and current-relaxation deep-level transient spectroscopy (DLTS). A peak related to electron emission from the ground size-quantization level in the conduction-band was observed in the DLTS spectra of quantum-confinement structures. The conduction–band offset parameter Q_C was determined from the DLTS and cathodoluminescence data. For $Zn_{1-x}Cd_xTe/ZnTe$ single-quantum-well structures with $x = 0.2$–0.22, Q_C equals 0.82 ± 0.05. The effect of internal elastic strain on the band offsets and Q_C at the $Cd_xZn_{1-x}Te$ quantum well interfaces was calculated; the results of calculations agree well with experimental data. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Quantum-confinement structures based on the $ZnTe/Zn_{1-x}Cd_xTe$ heteropair are promising for the development of visible-light sources and modulators. The conduction- and valence-band offsets ΔE_C and ΔE_V at the interface, which determine the potential barrier height for electrons and holes, are basic parameters characterizing a heterostructure. However, reliable data about them are still lacking. For instance, the values reported in the literature for the conduction-band offset parameter $Q_C = \Delta E_C/\Delta E_G$ (where $\Delta E_G = \Delta E_C + \Delta E_V$ is the bandgap discontinuity) in the ZnCdTe/CdTe heterojunction vary from 0.775 to 1.2 [1–4]. Among other reasons, this spread in the data may result from the significant dependence of the band offsets on the internal elastic strain that appears due to the lattice-constant mismatch (which amounts to 6% for the ZnTe/CdTe pair). Note that the strain distribution over the layers depends on the specific design of the structure.

Usually, optical diagnostics techniques are used to study the energy bands in the II–VI compounds. However, valuable information can also be obtained from electrical measurements. In particular, in recent years deep-level transient spectroscopy (DLTS) [7] has been used to determine the band discontinuities in AlGaAs/GaAs [5] and SiGe/Si [6] heterojunctions. The quantum well (QW) may be considered as a giant trap capable of capturing electrons and holes. Thermal emission of the captured charge carriers, which takes place in the course of a DLTS scan, gives information on the energy separation between the quantum-confinement levels in the QW and the barrier conduction- and valence-band edges. If the energies of the quantum-confinement levels are known, ΔE_C and ΔE_V can be determined. In addition, the DLTS technique provides information on the parameters of the deep-level defects; some of them cannot be detected by the commonly employed optical diagnostic methods.

2. EXPERIMENTAL

The following four samples were studied: a ZnTe epitaxial layer (sample 160) and three structures with individual ZnCdTe/ZnTe QWs of different compositions and widths (samples 195, 196, and 197) (see Tables 1, 2).

The samples were grown by molecular-beam epitaxy on n^+-GaAs(100) substrates misoriented by 3° to the ⟨110⟩ direction. The epitaxial ZnTe layer and buffer ZnTe layers on which the quantum-confinement structures were grown were 1.5-μm thick and were grown at 350°C; surface reconstructions $(1 \times 2) + c(2 \times 2)$ detected by high-energy electron diffraction (HEED) coexisted under growth conditions. The growth rate of ZnTe layers was 2 Å/s, and the ratio of Zn and Te molecular-beam equivalent pressures was 1 : 2. During the growth of the quantum-confinement structures, the growth rate and the substrate temperature were reduced to 1 Å/s and 280°C, respectively. The composition and thickness of the Cd-containing layers were determined both from the variation in the period of HEED reflection-intensity oscillations and from the dependences of the molecular beam intensities on the evaporation temperature. The thickness of the ZnTe cap layer was 250 nm. The main parameters of the single-QW samples ZnCdTe/ZnTe are given in the left-hand columns of Table 1.

1063-7826/00/3408-0960$20.00

Table 1. Parameters of single-QW $Zn_{1-x}Cd_xTe$ structures under study

Sample no.	CdTe content, x	L_w, nm	E_{CL}, eV	ΔE_{CL}, meV	E_{add}, meV	E_{e1}, meV	E_{hh1}, meV	Q_C
195	0.22	3.5	2.246	134	101	80	16	0.823 ± 0.045
196	0.22	6.5	2.206	174	133	38	7.5	0.816 ± 0.048
197	0.2	5.5	2.226	154	116	46	9	0.814 ± 0.05

Note: L_w is the QW width; E_{CL} and ΔE_{CL} are the QW emission peak energy position and shift relative to the free exciton emission line in the barrier, respectively; E_{add} is the activation energy corresponding to the current-relaxation DLTS peak related to electron emission from the QW; E_{e1} and E_{hh1} are the energies of the ground quantum-confinement levels in the conduction and the valence band, respectively, measured from the corresponding band edge; and Q_C is the conduction-band offset parameter.

The samples were studied by DLTS and low-temperature (14 K) cathodoluminescence. The cathodoluminescence was excited by a beam of 10-keV electrons with a current density of 0.1 mA/cm^2; the diameter of the electron spot on the sample was 1 mm. To perform the DLTS measurements, diode structures with a 1-mm^2 field-electrode area were fabricated. Contacts were deposited by the thermal evaporation of indium and nickel on the rear surface of the substrate and on the capping ZnTe layer, respectively.

All samples possessed high resistivity because of self-compensation in the epitaxial layers, and the structure capacitance was independent of the bias voltage. This means that the commonly used DLTS scheme, which relies on the capacitance transients, cannot be applied here. Thus, we made use of the current transients to record the DLTS curves (the current-relaxation DLTS technique); in other respects, the measurement procedure did not differ from the case of conventional DLTS.

3. RESULTS

The cathodoluminescence spectrum of an epitaxial ZnTe layer features a number of narrow peaks in the edge region, which originate from the emission of free and bound excitons. There is also a broad emission band at 600 nm caused by intrinsic defects related to structural imperfections in the film. In the cathodoluminescence spectra of single-QW samples, an additional intense line I_{qw}, whose spectral position depends on the QW parameters, is observed. This line exceeds in intensity the lines related to emission from the ZnTe buffer layer by 2–3 orders of magnitude. Its full width at half-maximum is 8–10 meV, which is indicative of the high quality of the structures. The energies E_{CL} of the I_{qw} lines in different samples are listed in Table 1.

Two peaks (labeled $E1$ and $E2$ in Fig. 1a) related to the deep-level centers were observed in the current-relaxation DLTS spectrum of the epitaxial ZnTe layer. Their activation energies ΔE_t equal 0.21 ± 0.02 ($E1$) and 0.58 ± 0.02 eV ($E2$), respectively. The intensity of the $E1$ peak is low, and it is absent in the spectra of the quantum-confinement structures. The $E2$ peak can be attributed to an electron trap located 0.6–0.62 eV below the bottom of the conduction band [8, 9], commonly related to Te vacancies or Zn interstitials present in the material.

In the DLTS spectrum of the samples with quantum-confinement structures (Fig. 1b), in addition to the $E2$ peak, a new, relatively intense peak E_{add} appears in the low-temperature region (77–120 K). With the lowest temperature attainable in a DLTS scan being equal to 77 K, this peak can be detected for the emission-rate windows corresponding to relaxation time constants less than 5×10^{-4} s. The activation energy of the trap related to this additional peak varies in different structures from 101 to 133 meV; the values of E_{add} are given in Table 1. We believe that these DLTS peaks are related to the electron emission from the size-quantization levels formed in the QW conduction band. In Table 2, we list the effective volume concentrations N_t of the carriers captured at the E_{add} levels, calculated using the conventional DLTS formalism. Taking into account the fact that the carriers are captured in the QWs, we can estimate the density n_s of the carriers captured at the quantum-confinement levels by the formula $n_s = N_t L_w$ (where L_w is the QW width). The value of n_s obtained in this way does not exceed 10^{10} cm^{-2}. This indicates that only the ground quantum-confinement level in the QW is occupied.

4. DISCUSSION

Next, we use the above results to evaluate the conduction-band offset parameter $Q_C = \Delta E_C/\Delta E_G$. ΔE_C can be determined as a sum of two quantities: (i) the energy separation E_{add} between the ground quantum-confinement level $e1$ and the bottom of the ZnTe barrier conduction band (obtained from the DLTS data) and (ii) the quantum-confinement energy E_{e1} of an electron at the ground level in ZnCdTe QW. The latter can be calculated from the known QW parameters (width, Cd content, and electron effective mass in the well and barrier layers) in the model of rectangular QW of a finite depth with wide barriers [10]. Taking into account that the electron effective masses in the well and barrier layers

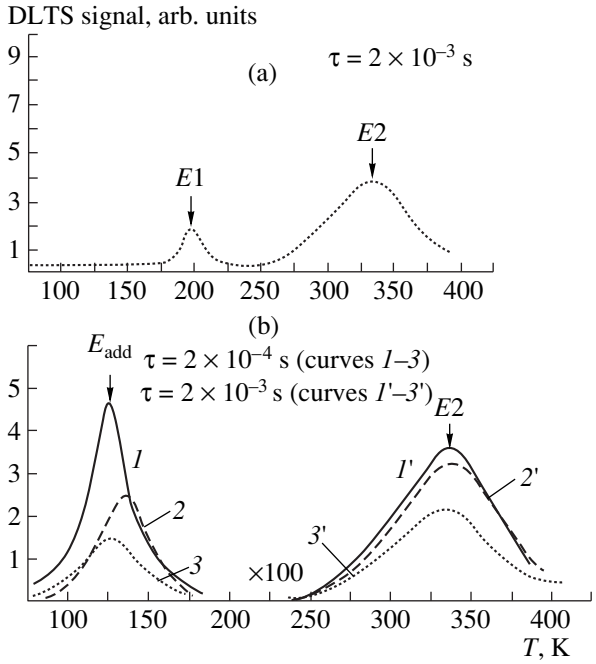


Fig. 1. DLTS spectra of (a) an epitaxial ZnTe layer (sample 160) and (b) single-QW ZnCdTe/ZnTe structures, samples 195 (curves *1* and *1*'), 196 (curves *2* and *2*'), and 197 (curves *3* and *3*') (see Table 1). The scans were performed under reverse bias voltage V_r = –1 V and filling voltage V_f = 0.

are virtually equal ($m_{e,qw} = m_{e,b} = m_e = 0.11m_0$ [4]), we obtain the following relationship for E_{e1}:

$$L_w(2m_eE_{e1}/\hbar^2)^{0.5} = \pi - 2\arcsin[(E_{e1}/(E_{e1} + E_{\text{add}}))^{0.5}]. \quad (1)$$

The band gap of the ZnCdTe layer, forming the QW, and the band-gap difference between the barrier and the well layers can be expressed in terms of the QW cathodoluminescence-peak energy E_{CL} and its shift ΔE_{CL} (Table 1) relative to the free exciton energy in the cathodoluminescence spectra of the epitaxial ZnTe film in the samples considered:

$$E_G(\text{ZnCdTe}) = E_{\text{CL}} + E_{hh}^b - E_{e1} - E_{hh1}, \quad (2)$$

$$\Delta E_G = \Delta E_V + \Delta E_C = \Delta E_V + E_{\text{add}} + E_{e1} = \Delta E_{\text{CL}} - \Delta E_{hh}^b + E_{e1} + E_{hh1}. \quad (2a)$$

Here, E_{hh}^b is the heavy-hole exciton binding energy and E_{hh1} is the size quantization energy (measured from the top of the valence band in the QW) of the heavy hole at ground level in the well. It is assumed that the QW cathodoluminescence line originates from the free heavy-hole exciton recombination. Next, we make use of the fact that in narrow QWs, in which the exciton Bohr radius (~5.5 nm) becomes comparable to the well width, the exciton binding energy E_{hh}^b increases by about 10 meV [11]. E_{hh1} is related to the valence band offset ΔE_V by the following equation:

$$L_w(2m_{hh}E_{hh1}/\hbar^2)^{0.5} = \pi - 2\arcsin[(E_{hh1}/\Delta E_V)^{0.5}]. \quad (3)$$

Here, the heavy-hole effective mass $m_{hh} = 0.6m_0$ [4]. Solving equation set (1)–(3) numerically, we can obtain E_{e1}, E_{hh1}, and Q_C from the cathodoluminescence and DLTS data without any need to calculate elastic strains. The Q_C values thus derived are listed in Table 1. We found that $Q_C \approx 0.82$ for the CdTe content in the QW x = 0.2–0.22.

We now calculate Q_C theoretically, taking into account the influence of the internal elastic strain on the band offsets. The conduction-band offset between two unstrained semiconductors is equal to the difference between the corresponding electron affinities [12]. Since the electron affinity equals 3.53 and 4.28 eV and the band gaps equal 2.39 and 1.60 eV for ZnTe and CdTe [1, 13], respectively, the values of ΔE_C, ΔE_V, and Q_{C0} for these two semiconductors in the absence of

Table 2. Parameters of the deep levels revealed by the current-relaxation DLTS measurements

Sample no.	Level type	ΔE_t, eV	Capture cross section $\sigma_\infty \exp(\Delta E_B/kT)$, cm^{-2}	N_t, cm^{-3}
160	*E*1	0.21 ± 0.01	4×10^{-16}	1.8×10^{13}
	*E*2	0.58 ± 0.02	8.6×10^{-17}	3.8×10^{13}
195	E_{add}	0.101 ± 0.01	6×10^{-16}	3.3×10^{15}
	*E*2	0.58 ± 0.02	1.3×10^{-16}	3.6×10^{13}
196	E_{add}	0.133 ± 0.01	2.6×10^{-14}	2.6×10^{15}
	*E*2	0.58 ± 0.02	3.4×10^{-16}	3.4×10^{13}
197	E_{add}	0.116 ± 0.01	3.5×10^{-14}	1.4×10^{15}
	*E*2	0.58 ± 0.02	5×10^{-16}	2.1×10^{13}

Note: ΔE_t and N_t are the activation energy and bulk concentration of deep-level centers; σ_∞ is the temperature-independent prefactor in the capture cross section; and ΔE_B is the activation energy for the temperature-dependent factor in the capture cross section.

strain would be 0.75, 0.04, and 0.95 eV, respectively. If the QW is composed of a $Zn_{1-x}Cd_xTe$ alloy, then

$$Q_{C0} = [E_G(\mathrm{ZnTe}) - E_G(x) - 0.04x]/[E_G(\mathrm{ZnTe}) - E_G(x)], \quad (4)$$

where allowance is made for nonlinearity of the $E_G(x)$ dependence (with the sagging parameter B)

$$E_G(x) = E_G(\mathrm{ZnTe}) - [E_G(\mathrm{ZnTe}) - E_G(\mathrm{CdTe})]x - Bx(1-x) \quad (5)$$

and linear variation of the valence-band offset is assumed. Using the value $B = 0.34$ calculated following [14], we obtain the dependence $Q_{C0}(x)$ plotted by the dashed line (curve *1*) in Fig. 2.

Next, let us consider the effect of the internal elastic stress. We assume that the structure is isomorphic in the vicinity of the QW, and the in-plane lattice constant (in general, different from the lattice constants a_b and a_{qw} of the free bulk ZnTe and ZnCdTe crystals) equals $a_\parallel$. Then, internal elastic stresses X_b and X_{qw} appear in the barrier and the well layers, according to the formula

$$X_i = (a_\parallel - a_i)/(a_i C_{11i}), \quad (6)$$

where C_{11i} is the elastic constant and subscript i stands for symbols b for the barrier and qw for the QW.

According to [15], elastic stresses result in the shifts ΔU_{ci}, ΔU_{hhi}, and ΔU_{lhi} of the conduction band and heavy- and light-hole valence bands, respectively, given by

$$\Delta U_{ci} = 2a_{ci}(C_{11i} - C_{12i})X_i,$$

$$\Delta U_{hhi} = -2a_{vi}(C_{11i} - C_{12i})X_i + b_i(C_{11i} + 2C_{12i})X_i,$$

$$\Delta U_{lhi} = -2a_{vi}(C_{11i} - C_{12i})X_i - 0.5b_i(C_{11i} + 2C_{12i})X_i - 0.5\delta_i + 0.5\{\delta_i^2 + 4.5[2b_i(C_{11i} + 2C_{12i})X_i]^2 - 2b_i(C_{11i} + 2C_{12i})X_i\delta_i\}^{0.5}, \quad (7)$$

where a_{ci} and a_{vi} are the hydrostatic deformation potentials of the conduction and valence bands, respectively; b_i is the shear deformation potential of the valence band; δ_i is the spin–orbit splitting; and C_{11i} and C_{12i} are the elastic constants. Deformation potentials, elastic constants, and lattice constants for the binary compounds ZnTe and CdTe are well known; they are given in Table 3. Parameters characterizing the $Cd_xZn_{1-x}Te$ layer were obtained by linear interpolation between the corresponding ZnTe and CdTe values.

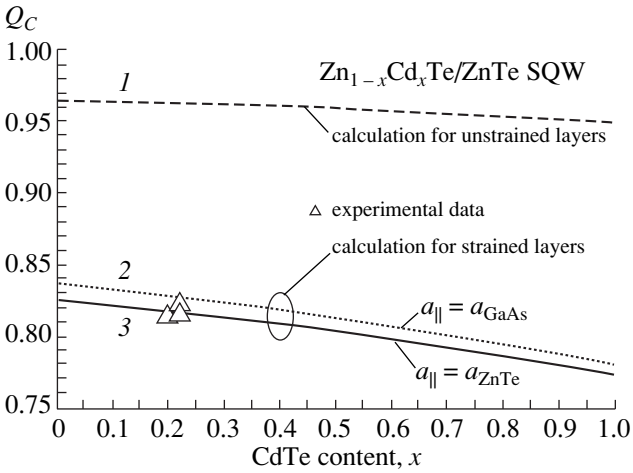


Fig. 2. Dependence of the conduction-band offset parameter Q_C on the CdTe content x in the QW. The calculation was carried out under the following conditions: (*1*) elastic strain in the layers is disregarded and (*2*, *3*) elastic strain is accounted for in isomorphic structures with $a_\parallel$ = (*2*) a_{GaAs} or (*3*) a_{ZnTe}. Experimental data are depicted by triangles.

First, we consider the case of $a_\parallel = a_{ZnTe}$. It corresponds satisfactorily to the properties of actual structures under investigation, because the thickness of the ZnTe buffer layer, 1.5 μm, is far in excess of the critical thickness for isomorphic growth of the ZnTe film on GaAs (which, according to [20], is about 1.5 nm). Thus, it is reasonable to assume that the ZnTe lattice in the vicinity of the QW is completely relaxed due to the generation of misfit dislocations at the GaAs–ZnTe interface. On the other hand, the QW width in the samples under study (3.5–6.5 nm) is much smaller than the critical thickness for isomorphic growth of $Zn_{0.78}Cd_{0.22}Te$ on ZnTe (about 17 nm). In this case, the QW is elastically compressed, and the barrier ZnTe layers remain virtually unstrained. The top of the heavy-hole valence band in the QW will always be at higher energies than the top of the light-hole band. The depth of the potential well for heavy holes increases due to the elastic strain, while the potential energy of the light hole in the QW becomes higher than in the ZnTe barrier (the discontinuity of the second type). Since the band gap in the QW material is controlled by the energy

Table 3. Parameters used in the calculation of the band-edge shifts caused by elastic strains [16–19]

Material	a, nm	C_{11}, rel. units	C_{12}, rel. units	a_c, eV	a_v, eV	b, eV
ZnTe	0.6104	7.13	4.07	–4.829	–0.654	–1.3
CdTe	0.648	5.62	3.94	–3.38	–0.4695	–1.15

Note: a is the lattice constant; C_{11} and C_{12} are the elastic constants; and a_c, a_v, and b are the deformation potentials.

position of the top of the heavy-hole valence band, Q_C can be calculated by the following formula:

$$Q_C(x) = [E_G(\mathrm{ZnTe}) - E_G(x) - 0.04x - \Delta U_c(x)]/[E_G(\mathrm{ZnTe}) - E_G(x) - \Delta U_c(x) + \Delta U_{hh}(x)]. \quad (8)$$

The dependence $Q_C(x)$ calculated from (8) is shown by the solid line in Fig. 2 (curve *3*). A steady decrease of Q_C with the Cd content in the QW is related to the difference between the conduction- and valence-band deformation potentials a_c and a_v in the QW material and between the uniaxial-deformation potentials b for ZnTe and CdTe. The values of Q_C determined from the experimental data are shown in Fig. 2 by triangles. One can see that the calculated curves fit the experimental points well. It should be mentioned that the position of the curve $Q_C(x)$ in the plot depends only slightly on the value of $a_\parallel$. This is illustrated by curve *2* (dotted line in Fig. 2), calculated for the hypothetical case where the lattice constant is defined by a highly mismatched substrate: $a_\parallel = a_{\mathrm{GaAs}}$. Even then, the shift of the $Q_C(x)$ curve is insignificant. This remains all the more true if $a_\parallel$ varies between a_b and a_{qw}. This means that, for ZnCdTe/ZnTe multiple-well structures or superlattices, the band offsets are nearly independent of the relation between the QW and barrier widths.

5. CONCLUSION

A signal in the current-relaxation DLTS spectra originating from the electron emission from the ground quantum-confinement level in the conduction band of a high-resistivity ZnCdTe/ZnTe single-QW heterostructure has been observed. The activation energy of this level correlates with the energy position of the QW emission line in the cathodoluminescence spectra. A method to determine the conduction-band offset parameter Q_C from the DLTS and cathodoluminescence data and the known QW width is proposed. It is established that, for single-QW $Zn_{1-x}Cd_xTe/ZnTe$ heterostructures with $x = 0.2$–0.22, $Q_C = 0.82 \pm 0.05$. It is demonstrated that the band offset is primarily determined by the internal elastic strain that appears in the QW layer due to the lattice mismatch with the barrier layers. The calculated $Q_C(x)$ curve agrees well with the experimental data.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, project no. 98-02-16890.

REFERENCES

1. H. Mariette, F. Dal'bo, N. Magnea, *et al.*, Phys. Rev. B **38**, 12443 (1988).
2. P. Peyla, Y. Merle d'Aubigné, A. Wasiela, *et al.*, Phys. Rev. B **46**, 1557 (1992).
3. Tran Min Duc, C. Hsu, and J. P. Faurie, Phys. Rev. Lett. **58**, 1127 (1987).
4. T. W. Kim and H. L. Park, J. Cryst. Growth **159**, 467 (1996).
5. O. Chretien, R. Apetz, L. Vescan, *et al.*, J. Appl. Phys. **78**, 5439 (1995).
6. K. L. Jiao and W. A. Anderson, J. Appl. Phys. **73**, 271 (1993).
7. D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).
8. C. B. Norris, J. Appl. Phys. **53**, 5172 (1982).
9. P. S. Kireev, A. G. Kornitskiĭ, V. N. Martynov, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **4**, 900 (1970) [Sov. Phys. Semicond. **4**, 762 (1970)].
10. G. Bastard and J. A. Brum, IEEE J. Quantum Electron. **QE-22**, 1625 (1986).
11. H. Mathieu, P. Lefebvre, and P. Christol, Phys. Rev. B **46**, 4092 (1992).
12. H. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers* (Academic, New York, 1978; Mir, Moscow, 1981), Part B.
13. R.K. Swank, Phys. Rev. **153**, 844 (1967).
14. H. J. Lozykowski and V. K. Shastri, J. Appl. Phys. **69**, 3235 (1991).
15. F. K. Pollak and M. Cardona, Phys. Rev. B **172**, 816 (1968).
16. C. G. Van de Walle, Phys. Rev. B **39**, 1871 (1989).
17. W. Wardzynski, W. Giriat, H. Szymaczk, and R. Kowalczyk, Phys. Status Solidi B **49**, 71 (1972).
18. M. Zigone, H. Roux-Buisson, H. Tuffigo, *et al.*, Semicond. Sci. Technol. **6**, 454 (1991).
19. D. J. Thomas, J. Appl. Phys. **32**, 2298 (1961).
20. P. J. Parbrook, B. Henderson, K. P. O'Donnel, *et al.*, J. Cryst. Growth **117**, 492 (1992).

Translated by M. Skorikov

Semiconductors, Vol. 34, No. 8, 2000, pp. 965–970. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 1004–1009.

LOW-DIMENSIONAL SYSTEMS

The Influence of Irradiation and Subsequent Annealing on Si Nanocrystals Formed in SiO_2 Layers

G. A. Kachurin*, S. G. Yanovskaya*, M.-O. Ruault, A. K. Gutakovskiĭ*, K. S. Zhuravlev*, O. Kaitasov**, and H. Bernas****

** Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, pr. akademika Lavrent'eva 13, Novosibirsk, 630090 Russia*

e-mail: kachurin@isp.nsc.ru

*** CSNSM-CNRS/IN2P3, 91405 Orsay, France*

Submitted February 1, 2000; accepted for publication February 14, 2000

Abstract—Luminescent Si nanocrystals formed in SiO_2 layers were irradiated with electrons and He^+ ions with energies of 400 and 25–130 keV, respectively. The effects of irradiation and subsequent annealing at 600–1000°C were studied by the methods of photoluminescence and electron microscopy. After irradiation with low doses (~1 displacement per nanocrystal), it was found that photoluminescence of nanocrystals was quenched but the number of them increased simultaneously. After irradiation with high doses (~10^3 displacements per nanocrystal), amorphization was observed, which is not characteristic of bulk Si. The observed phenomena are explained in terms of the generation of point defects and their trapping by Si–SiO_2 interfaces. Photoluminescence of nanocrystals is recovered at annealing temperatures below 800°C; however, an annealing temperature of about 1000°C is required to crystallize the precipitates. An enhancement of photoluminescence observed after annealing is explained by the fact that the intensities of photoluminescence originated from initial nanocrystals and from nanocrystals formed as a result irradiation are summed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The persistent reduction of sizes of semiconductor devices and the discovery of intense visible luminescence in silicon quantum-confinement structures have made the study of Si nanocrystals very topical. The properties of nanocrystals, including the radiation effects, may differ markedly from the properties of bulk material owing to the size effects and to the drastically increasing role of the surface.

Until now, the influence of radiation on silicon nanocrystals has been poorly understood; in most cases, porous Si has been studied. It was reported that ion bombardment detrimentally affects the visible luminescence of porous Si [1–7]. Quenching of photoluminescence (PL) was related to amorphization [1–5], hydrogen desorption [6], and the formation of nonradiative-recombination centers [7]. A decrease in the PL intensity was also observed after irradiation with electrons, positrons, alpha particles, and X- and gamma-rays [8–12]. In contrast with ions of medium masses, the above types of radiation do not amorphize conventional Si. Quenching of PL was related to a decrease in the number of nanocrystals [8–10]. However, a decrease in the luminescence intensity as a result of irradiation with 20-keV electrons was observed [11]; this energy of electrons is much lower than that corresponding to the threshold of atomic displacements. This decrease in intensity was attributed to the evaporation of hydrogen under the effect of an electron beam.

There is also evidence that radiation may stimulate the PL. Thus, it was observed that the gamma radiation enhanced PL and made it more stable [13]. However, this effect was observed only in the case of irradiation under atmospheric conditions and was not observed if the irradiation was performed in vacuum. An enhancement of PL was also observed following irradiation with gamma quanta, electrons, and Si ions and was explained by an increase in the number of surface O–Si–H bonds [14]. Finally, it was concluded [15] that the luminescent porous Si is more radiation-resistant than the bulk material.

Thus, data on the interaction of radiation with nanocrystals are scarce and contradictory; the contradictions are related to a large extent to the fact that the surfaces in porous Si are unprotected. The question of the annealing of radiation damage also remains open. Recently, methods for forming nanostructures with the nanocrystals' surface reliably protected have been suggested. In particular, one of these methods is the ion-implantation synthesis of Si nanocrystals in SiO_2 [16–19]. Therefore, the aim of this study was to investigate the influence of radiation and subsequent annealing on the properties of silicon nanocrystals formed within SiO_2. The samples were irradiated with fast electrons and He^+ ions that produce only mobile point defects in Si.

1063-7826/00/3408-0965$20.00 © 2000 MAIK "Nauka/Interperiodica"

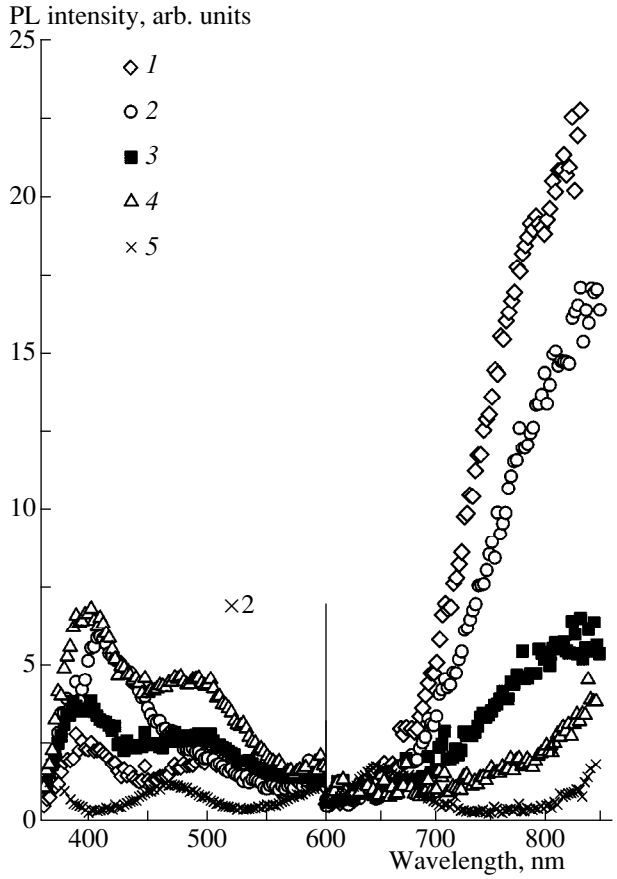


Fig. 1. PL spectra of SiO_2 layers with embedded Si nanocrystals (*1*) before and after irradiation with He^+ ions with doses of (*2*) 3×10^{12}, (*3*) 1×10^{13}, and (*4*) 3×10^{13} cm^{-2}. Spectrum *5* was measured after the SiO_2 layer without Si nanocrystals had been irradiated with He^+ ions at a dose of 1×10^{15} cm^{-2}.

EXPERIMENTAL

In order to form the nanocrystals, Si^+ ions were implanted into the submicrometer SiO_2 layers grown thermally on silicon. Doses of ions of 5×10^{16}–1×10^{17} cm^{-2} and energies of 25–130 keV ensured an excess concentration of Si in SiO_2 that amounted to 15–20 at. % at a depth down to 200 nm. After implantation, the samples were annealed for 30 min at 1150°C in the nitrogen atmosphere. The high-resolution electron microscopy (HREM) study showed that, as a result of annealing, silicon nanocrystals were formed in SiO_2. Their parameters, shape, and sizes were consistent with the data reported previously [20, 21]. The nanocrystals formed in SiO_2 were subjected to room-temperature irradiation with 400-keV electrons (in the electron-microscope column) and He^+ ions with energies of 25–130 keV. The energies of He^+ ions were chosen such that the ions penetrated through the entire SiO_2 layer and only the electronic energy losses were significant. The ion-current densities did not exceed 1 μA/cm^2, and the electron beam was focused to a spot ~2 μm in diameter, with a beam current of 2.5 μA. In both cases, heating of the samples by the beam was insignificant. The postirradiation annealings were performed either in N_2 or in vacuum. A fraction of experiments with irradiation with He^+ ions and subsequent annealing were carried out *in situ* in the electron-microscope column. The samples were studied by the method of PL at 20°C and by electron microscopy (including the HREM). PL was excited with a nitrogen laser with $\lambda = 337$ nm and was detected with an FEU-79 photomultiplier. The spectra reported below were normalized to the spectral sensitivity of the setup.

RESULTS

Figure 1 shows the PL spectra of the samples before and after bombardment with He^+ ions with different doses. An intense red–infrared (red–IR) band is evident in the spectrum of the unirradiated sample; this band is typically attributed to quantum-confinement effects in Si nanocrystals [16–19]. The intensity of this band decreases progressively as the dose increases from 3×10^{12} to 3×10^{13} cm^{-2}, with the peak position remaining unchanged. In addition, a weak emission in the short-wavelength region of the spectrum becomes more intense with increasing dose (Fig. 1). For the sake of reference, the SiO_2 layer that did not contain the Si nanocrystals was subjected to ion bombardment with a dose as high as 10^{15} cm^{-2}. The short-wavelength emission was observed in this case as well; however, it featured a significantly lower intensity.

The influence of annealing on PL in the irradiated samples is illustrated in Fig. 2. For the sake of comparison, the PL spectra were measured after each annealing stage in combination with measuring the spectra of the reference (unirradiated) sample (see Fig. 1, curve *1*). The intensity of the red–IR band in the spectrum of this sample was taken as unity, and the other curves in Fig. 2 were normalized to this intensity. The bombardment-induced short-wavelength PL was virtually quenched after annealing at a temperature as low as 600°C. The behavior of the long-wavelength PL related to silicon nanocrystals depended on the He^+-ion dose. After minimal-dose irradiation and annealing at 600°C, the intensity of this band was even higher than that in the unirradiated sample. This band was also comparatively more intense in the sample irradiated with a dose of 10^{13} cm^{-2}. However, for the dose of 3×10^{13} cm^{-2}, an annealing temperature of 600°C was insufficiently high. After an annealing at 800°C, the intensities of the long-wavelength PL were in all cases somewhat higher than that in the unirradiated sample. Furthermore, an increase in the annealing temperature from 600 to 800°C did not bring about any significant increase in the maximal PL intensity. However, annealing at 1000°C resulted in an abrupt increase in the PL inten-

sity, so that the latter exceeded several times the initial PL intensity.

The HREM study showed than an increase in the He^+-ion dose from 10^{13} to 10^{15} cm^{-2} does not eliminate the nanocrystals. We also performed *in situ* electron-microscopy studies of the samples in the case where the He^+-ion dose was successively increased to 10^{16} cm^{-2}. The nanocrystals in the sample irradiated with a dose of 10^{16} cm^{-2} were found to be amorphized (Fig. 3). It is noteworthy that Si inclusions with dimensions of about 10 nm were found accidentally at the surface of some of the samples. The crystallinity of these inclusions was retained, which is inherent in bulk Si irradiated with ions as light as He^+ at room temperature. We also performed *in situ* postimplantation annealings for 30 min at temperatures up to 775°C. These annealing conditions were found to be inadequate to restore the nanocrystals (Fig. 3). In the case of electron irradiation, the degree of preservation of nanocrystals was checked after 2, 20, 40, 60, and 90 min of irradiation. The nanocrystals cease almost completely to exist only after a 90-min irradiation (Fig. 3). It was observed that larger nanocrystals require a longer irradiation for their amorphization.

The result of counting the number of nanocrystals per unit surface area before and after bombardment with He^+ ions and electrons was unexpected. In both cases, the number of nanocrystals increased, rather than decreased, after the low-dose irradiation (Fig. 4). In the case of electron irradiation, for the sake of reliability, the nanocrystals were counted over several surface areas of the sample under investigation (Fig. 4). It can also be seen from the histogram in Fig. 4 that the disappearance of nanocrystals occurs rather drastically after a certain irradiation dose is attained.

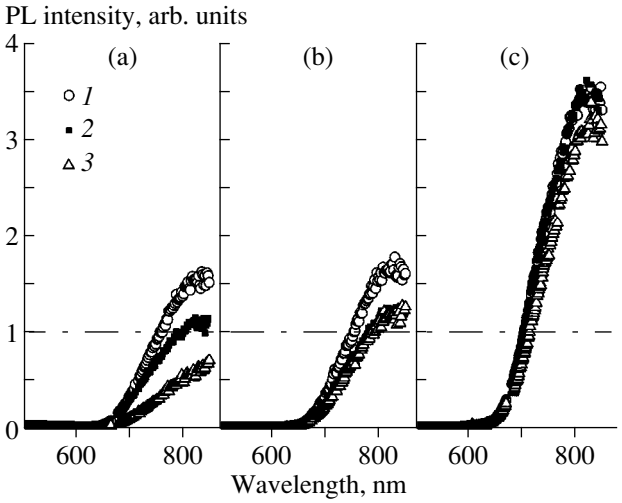


Fig. 2. PL spectra of the samples irradiated with He^+ ions at doses of (*1*) 3×10^{12}, (*2*) 1×10^{13}, and (*3*) 3×10^{13} cm^{-2} and then annealed for 30 min at temperatures of (a) 600, (b) 800, and (c) 1000°C. The horizontal dash-and-dot line indicates the PL intensity of the samples before irradiation.

DISCUSSION OF THE RESULTS

Quenching of the red-IR PL and amorphization are clearly related to the defect production but require distinctly different levels of elastic energy-loss deposition. Quenching of PL is observed for alpha-particle doses on the order of 10^{13} cm^{-2}, which, for a displacement threshold energy of $E_{th} \cong 15$ eV, corresponds to about 6×10^{-5} displacements per atom in the layer under investigation. Since the nanocrystals with dimensions of ~5 nm contain ~5000 atoms, a decrease in the PL intensity occurs if a mere ~1 displacement is introduced into a crystallite. This means that a fixed sink with a high trapping efficiency exists for mobile defects produced by irradiation. The surface of the nanocrystal evidently plays the role of such a sink. The key role of the surface in trapping the point defects is also quite evident from experiments with PL in Si irradiated before and after the formation of the nanocrystalline structure [9].

Another effect, namely, the emergence of new Si nanocrystals after irradiation at 20°C, was evident for the same dose range. It follows from this, first, that, after an initial annealing at 1150°C, the SiO_2 layer contained Si nanoprecipitates, which remained amorphous, apart from Si nanocrystals. Second, such nanoprecipitates may crystallize abruptly (without heating) owing to individual atomic displacements. At least it is elastic losses, rather than ionization, that initiates the crystallization. For electrons with $j \cong 60$ A/cm^2, the ionization level was by four orders of magnitude higher than that for He^+ ions with $j \cong 1$ $\mu A/cm^2$; nevertheless, in both cases, the effect of crystallization was observed for elastic losses that corresponded to unit displacements in the precipitates. We are possibly dealing here with the phenomenon known as impact crystallization [22, 23].

Amorphization by He^+ ions occurs for large doses, such that there are 0.2 displacements for a target atom (~100 displacements per nanocrystal). As is known, silicon is amorphized at 20°C by heavier ions alone, with the defect-formation expenditure amounting to as high as ~1 displacement per atom. Since the boundaries of nanocrystals efficiently trap the mobile defects and, as follows from our observations, the dose for amorphization increases with the crystallite size, the loss of stability of the lattice is caused by the accumulation of point defects at the boundaries up to a certain critical level. The results reported in [24, 25] indicate that Si can be amorphized owing to the accumulation of point defects, provided that there are binding boundaries.

Calculations show that expenditures of elastic losses spent for amorphization are much larger in the case of electrons than in the case of He^+ ions and amount to about 20 displacements per atom. This may be caused by the difference in the displacement rate for electrons

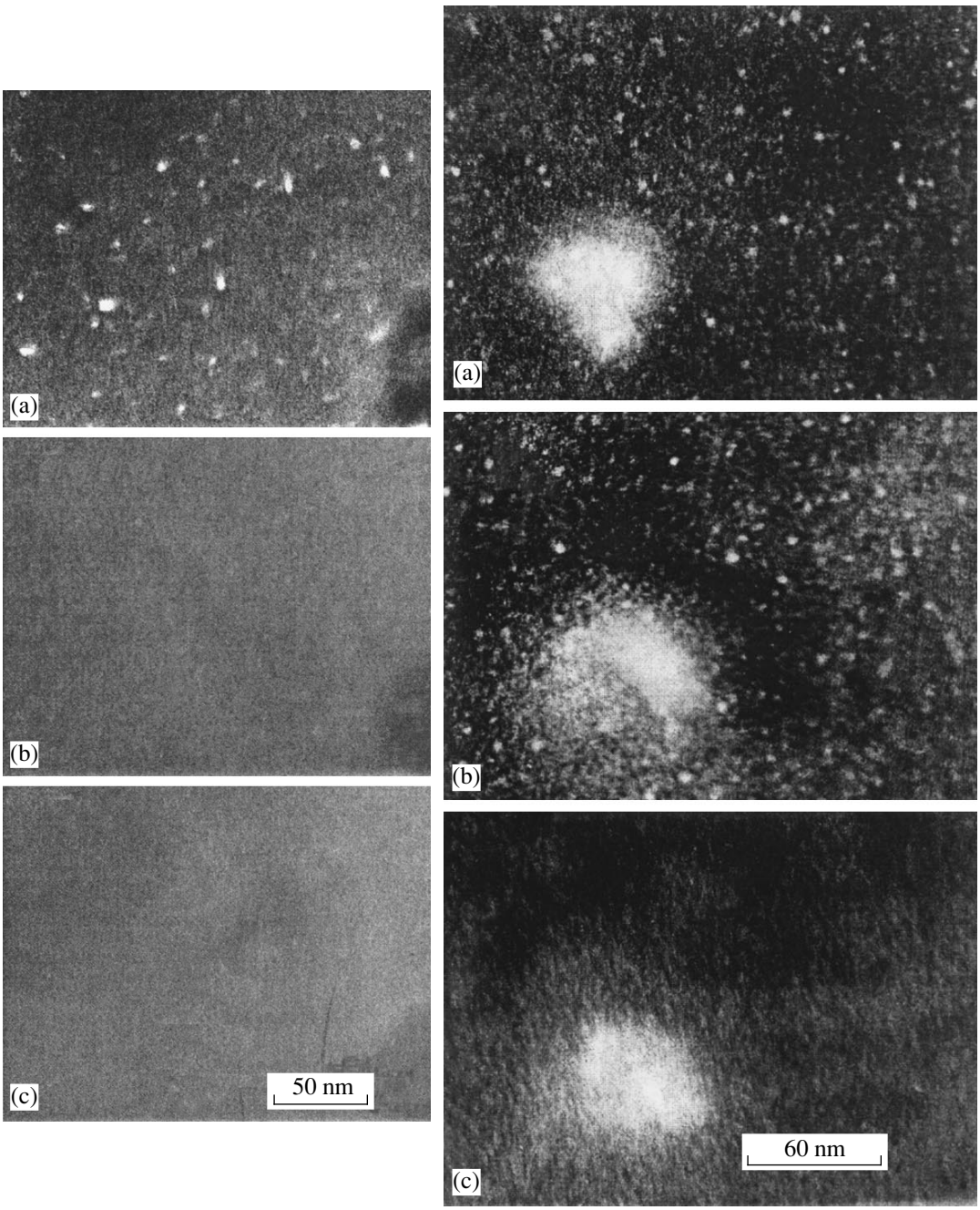


Fig. 3. The dark-field electron-microscopy images of SiO_2 layers with embedded Si nanocrystals (light dots). On the left: (a) the sample before irradiation, (b) after irradiation with He^+ ions at the dose of 1×10^{16} cm^{-2}, and (c) the same sample after subsequent annealing for 30 min at 775°C. On the right: the sample irradiated with electrons for (a) 2, (b) 40, and (c) 90 min.

and alpha particles and by the difference in the energy transferred to the target atoms. The generation rate for displacements by electrons with a current density of ~60 A/cm^2 is higher by two orders of magnitude than that by alpha particles for the current densities ~1 μA/cm^2, and, in this regard, it would be easier for electrons to amorphize Si. The homogeneous formation of stable damage within a single displacement spike may be ignored for light ions like He^+. Their elastic energy losses amount to ~10 eV/nm, and no more than three Frenkel pairs can be formed within a crystallite ~5 nm in size. Therefore, the difference in the energy spent by the particles on amorphization is primarily caused by differences in the energy transferred to the target atoms. For He^+ ions, the mean value of this energy is ~100 eV, and the displaced atoms may be found to be at a distance of several lattice constants from the lattice site. By contrast, electrons with an

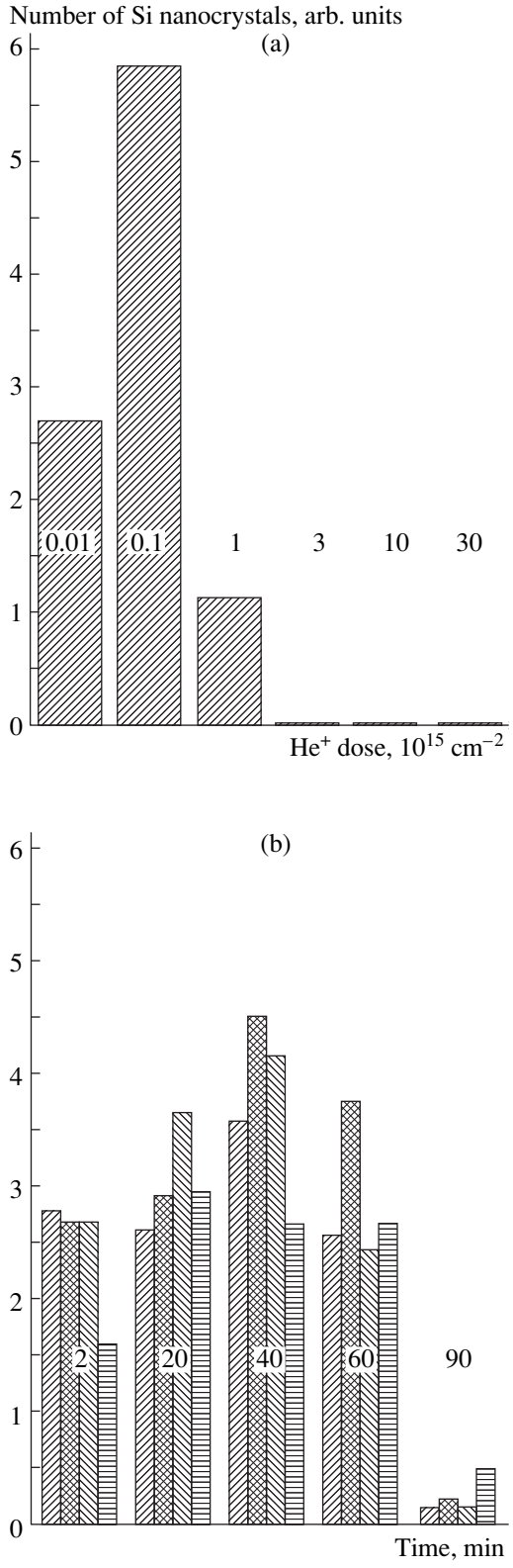


Fig. 4. Histograms illustrating variations in the number of Si nanocrystals as (a) the dose of He^+ ions increases and (b) the duration of electron irradiation increases (in the latter case, the results of counting the nanocrystals over the four different surface areas are shown).

energy of 400 keV transfer a mere 20 eV to the Si atoms and produce close Frenkel pairs. For the latter, recombination is much more probable than escape of the components to the surface.

In order to anneal out the defects in nanocrystallites, temperatures below 800°C are sufficient; at the same time, crystallization of amorphous nanoprecipitates requires temperatures between 800 and 1000°C. This range corresponds to temperatures at which the nanocrystals are formed in SiO_2 layers implanted with Si ions [16–21]. The data we obtained indicate that high annealing temperatures are required not only for diffusion-limited growth of precipitates but also for crystallization itself. Such temperatures are appreciably higher than those corresponding to the onset of crystallization of conventional amorphous Si (~500–600°C). The mechanism for crystallization of nanoprecipitates needs further investigation. Possibly, this mechanism is somehow related to the melting temperature that is lowered for nanocrystals owing to an increase in the role of the surface energy.

It would also be reasonable to relate an enhancement of PL in the irradiated samples after an annealing at 1000°C as observed in this study to an increase in the number of nanocrystallites. This enhancement is probably a result of the summation of intensities of the initial and new emitters. New nanocrystals might be formed not only owing to crystallization under the effect of irradiation at 20°C. The emergence of short-wavelength PL as a result of irradiation (Fig. 1) is indicative of the radiation-stimulated segregation of Si from SiO_2 with the formation of Si–Si pairs (blue PL) and more intricate luminescent complexes [26, 27]. It can be seen from Fig. 1 that, in the layers enriched in Si, the intensity of short-wavelength PL increases much faster and exhibits a larger number of spectral features. Thus, in addition to the direct irradiation-induced modification of silicon precipitates, these precipitates obviously interact with the displaced Si atoms. As a result, the nonocrystalline silicon nanoprecipitates present in the SiO_2 layer could crystallize following repeat annealing at 1000°C.

CONCLUSION

In the initial stages of irradiation of Si nanocrystals in an SiO_2 layer with light particles, the quenching of PL related to quantum-confinement effects occurs; this quenching is not related to amorphization. PL is quenched after irradiation with ion doses corresponding to a mere ~1 displacement per nanocrystal. This quenching is apparently due to the introduction of nonradiative-recombination centers. In this case, the nanocrystals are not amorphized; rather, their numbers increase as a result of irradiation. Crystallization of silicon nanoprecipitates in the SiO_2 layers irradiated at 20°C occurs in the same dose range as the PL quenching, i.e., at a level of ~1 displacement per nanocrystal. This phenomenon is not related to ionization, and its

mechanism is possibly similar to impact crystallization. In contrast with the bulk material, the Si nanocrystals may be amorphized at room temperature as a result of irradiation with electrons and light ions. This occurs for alpha-particle doses that correspond to about 0.2 displacements per atom in nanocrystals ~5 nm in size (instead of ~1 displacement per atom needed for the amorphization of bulk Si by ions with medium and large mass). The PL quenching and amorphization occur owing to the interaction of generated mobile defects with the surface of nanocrystals. Since electrons mainly produce rapidly recombining Frenkel pairs, the energy expenditure for amorphization is larger for electrons than for alpha particles. Radiation defects in Si nanocrystals are annealed out at temperatures below 800°C; however, higher temperatures are required to recover the crystallinity of amorphized nanoprecipitates. After an annealing of irradiated SiO_2 layers at 1000°C, the PL intensity for nanocrystals is several times higher than that after the nanocrystals were initially formed at 1150°C. This phenomenon is explained by the fact that the PL intensities for initially formed nanocrystals and for those produced by irradiation are summed.

ACKNOWLEDGMENTS

We are grateful to S.V. Shemyakin for his help in conducting this study.

REFERENCES

1. L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
2. L. Wang, C. Lin, P. Liu, *et al.*, Mater. Res. Soc. Symp. Proc. **316**, 445 (1994).
3. X.-M. Bao, H.-Q. Yang, and F. Yan, J. Appl. Phys. **79**, 1320 (1996).
4. X.-M. Bao and H.-Q. Yang, Appl. Phys. Lett. **63**, 2246 (1993).
5. J. S. Barbour, D. Dimos, T. R. Guillinger, *et al.*, Appl. Phys. Lett. **59**, 2088 (1991).
6. B. R. Mehta, M. K. Sahay, L. K. Malhotra, *et al.*, Thin Solid Films **289**, 95 (1996).
7. S. Tanaka, H. Koyama, and N. Koshida, Appl. Phys. Lett. **73**, 2334 (1998).
8. R. R. Kunz, P. M. Nitishin, H. R. Clark, *et al.*, Appl. Phys. Lett. **67**, 1766 (1995).
9. E. V. Astrova, V. V. Emtsev, A. A. Lebedev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1301 (1995) [Semiconductors **29**, 674 (1995)].
10. A. A. Lebedev, A. M. Ivanov, A. D. Remenyuk, and Yu. V. Rud', Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 188 (1996) [Semiconductors **30**, 107 (1996)].
11. J.-L. Maurice, A. Riviere, A. Alapini, and C. Levy-Clement, Appl. Phys. Lett. **66**, 1665 (1995).
12. Y. M. Huang, Appl. Phys. Lett. **71**, 3850 (1997).
13. J. S. Fu, J. C. Mao, E. Wu, *et al.*, Appl. Phys. Lett. **63**, 1830 (1993).
14. T. M. Bhave, S. V. Bhoraskar, P. Singh, and V. N. Bhoraskar, Nucl. Instrum. Methods Phys. Res. B **132**, 409 (1997).
15. V. V. Ushakov, V. A. Dravin, N. N. Mel'nik, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **31**, 1126 (1997) [Semiconductors **31**, 966 (1997)].
16. T. Shimizu-Iwayama, K. Fujita, S. Nakao, *et al.*, J. Appl. Phys. **75**, 7779 (1994).
17. P. Mutti, G. Ghislotti, S. Bertoni, *et al.*, Appl. Phys. Lett. **66**, 851 (1995).
18. G. A. Kachurin, I. E. Tyschenko, K. S. Zhuravlev, *et al.*, Nucl. Instrum. Methods Phys. Res. B **122**, 571 (1997).
19. G. A. Kachurin, K. S. Zhuravlev, N. A. Pazdnikov, *et al.*, Nucl. Instrum. Methods Phys. Res. B **127/128**, 583 (1997).
20. G. A. Kachurin, I. E. Tyschenko, V. Skorupa, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **31**, 730 (1997) [Semiconductors **31**, 626 (1997)].
21. G. A. Kachurin, A. F. Leĭer, K. S. Zhuravlev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 1371 (1998) [Semiconductors **32**, 1222 (1998)].
22. T. Takamori, R. Messier, and R. Roy, Appl. Phys. Lett. **20**, 201 (1972).
23. A. Mineo, A. Matsuda, T. Kurosu, and M. Kikuchi, Solid State Commun. **13**, 329 (1973).
24. S. I. Romanov and L. S. Smirnov, Fiz. Tekh. Poluprovodn. (Leningrad) **10**, 876 (1976) [Sov. Phys. Semicond. **10**, 519 (1976)].
25. H. A. Atwater and W. L. Brown, Appl. Phys. Lett. **56**, 30 (1990).
26. L.-S. Liao, X.-M. Bao, N.-S. Li, *et al.*, J. Lumin. **68**, 199 (1996).
27. G. A. Kachurin, L. Rebole, V. Skorupa, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 439 (1998) [Semiconductors **32**, 392 (1998)].

Translated by A. Spitsyn

Semiconductors, Vol. 34, No. 8, 2000, pp. 971–974. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 1010–1013.

PHYSICS OF SEMICONDUCTOR DEVICES

High-Frequency Properties of Avalanche Multiplication of Photocarriers in Structures with Negative Feedback

T. M. Burbaev*, V. A. Kurbatov*, N. E. Kurochkin, and V. A. Kholodnov****

** Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia*

*** NPO Orion State Research Center, Theoretical Department, Moscow, 111123 Russia*

Submitted February 14, 2000; accepted for publication February 24, 2000

Abstract—The high-frequency characteristics of photosensitive avalanche Si–SiC structures were studied. It is shown that their high-speed operation is substantially superior to that of silicon avalanche photodiodes. A theoretical analysis of the high-frequency properties of avalanche photodiodes is carried out and analytical expressions for the gain–bandwidth product are obtained. It is shown that this product is not a universal parameter for a metal–insulator–semiconductor structure with a negative feedback, since, for high amplification factors, the effective value of the relation of the impact-ionization coefficients for different types of charge carriers in such structures turns out to be significantly different from that in avalanche photodiodes. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is traditionally believed that the best characteristics of avalanche amplification are realized in avalanche photodiodes (APD) made of silicon. The reason for this is a substantial difference of the impact ionization coefficients by electrons and holes in Si, which weakens the avalanche-process positive feedback, which induces a spread of the single-electron amplification amplitudes. This spread manifests itself naturally in the form of excess noise whose behavior in avalanche photodiodes (APD) obeys the known McIntyre theory [1] closely.

However, a number of studies [2–5] show that the Si-based metal–insulator–semiconductor (MIS) structures exhibit far better characteristics than silicon APD; until now, most attention has been drawn to the amplification and noise characteristics of such structures. The main results of this paper concern the studies of temporal characteristics, but we also present the main results of the noise studies, since, as will be shown below, they are correlated to a large extent.

NOISE IN AVALANCHE STRUCTURES OF THE MIS TYPE

One can attribute to MIS-type structures both silicon–insulator structures proper and silicon–wide-bandgap semiconductor heterodiodes, in which the effective depletion of the silicon region under a constant electrical bias is provided due to the high concentration of deep centers in a wide-bandgap semiconductor [2]. In this case, the outflow of carriers from the potential well occurs across the layer of a wide-bandgap semiconductor. In the MI(p–n)/MI(n–p)-type structures [3], with light doping of the thin semiconductor layer adjacent to the insulator, the outflow of carriers from the interface occurs over the plane of a structure.

The values of the noise factor (F) measured for various structures are shown in Fig. 1 in relation to the multiplication factor (M) and also for Si–SiC structures among them, the temporal characteristics of which will be discussed afterwards. The silicon–insulator (Si–SiO_2) structures do not provide a flow of constant current; therefore, they were tested in the pulsed mode of the electrical bias. The calculated dependences of the noise factor according to the McIntyre theory are given in Fig. 1 for the ratio of impact ionization coefficients by holes and electrons of $k = 0.001$. It can be seen that the McIntyre theory does not describe the behavior of the MIS structure noise even for such a low value of k. One should note that the values $k < 0.01$ even for silicon APDs are practically unrealistic. The results of calculation for a MIS-type structure are shown too; these calculations were performed by us previously [6].

A qualitative explanation of the noise behavior observed consists in the following. As distinct from an APD, where the carrier multiplication as a result of impact ionization represents an avalanche (i.e., self-damped) process in a constant electric field, in the case of a MIS structure, the impact ionization begins in a field substantially exceeding the breakdown field similar to a Geiger mode of an APD. As distinct from the latter, the breakdown quenching in our case occurs locally for each "one-electron" process owing to the

1063-7826/00/3408-0971$20.00 © 2000 MAIK "Nauka/Interperiodica"

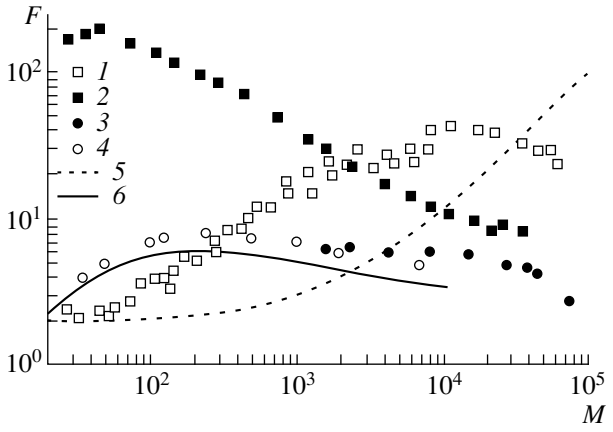


Fig. 1. Experimental values of the noise factor and calculated dependences $F(M)$ for a series of MIS-type structures: (*1*) Si–TiO_x [5]; (*2*) Si–SiC; (*3*) Si–SiO_2 [4]; (*4*) Si(*p*–*n*)–SiO_2 [3]; (*5*) calculation for the case of a *p*–*n* junction [1], $k = 0.001$; and (*6*) calculation for a MIS-type structure [6].

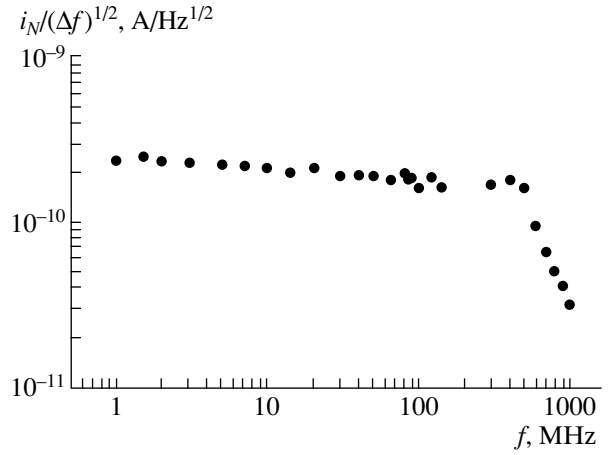


Fig. 2. Frequency dependence of the noise current spectral density for a Si–SiC heterojunction diode; the noise current was generated by radiation with $\lambda = 0.7$ μm.

internal mechanism of negative feedback—screening of the electric field in the space-charge region (SCR) due to accumulation of the charge at the interface or in the bulk of a wide-bandgap semiconductor.

A decrease in the electric field in the course of the avalanche process results in a reduction of the probability of each subsequent generation of electrons, and, in this case, the effective value k should decrease; the higher the starting field is (i.e., M), the larger this decrease is. In the limit, one can conceive that the first generation of electrons is sufficient for field screening to such an extent that a further multiplication in SCR is practically impossible. In this case, the noise factor tends to 2 (since we can assume that the primary holes do not have time enough to produce some noticeable generation, which is similar to the case of $k = 0$).

SPEED OF RESPONSE OF AVALANCHE MIS STRUCTURES

Previously [5], we already noted that the temporal characteristics of MIS structures, as well as noise, differ appreciably from APD characteristics. Here, we present more detailed results and their theoretical grounds.

We measured the amplitude–frequency characteristics of the heterojunction avalanche Si–SiC photodiode noise in the frequency range of 1–1000 MHz. The dependence of the noise-current spectral density $i_N/(\Delta f)^{1/2}$ on the frequency f of a photodiode operating in the multiplication mode with $M \approx 10000$ under continuous illumination with radiation with a wavelength $\lambda = 0.7$ μm is shown in Fig. 2. The experimental dependence obtained points to the fact that the gain frequency band of the photodiode is no less than 500–600 MHz (with account taken of the *RC*-limitation of the electric circuit). The product of the bandwidth by the multiplication coefficient in this case exceeds 5000 GHz, which is by more than an order of magnitude higher than the value obtained for silicon APD (320 GHz [7]). Similar results were also obtained by us by studying a Si–SiO_2 structure in the pulsed mode. We note that in the frequency range employed, a slight decay begins at 1 MHz. The measurements for a series of samples showed that in most cases the amplitude–frequency characteristic exhibits an even larger increase at low frequencies. This points to the presence of long-term "tails" in one-electron pulses.

GAIN–BANDWIDTH PRODUCT

The avalanche process in a conventional abrupt *p*–*n* junction occurs under conditions of positive feedback, which is realized owing to ionization by secondary carriers (of another sign than the carrier which initiates the avalanche).

The electron initiating the avalanche generates M_0 electrons and, correspondingly, $M_0 - 1$ holes. We may assume that, for small k, practically all the holes are generated at the metallurgical boundary. With a certain probability (ρ), the holes produce an electron–hole pair. Due to the smallness of ρ, the secondary electron is formed with a high probability at the SCR boundary, and with the probability $(M_0 - 1)\rho$ the process of avalanche multiplication repeats. Since all the electrons have similar conditions for multiplication, we can pass from temporal averaging to ensemble averaging; i.e., we assume that the current diminishes from its maximum value by e times when the probability of the appearance of the nth-electron generation is equal to $1/e$:

$$[(M_0 - 1)\rho]^n = 1/e.$$

The time of avalanche development is given by

$$T = n\tau,$$

where τ is the doubled time of transit across the SCR.

Under steady-state illumination, the total number of electrons and holes generated in the avalanche process (i.e., the gain M) is

$$M = M_0 + M_0 - 1 + M_0(M_0 - 1)\rho$$

$$+ (M_0 - 1)(M_0 - 1)\rho + \ldots = \frac{2M_0 - 1}{1 - (M_0 - 1)\rho}.$$

Thus, the gain-bandwidth product is given by

$$\Delta f M = \frac{2M_0 - 1}{1 - (M_0 - 1)\rho}\frac{1}{2\pi n\tau}$$

$$= (2M_0 - 1)\frac{-\ln[(M_0 - 1)\rho]}{1 - (M_0 - 1)\rho}\frac{1}{4\pi\tau}.$$

For $M \gg 1$, $(M_0 - 1)\rho$ tends to 1, and, respectively, the gain-bandwidth product tends to $(2M_0 - 1)/4\pi\tau$. Numerical calculations show that, for $M > 10$, the value of M_0 practically does not change (for constant k). One can also show that in this case $M_0 \approx 1/k$ also for $k \ll 1$. This means that, for Si, the gain-bandwidth product for large multiplication coefficients is defined by a simple relation

$$\Delta f M \approx 1/2\pi k\tau.$$

Notwithstanding the fact that this expression is rather approximate, it accounts for the behavior of the value ΔfM as a characteristic parameter for a certain type of structure. The accuracy of the results obtained can be judged from comparison with the paper by Emmons [8], where computational methods were used to show that the avalanche development time changes from $(1/3)M\tau$ for small M to $2kM\tau$ for $M \gg 1$. The calculation according to the above formula yields the corresponding change from $(1/e)M\tau$ to $kM\tau$. We remind the reader that these results are obtained assuming that the value of k is constant and independent of the field.

As was noted above, for MIS structures, the avalanche multiplication of carriers occurs in the field that decreases during the avalanche process, which results in a change in the value of k during the process. For this reason, in the presence of the negative feedback mechanism that reduces the probability of ionization by each subsequent hole generation, the product ΔfM is no longer a universal index for any (at least, for large) M and changes with the value of M.

One can assess the effective value of k at the operating point (k_{eff}) from the relation between the multiplication coefficient and the noise factor. Without claiming to perform a rigorous treatment, we assume that, at each operating point, we can use the McIntyre formula to determine $k_{\text{eff}} = k(M)$. Then, for $F \gg 1$, we have $k_{\text{eff}} = F(M)/M$, which yields

$$\Delta f = \frac{1}{2\pi k_{\text{eff}}\tau M} = \frac{1}{2\pi F(M)\tau}.$$

In such a form, the expression for Δf is also applicable to a structure with negative feedback. Substituting the values of parameters for the heterostructure under study ($M = 10000$, $F = 10$, the thickness of SCR $w = 2$ μm, and the carrier saturation velocity $v_s = 10^7$ cm/s) into this expression, we obtain

$$\Delta f = \frac{1}{2\pi F(M)(2w/v_s)} \approx 500 \text{ MHz},$$

which is consistent with the experimental result.

Taking into account the approximate character of the calculation, it is reasonable to expect a rather significant difference between the calculated and experimental values; however, taking into account the *RC* limitation, one can consider that the experimental value is at least no smaller than the calculated one.

The effective ratio of the ionization coefficients by holes and electrons is in this case $k_{\text{eff}} = F(M)/M = 0.001$, and the number of generations of secondary carriers is given by

$$n = \frac{T}{\tau} = \frac{1}{2\pi\Delta f(2w/v_s)} \approx 8.$$

In the limiting case, for $n = 1$, the response time to a single pulse will be equal to the time of the hole drift from the heteroboundary to the SCR boundary. Here, it is of importance to note that, as the calculations show for the case of a constant doping level (linear coordinate dependence of the field), about 0.1–0.3 of the SCR length affects the high-speed performance of the abrupt p–n-junction for large M ($M > 10$) (depending on the approximation employed for the impact ionization coefficient dependence on the intensity of the electric field).

CONCLUSION

In heterostructures similar to MIS structures, an appreciable decrease in the effective ratio of the probabilities of impact ionization by electrons and holes for an avalanche process in silicon is achieved, which results in a radical improvement in both the threshold and temporal characteristics of the avalanche photodetectors.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Science and Technology within the framework of the Program "Promising Devices and Technologies of Microelectronics and Nanoelectronics" (project no. 02.04.1.1.16. É.1), the State Program for the Support of Scientific Schools (project no. 96-15-96341), and the Russian

Foundation for Basic Research (project no. 99-02-16675).

REFERENCES

1. R. McIntyre, IEEE Trans. Electron Devices **ED-13** (1), 164 (1966).
2. T. M. Burbaev and V. A. Kurbatov, Kratk. Soobshch. Fiz., No. 11/12, 38 (1994).
3. N. Bacchetta, D. Bisello, Z. Sadygov, *et al.*, Nucl. Instrum. Methods Phys. Res. A **387** (1/2), 225 (1997).
4. T. M. Burbaev, V. V. Kravchenko, V. A. Kurbatov, and V. É. Shubin, Kratk. Soobshch. Fiz., No. 4, 19 (1990).
5. A. P. Boltaev, T. M. Burbaev, G. A. Kalyuzhnaya, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29** (7), 1220 (1995) [Semiconductors **29**, 630 (1995)].
6. N. E. Kurochkin and V. A. Kholodnov, Pis'ma Zh. Tekh. Fiz. **25** (9), 70 (1999) [Tech. Phys. Lett. **25**, 369 (1999)].
7. T. Kaneda *et al.*, J. Appl. Phys. **47**, 4960 (1976).
8. R. B. Emmons, J. Appl. Phys. **38**, 3705 (1967).

Translated by T. Galkina

Semiconductors, Vol. 34, No. 8, 2000, pp. 975–980. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 1014–1019.

PHYSICS OF SEMICONDUCTOR DEVICES

Parameters of Metal One-Electron Transistors Based on Various Materials

I. I. Abramov* and E. G. Novik

Belarussian State University of Information Science and Radio Engineering, ul. Brovki 17, Minsk, 220072 Belarus

** e-mail: device@micro.rei.minsk.by*

Submitted February 21, 2000; accepted for publication February 24, 2000

Abstract—Limiting parameters (operating temperature and cutoff frequency) and current–voltage characteristics of one-electron transistors based on various metal compounds ($Al/AlO_x/Al$, $Al/SiO_2/Al$, $Au/Al_2O_3/Au$, $Nb/Al_2O_3/Nb$, $Ti/TiO_x/Ti$, $Cr/Cr_2O_3/Cr$, and $Nb/NbO_x/Nb$) were theoretically studied. Practical recommendations for the choice of materials and structure sizes were formulated. The characteristics were calculated using a SET-NANODEV software package based on the effect of one-electron tunneling and developed for structure simulation according to a technique for estimating the limiting parameters and a two-dimensional numerical model of the metal one-electron transistor. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The present state of the art of solid-state technology is adequate to produce structures with nanometer regions. Electron transport in such structures can substantially depend on the effect of one-electron tunneling [1]. Currently, one-electron structures are actively studied experimentally. Operating devices based on the effect of one-electron tunneling have already been described [1–3]. One of the most promising is the metal one-electron transistor [4]. Along with experimental studies of the effect of one-electron tunneling, a demand arises for models adequately describing processes in such structures. Furthermore, device parameters should be analyzed depending on design, technological, and electrical parameters. Such an investigation would allow the determination of necessary sizes, the selection of appropriate materials when developing one-electron devices, and the prediction of their output characteristics.

This study is aimed at analyzing the basic parameters of metal one-electron transistors (operating temperature, cutoff frequency, and current–voltage characteristic) for various materials and at revealing the most preferable systems.

MODELS

To study the characteristics of one-electron transistors, we used a technique developed for estimating limiting parameters and a two-dimensional (2D) numerical model, described elsewhere [5–7] in detail. Therefore, below we only briefly dwell on the above technique and model.

The technique allows calculation of a number of one-electron transistor parameters (operating temperature, cutoff frequency, optimum operating voltage, capacitance and area of tunnel junctions, and total structure capacitance) depending on material parameters and device configuration. Applicability of the technique was confirmed by comparing the results of simulation and experimental data [5].

The developed 2D numerical model of the metal one-electron transistor is intended for calculating the current–voltage (I–V) characteristics of a device with a block diagram typical of one-electron transistors [4]. As distinct from known models based on the semiclassical approximation [1] and the Monte Carlo method [8], initial data in the proposed model are geometrical, electrical (voltages at contacts), and external (temperature) parameters. The developed model is based on a numerical solution to the 2D Poisson equation (the impact of a magnetic field is neglected)

$$\nabla\varepsilon\nabla\varphi(x, y) = -q_M, \qquad (1)$$

where ε is the dielectric constant, φ is the electrostatic potential, and q_M is the volume density of the electric charge. The model also assumes the electroneutrality of plasma of bound charges at an "island," except for free carriers. Charges in an insulator are neglected. Thus, q_M is defined only by the averaged density of excess charges directly at the island.

The model includes a transition from the continuous Poisson equation (1) to its discrete approximation for the 2D case. The Dirichlet and Neuman boundary conditions were chosen as the boundary ones at contacts and free boundaries, respectively. The proposed finite-difference approximation of the Poisson equation, based on the Tikhonov–Samarskiĭ approach, and a number of assumptions were used to derive a set of linear equations, solved by the matrix sweep method. The solution yields a potential distribution in the structure depending on its geometrical, electrical, and external

1063-7826/00/3408-0975$20.00

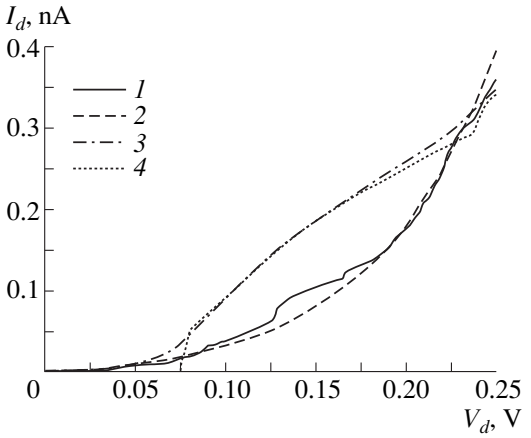


Fig. 1. Current–voltage characteristics of the $Nb/NbO_x/Nb$-based device: (*1*) experimental data and the results of calculations based on (*2*) the 2D numerical model, (*3*) the MOSES code, and (*4*) the semiclassical model.

parameters, as well as on the number of excess charges at the island.

To calculate the device *I–V* characteristics, we use the master equation for one-electron structures [1], written directly for currents through tunnel barriers rather than for tunneling rates. To calculate partial currents, we used the tunnel-junction voltages (determined taking into account the electrostatic potential distribution) and resistances calculated according to [9].

The developed model allows better agreement with experimental data compared to the most adequate of the known models [1, 8] for actual transistors, in particular, those based on $Ti/TiO_x/Ti$ and $Nb/NbO_x/Nb$ [7]. As an example, Fig. 1 displays the current–voltage characteristic $I_d(V_d)$ of the $Nb/NbO_x/Nb$-based device (curve *1* corresponds to the experimental data [10]; curves *2–4* represent the results of calculations based on the developed model, known MOSES code [8], and semiclassical model [1], respectively). It is evident from Fig. 1 that the curve constructed using the 2D numerical simulation [2] conforms well to the experimental data [10]. The error was estimated at 10%. Hence, it is expedient to employ the model to predict electrical parameters of one-electron transistors, since it yields a more complete relation between output parameters and device design (and material parameters) compared to known models.

The developed technique and model were implemented in the SET-NANODEV software package for simulating one-electron structures [7, 11].

LIMITING PARAMETERS OF ONE-ELECTRON TRANSISTORS

We studied limiting parameters (operating temperature T_{op} and cutoff frequency F_t) of one-electron transistors with a planar configuration, based on various metal compounds ($Al/AlO_x/Al$, $Al/SiO_2/Al$, $Au/Al_2O_3/Au$, $Nb/Al_2O_3/Nb$, $Ti/TiO_x/Ti$, $Cr/Cr_2O_3/Cr$, and $Nb/NbO_x/Nb$) depending on electrical parameters of the structure (insulator dielectric constant ε_d and potential barrier height ϕ for tunneling). Relatively high T_{op} and F_t were established to be inherent in a structure with a small insulator dielectric constant and (or) with a low potential barrier ϕ of tunnel junctions; the latter factor affects the parameters more strongly than the former one for a fixed width of source d_s (or drain) tunnel junctions. Therefore, we should rely on the second factor when choosing materials for a device to be designed. According to the first factor, the following systems are preferable (in order of deteriorating parameters): $Al/SiO_2/Al$ ($\varepsilon_d = 4$); $Nb/Al_2O_3/Nb$, $Au/Al_2O_3/Au$, and $Al/AlO_x/Al$ ($\varepsilon_d = 8$–10); $Cr/Cr_2O_3/Cr$ ($\varepsilon_d = 12$); $Ti/TiO_x/Ti$ ($\varepsilon_d = 24$); and $Nb/NbO_x/Nb$ ($\varepsilon_d = 64$). According to the second factor, we have the following sequence: $Cr/Cr_2O_3/Cr$ ($\phi = 0.06$ eV), $Nb/NbO_x/Nb$ ($\phi = 0.133$ eV), $Ti/TiO_x/Ti$ ($\phi = 0.285$ eV), $Nb/Al_2O_3/Nb$ ($\phi = 0.74$ eV), $Al/AlO_x/Al$ ($\phi = 1.6$ eV), $Au/Al_2O_3/Au$ ($\phi = 1.65$ eV), and $Al/SiO_2/Al$ ($\phi = 3.2$ eV). According to the operating-temperature and cutoff-frequency parameters calculated for devices with a planar configuration, the above systems can be arranged in order of decreasing parameters as $Cr/Cr_2O_3/Cr$, $Nb/NbO_x/Nb$, $Ti/TiO_x/Ti$, $Nb/Al_2O_3/Nb$, $Al/AlO_x/Al$, $Au/Al_2O_3/Au$, and $Al/SiO_2/Al$. We illustrate the above trends by specific examples.

Figure 2 displays dependences of the operating temperature T_{op} on the source (drain) tunnel-junction width d_s for one-electron transistors with a planar configuration, based on several metal compounds ($Al/SiO_2/Al$, $Au/Al_2O_3/Au$, $Nb/Al_2O_3/Nb$, and $Ti/TiO_x/Ti$). The largest values T_{op} are inherent in a transistor based on the $Ti/TiO_x/Ti$ structure whose relative insulator dielectric constant is larger, while the junction potential barrier height is much lower than in the other three structures. Thus, the barrier height ϕ affects the device operating temperature more strongly than the dielectric constant. Variation of ϕ can lead to a large change in T_{op} for an unchanged junction width d_s. For example, for d_s = 1 nm, the difference in T_{op} is about 85 times for transistors based on $Au/Al_2O_3/Au$ and $Nb/Al_2O_3/Nb$ (with identical ε_d) and about 10^4 times for those based on $Al/SiO_2/Al$ or $Ti/TiO_x/Ti$. Thus, the most optimal operating temperature is inherent in devices with a lower potential barrier of tunnel junctions.

Figure 3 shows dependences of the cutoff frequency F_t on the tunnel junction width d_s for one-electron transistors with a planar configuration, based on tunnel junctions $Cr/Cr_2O_3/Cr$, $Nb/NbO_x/Nb$, and $Ti/TiO_x/Ti$. The higher boundary frequency for fixed d_s is inherent in a device based on $Cr/Cr_2O_3/Cr$ due to a very low ϕ. However, when choosing devices with low potential barriers, one should take into account that parameters are limited from above by the smallest feasible structure size, for instance, by the tunnel junction width,

since the operating voltage should not exceed the breakdown voltage. This limitation is most pronounced for the $Cr/Cr_2O_3/Cr$-based transistor, since it is characterized by the lowest potential barrier among the considered structures (see Fig. 3).

The impact of electrical parameters was also studied for a transistor with a spherical configuration (with source, drain, and gate surrounding an island). In the case of fixed widths of source (drain) tunnel junctions, the same trends in the impact of electrical parameters on studied device characteristics are valid as for the transistor with a planar configuration. For instance, the operating temperature and cutoff frequency are greater at lower junction potential barriers and smaller dielectric constants if the barrier height is fixed. The largest operating parameters T_{op} and F_t are inherent in devices based on tunnel junctions $Cr/Cr_2O_3/Cr$, $Nb/NbO_x/Nb$, $Ti/TiO_x/Ti$, and $Nb/Al_2O_3/Nb$ due to a fairly low potential barrier. These systems are preferable to $Al/AlO_x/Al$, $Au/Al_2O_3/Au$, and $Al/SiO_2/Al$. The latter system is characterized by the smallest values of T_{op} and F_t.

Using the example of the operating temperature calculated in relation to the conducting island radius r_0 for transistors with a spherical configuration, based on various metal compounds, we studied the impact of electric parameters for fixed radius r_0. As an example, the dependence of T_{op} on r_0 for the $Al/AlO_x/Al$-based transistor is plotted in Fig. 4. For fixed r_0, the dielectric constant variation strongly affects the device operating temperature. For $r_0 = 2$ nm and $\phi = 1.6$ eV, the operating temperature in the case of $\varepsilon_d = 8$ is approximately 1.3 times higher than for $\varepsilon_d = 10$. For $r_0 = 2$ nm and $\varepsilon_d = 8$, the value of T_{op} for $\phi = 1.6$ eV is 1.2 times larger than for $\phi = 2.5$ eV. Thus, both electrical parameters affect the transistor operating temperature for a fixed radius of the conducting island.

Dependences $T_{op}(r_0)$ in Fig. 5 represent an example of the impact of electrical parameters on a characteristic of the device with a spherical configuration, based on various metal compounds for a fixed radius of the conducting island. The highest operating temperature is inherent in the $Cr/Cr_2O_3/Cr$-based transistor with the lowest value of ϕ (0.06 eV) and relatively small ε_d (12). The lowest value of T_{op} corresponds to the $Nb/NbO_x/Nb$-based transistor with relatively low ϕ (0.133 eV) and the largest value of ε_d (64) among considered systems. For $r_0 = 10$ nm, the operating temperature for the $Cr/Cr_2O_3/Cr$-based transistor is higher approximately by an order of magnitude than for the $Nb/NbO_x/Nb$-based transistor. The fairly high value of T_{op} is typical of the transistor based on $Al/SiO_2/Al$ tunnel junctions with the smallest ε_d. Comparison of the calculated data in Fig. 5 allows the following conclusion: for a fixed radius of the conducting island, the parameters are strongly affected by the potential barrier height and the insulator dielectric constant; as these decrease, the values of T_{op} and F_t increase. The studied

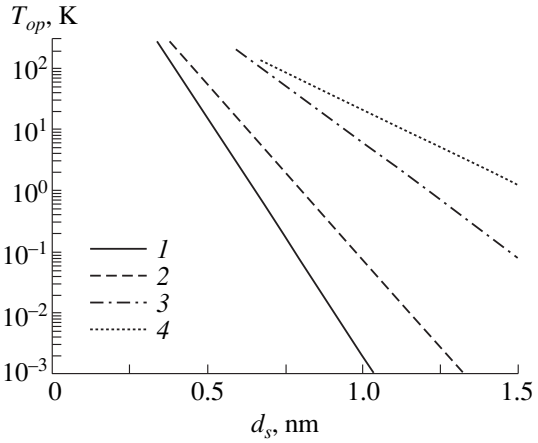


Fig. 2. Dependence of the operating temperature on the width of the source (drain) tunnel junctions for transistors that have a planar configuration and are based on (*1*) $Al/SiO_2/Al$, (*2*) $Au/Al_2O_3/Au$, (*3*) $Nb/Al_2O_3/Nb$, and (*4*) $Ti/TiO_x/Ti$.

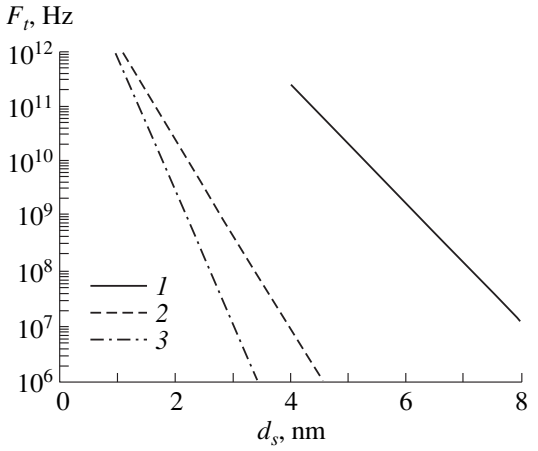


Fig. 3. Dependence of the cutoff frequency on the width of the source (drain) in tunnel junctions for transistors with a planar configuration, based on (*1*) $Cr/Cr_2O_3/Cr$, (*2*) $Nb/NbO_x/Nb$, and (*3*) $Ti/TiO_x/Ti$.

systems can be arranged in the following order of decreasing T_{op} and F_t: $Cr/Cr_2O_3/Cr$, $Al/SiO_2/Al$, $Nb/Al_2O_3/Nb$, $Au/Al_2O_3/Au$ ($Al/AlO_x/Al$), $Ti/TiO_x/Ti$, and $Nb/NbO_x/Nb$. However, the upper limitation related to the smallest feasible structure size (in this case, the conducting island size) also affects the parameters of the $Cr/Cr_2O_3/Cr$-based device, similar to the case in Fig. 3.

The analyzed parameters of one-electron transistors with planar and spherical configurations can significantly depend on the structure design and technological characteristics, in particular on the source (drain) tunnel-junction width. A decrease in d_s leads to the operating temperature and cutoff frequency increase; as the barrier height ϕ increases, the parameters become more

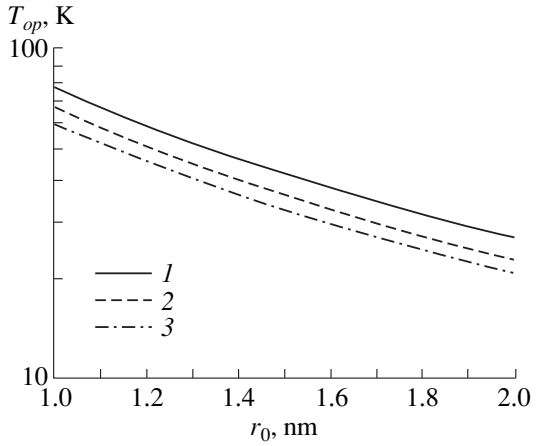


Fig. 4. Dependence of the operating temperature on the conducting island radius for the transistor with a spherical configuration, based on $Al/AlO_x/Al$ tunnel junctions for $\phi = 1.6$, 2.5, and 1.6 eV; and $\varepsilon_d = 8$, 8, and 10 (curves *1–3*, respectively).

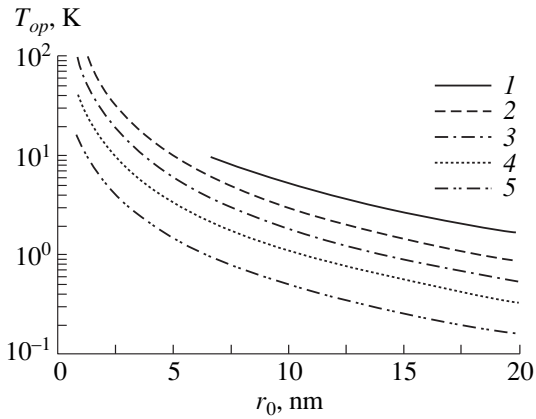


Fig. 5. Dependence of the operating temperature on the conducting island radius for spherical-configuration transistors based on various metals: (*1*) $Cr/Cr_2O_3/Cr$, (*2*) $Al/SiO_2/Al$, (*3*) $Au/Al_2O_3/Au$, (*4*) $Ti/TiO_x/Ti$, and (*5*) $Nb/NbO_x/Nb$.

sensitive to d_s variation. To achieve a higher ease of device production, one should choose systems with a lower sensitivity to structure parameter variation. Therefore, the systems with a lower potential barrier are preferable: $Cr/Cr_2O_3/Cr$, $Nb/NbO_x/Nb$, $Ti/TiO_x/Ti$, and $Nb/Al_2O_3/Nb$. If the device parameters should be more sensitive to tunnel-junction width variations, it is expedient to choose the following systems: $Al/SiO_2/Al$, $Au/Al_2O_3/Au$, and $Al/AlO_x/Al$.

For instance, it is evident from Fig. 2 that the decrease in d_s from 1 to 0.7 nm for the $Al/SiO_2/Al$-, $Au/Al_2O_3/Au$-, and $Ti/TiO_x/Ti$-based transistors (with heights of the potential barrier of tunnel junctions ϕ = 3.2, 1.65, and 0.285 eV, respectively) leads to a T_{op} increase by approximately 220, 52, and 6 times, respectively.

Larger limiting parameters T_{op} and F_t of the one-electron transistor with a spherical configuration can also be attained by reducing the conducting island radius. For example, an r_0 decrease from 2 to 1 nm for the $Al/AlO_x/Al$-based transistor leads to an approximately threefold increase in T_{op} (see Fig. 4).

CURRENT–VOLTAGE CHARACTERISTICS OF ONE-ELECTRON TRANSISTORS

We also studied the *I–V* characteristics of one-electron transistors with a planar configuration, based on the $Nb/NbO_x/Nb$ and $Ti/TiO_x/Ti$ tunnel junctions. Such devices were shown to have acceptable limiting parameters (operating temperature and cutoff frequency) compared to a number of structures based on other metal compounds with the same configuration. Calculations carried out using the proposed 2D numerical model demonstrated the following. The difference in currents of devices of identical sizes for fixed biases and temperatures can be significant for dissimilar materials, in particular, can be as large as three orders of magnitude for Nb- and Ti-based structures. A lateral gate, to which a bias is applied, significantly changes the Coulomb blockade region and only slightly affects the drain current. A decrease in the source and drain tunnel-junction width and an increase in the conducting island sizes and in temperature leads to the drain current growth. The difference in tunnel-junction widths for the source and drain and a lowering of the environment temperature lead to a pronounced stepped *I–V*-characteristic shape, which also depends on the used material. For instance, the $Ti/TiO_x/Ti$-based transistor is characterized by a longer period and a more distinct shape of steps compared to the $Nb/NbO_x/Nb$-based device for identical design and technological parameters, biases, and operating temperatures. We demonstrate the established trends with specific examples.

We calculated the *I–V* characteristics for two $Nb/NbO_x/Nb$-based structures of identical size, one of which has no lateral gate and another which has one with an applied bias of 0.2 V. The constructed dependences differ by no more than 10%. Furthermore, the obtained *I–V* characteristic of the transistor with a side gate, as distinct from that of the device without a gate, does not exhibit the Coulomb blockade region (a characteristic portion in the range of low drain voltages with a zero current).

The transistor *I–V* characteristic is substantially affected by the width of the source and drain tunnel junctions, as well as by the conducting island sizes. An increase in the width of the source and drain tunnel junctions lowers the structure current, while an increase in the island size causes the current to grow. A combined variation of these parameters can mutually compensate for their impact. For instance, calculations of the $Nb/NbO_x/Nb$-based device *I–V* characteristic yielded the following results. For a drain voltage of

0.1 V and widths of source and drain tunnel junctions $d_s = d_d = 17.5$ nm, the transistor current is approximately six times higher than for $d_s = d_d = 18$ nm; for island sizes of 50×50 nm^2, the current is approximately by an order of magnitude higher than for a 20×20 nm^2 island.

Figure 6a displays *I–V* characteristics for one-electron Ti/TiO$_x$/Ti-based transistors for different widths of the source and drain tunnel junctions ($d_s = 17.5$ nm and $d_d = 18$ nm) and for equal parameters $d_s = d_d = 18$ nm (curves *2* and *1*, respectively). The calculations were carried out for an operating temperature of 103 K. It is evident that curve *2* has a more pronounced step-type shape compared to curve *1*. We also studied the impact of structure asymmetry (the difference in widths of the source and drain tunnel junctions) for the device based on Nb/NbO$_x$/Nb (see Fig. 6b). Characteristics were calculated for the same input data as for the Ti/TiO$_x$/Ti-based transistor: curves *1* and *2* correspond to the source and drain junction widths $d_s = d_d = 18$ nm and $d_s = 17.5$ nm and $d_d = 18$ nm. The dependences (Fig. 6) show that the structure asymmetry gives rise to steps in the *I–V* characteristic which represent an attribute of the one-electron tunneling effect at low temperatures and (or) small structure sizes. The step period and size differ for dissimilar systems (Nb/NbO$_x$/Nb and Ti/TiO$_x$/Ti). The asymmetric Ti/TiO$_x$/Ti-based transistor is characterized by a longer period (about 0.1 V) and a more distinct step shape than the Nb/NbO$_x$/Nb-based device (a period of about 0.045 V). Thus, to obtain a steplike dependence of current on the drain voltage, one should choose an asymmetric structure (with different widths of the source and drain tunnel junctions), which is, as a rule, observed experimentally.

Calculations of the *I–V* characteristics for two Nb/NbO$_x$/Nb- and Ti/TiO$_x$/Ti-based structures of identical size at a fixed operating temperature yielded dependences with currents differing by several orders of magnitude for the same drain voltage. In the case of a lower potential barrier and a higher insulator dielectric constant, the structure current is larger (for the Nb/NbO$_x$/Nb-based transistor) (see Fig. 6).

One of the factors affecting the one-electron transistor's *I–V* characteristic is the operating temperature: the structure current increases with T_{op}. For instance, the current changes by approximately 30% as the operating temperature is elevated from 100 to 110 K in the Nb/NbO$_x$/Nb-based device with source and drain tunnel junction widths $d_s = d_d = 17.5$ nm, an island area of 20×20 nm^2, and a drain voltage of 0.15 V. In the asymmetric Ti/TiO$_x$/Ti-based transistor (the widths of the source and drain tunnel junctions are $d_s = 17.5$ nm and $d_d = 18$ nm) with a conducting island area of 50×50 nm^2, an operating temperature increase from 103 to 110 K changes the current by approximately 20% for a drain voltage of 0.3 V. The steps in the *I–V* characteristic disappear when T_{op} is elevated to room temperature. This is a manifestation of the one-electron tunneling suppression by thermal fluctuations in the one-electron transistor as its operating temperature grows.

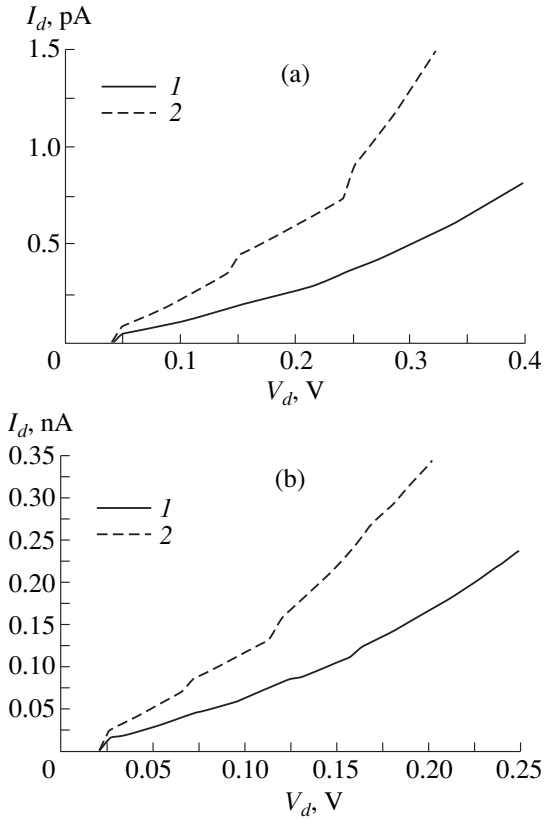


Fig. 6. Current–voltage characteristics of one-electron transistors based on tunnel junctions (a) Ti/TiO$_x$/Ti and (b) Nb/NbO$_x$/Nb for identical (curve *1*) and different (curve *2*) widths of the source and drain tunnel junctions.

CONCLUSION

A general pattern in the effect of structural, technological, and electrical parameters on basic characteristics (operating temperature, cutoff frequency, and *I–V* characteristic) of one-electron transistors based on various metal compounds (Al/AlO$_x$/Al, Al/SiO$_2$/Al, Au/Al$_2$O$_3$/Au, Nb/Al$_2$O$_3$/Nb, Ti/TiO$_x$/Ti, Cr/Cr$_2$O$_3$/Cr, and Nb/NbO$_x$/Nb) was established. The most preferable systems for one-electrode transistors were found to be Cr/Cr$_2$O$_3$/Cr, Nb/NbO$_x$/Nb, Ti/TiO$_x$/Ti, Nb/Al$_2$O$_3$/Nb and Cr/Cr$_2$O$_3$/Cr, Al/SiO$_2$/Al, Nb/Al$_2$O$_3$/Nb for devices with planar and spherical configurations (at a fixed width of tunnel junctions) and for devices with a spherical configuration (at a fixed radius of the conductive islet), respectively. If a lower sensitivity of limiting parameters to variations in the tunnel junction width is required, the best choice is Cr/Cr$_2$O$_3$/Cr, Nb/NbO$_x$/Nb, Ti/TiO$_x$/Ti, and Nb/Al$_2$O$_3$/Nb.

The above calculations of parameters of one-electron transistors using the developed technique, numerical model, and SET-NANODEV software package allow selection of appropriate materials to develop devices with required output parameters: operating temperature, cutoff frequency, and *I–V* characteristic. Such studies are expedient at the initial stages of one-electron transistor design.

ACKNOWLEDGMENTS

This work was supported in part by the Republican Scientific and Technical Programs "Computer Science," "Low-Dimensional Systems," and "Nanoelectronics."

REFERENCES

1. *NATO ASI Series B: Physica*, Vol. 294: *Single Charge Tunneling: Coulomb Blockade Phenomena in Nanostructures*, Ed. by H. Grabert and M. H. Devoret (Plenum, New York, 1992).
2. Zh. I. Alferov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 3 (1998) [Semiconductors **32**, 1 (1998)].
3. Int. J. Electron. **86** (5) (1999).
4. I. I. Abramov and E. G. Novik, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33** (11), 1388 (1999) [Semiconductors **33**, 1254 (1999)].
5. I. I. Abramov, I. A. Goncharenko, and E. G. Novik, Pis'ma Zh. Tekh. Fiz. **24** (8), 16 (1998) [Tech. Phys. Lett. **24**, 293 (1998)].
6. I. I. Abramov, I. A. Goncharenko, S. S. Ivashkevich, and E. G. Novik, in *Proceedings of the 7th Crimean Microwave Conference, KryMiKo'97, Sevastopol, 1997*, p. 398.
7. I. I. Abramov and E. G. Novik, *Numerical Simulation of Metal One-Electron Transistors* (Minsk, 2000).
8. R. H. Chen and K. K. Likharev, Appl. Phys. Lett. **72** (1), 61 (1998).
9. J. G. Simmons, J. Appl. Phys. **34** (6), 1793 (1963).
10. J. Shirakashi, K. Matsumoto, N. Miura, and M. Konagai, Jpn. J. Appl. Phys. **36** (9A/B), L1257 (1997).
11. I. I. Abramov and E. G. Novik, in *Proceedings of the 9th International Crimean Conference, KryMiKo'99, Sevastopol, 1999*, p. 292.

Translated by A. Kazantsev

Semiconductors, Vol. 34, No. 8, 2000, pp. 981–982. Translated from Fizika i Tekhnika Poluprovodnikov, Vol. 34, No. 8, 2000, pp. 1020–1021.

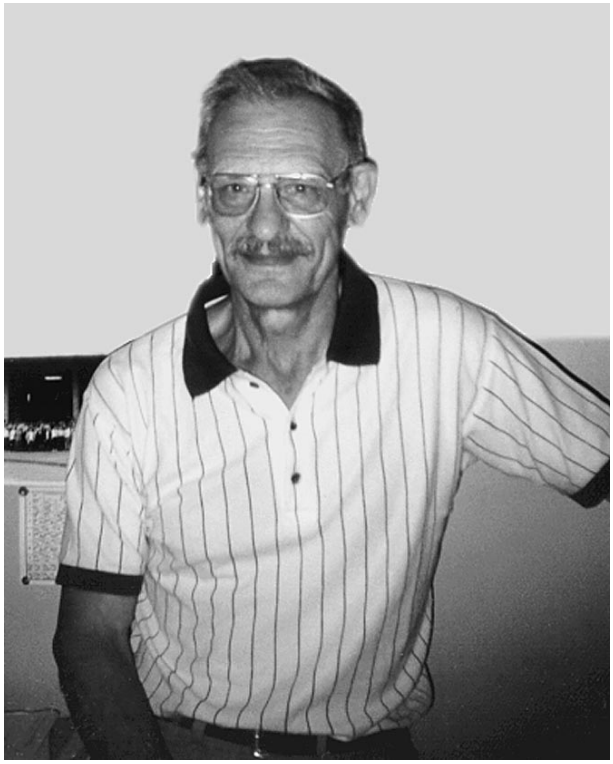
IN MEMORIAM

Izmail Arturovich Abroyan

Doctor of Physics and Mathematics Izmail Arturovich Abroyan passed away on March 1, 2000 at the age of 67; he was an outstanding scientist in the field of interaction of charged particles with solids and was a professor and the head of the Department of Applied Physics and Optics of the Solid State at St. Petersburg State Technical University.

I.A. Abroyan was born on August 26, 1933 in Petrozavodsk; however, he spent most of his life in Leningrad (St. Petersburg), except at the time of the siege of Leningrad during the Second World War. In 1957, he graduated from the Radio Engineering (now, Radiophysical) Faculty of Leningrad Polytechnical Institute (now, St. Petersburg State Technical University) with his major in physical and industrial electronics. Beginning at that time, Abroyan worked continuously at the Leningrad Polytechnical Institute, first at the Department of Physical Electronics as an assistant lecturer (1957–1964), then as an associate professor (1964–1972), and, finally, as a professor (1972–1982). Starting from 1982, he was a professor at the Department of Plasma Physics and, from 1986, he was a professor and the head of the department of the Physics of Dielectrics and Polymers; in 1994, the latter department became the Department of Applied Physics and Optics of the Solid State of St. Petersburg State Technical University. During 1981–1987, Abroyan was the dean of the Radiophysical Faculty. In 1963, Abroyan received the degree of Candidate of Science and, in 1971, the degree of Doctor of Physics and Mathematics.

The beginning of Abroyan's scientific activity was related to studies of the ion–electron emission from semiconductors, alkali halide compounds, and oxide cathodes. As a result, it became possible to gain insight into the relation of parameters that characterize the given phenomenon with special features of the band structure of materials under consideration. These studies naturally brought Abroyan to the next stage of investigations, whose aim was to determine the efficiency of excitation of electron subsystem in semiconductors under irradiation with 0.1–10-keV ions, in which case the probability of this excitation (as was assumed at that time) was very low. However, as was demonstrated in Abroyan's first study, performed in collaboration with V.A. Zborovskiĭ (1962), generation of electron–hole pairs in semiconductors is quite important even for the ions with energies that are lower by 1–2 orders of magnitude than the Seitz threshold, which was generally accepted at that time. It was this publication that was cited by I. Lindhard when he made known his classical theory of ion stopping in materials.



The subsequent studies performed with the participation of V.A. Koryukin, A.I. Titov, L.A. Tsikhnovicher, and others made it possible to conceive the entire pattern of electron excitation in semiconductors irradiated with various ions with energies in the above range.

The beginning of the 1960s was remembered for the discovery of phenomena that accompanied the interaction of accelerated ions with single crystals, such as channeling and blocking. Abroyan was one of the first to study these phenomena. In particular, he and coworkers discovered an anisotropy in the excitation of electrons in single-crystal semiconductors and dielectrics if these materials are irradiated with ions and electrons. On the basis of extensive studies of anisotropy in interaction of ions and electrons with crystalline solids, Abroyan and coworkers (V.S. Belyakov, O.A. Podsvirov, and A.I. Titov) suggested and developed two new methods for monitoring the state of the structure of near-surface layers in crystals.

The use of these unconventional methods and also of other conventional methods made it possible in the late 1960s to start fundamental studies of radiation damage in semiconductors under ion bombardment.

1063-7826/00/3408-0981$20.00

These studies, performed by Abroyan in cooperation with his coworkers and followers (V.S. Belyakov, L.M. Nikulina, A.I. Titov, A.V. Khlebalkin, and others), were very fruitful and formed the basis of Abroyan's scientific activity for the remainder of his life. As a result, the basic laws pertaining to the structural damage of semiconductors subjected to irradiation with slow ions and to implantation of light ions of medium energies into semiconductors were revealed and insight was gained into the mechanisms of the phenomena observed. In particular, the nonadditivity and noncommutativity of accumulation of defects were discovered and explained. The dependences of the effect of ion-stimulated recovery of the crystal structure (the ion-stimulated low-temperature annealing of defects and epitaxial crystallization of amorphous layers) on the ion-irradiation parameters and the target properties were clarified. Specific features of structural damage of silicon by light atomic ions as compared to the case of bombardment with molecular ions were determined.

On the basis of the results of his studies, Abroyan and his colleagues have published more than 150 papers, including two reviews in the journal *Uspekhi Fizicheskikh Nauk*, a monograph, and the first textbook in this country (1984) on the physical foundations of electron- and ion-related technology. Abroyan has participated in more than 60 International and All-Union conferences and symposia; at many of these, he was a member or the head of program or organizing committees. Abroyan was a member of the councils of the Russian Academy of Sciences on the problems in Plasma Physics and Radiation Physics of the Solid State.

Abroyan's pedagogical duties were no less important to him than his scientific activities. He was one of the favorite and most respected lecturers at St. Petersburg State Technical University; Abroyan prepared and taught a large number of unconventional courses of lectures. The feedback, the level of understanding of the subject matter by students, and their opinion about the lectures were always extremely important to him. Abroyan was scientific supervisor of a large number of students' university-degree projects and of master's degree theses and was the scientific consultant for ten candidate and two doctoral dissertations. In addition, Abroyan, as an expert of the highest level in his field of research, had a deep knowledge of other fields of physics. A combination of adherence to high principles and, at the same time, benevolence was characteristic of Abroyan; as a result, he was one of the most desired external examiners for scientific-degree candidates and enjoyed extremely high authority in all academic councils that conferred degrees. As a consequence, Abroyan led a great number of candidates to the scientific community. Abroyan himself was the chairman or a member of many doctor- and candidate-degree conferring councils at St. Petersburg State Technical University and the Russian Academy of Sciences.

Abroyan spared no efforts to improve the teaching methodology in higher education institutions. For many years, he was a member and the chairman of the Methodology Council of the Radiophysical Faculty, a member of Methodology Council at St. Petersburg State Technical University, and (for 15 years) a member of the Scientific–Methodological Council on Electronics of the Ministry of Higher-Education Institutions of the USSR. It was Abroyan who suggested the introduction of a new course in training specialists in Russia: technical physics. Students are taught according to this course in many leading higher education institutions of the country. Abroyan was the deputy chairman of the Scientific–Methodological Council dedicated to this subject in the teaching of the Ministry of Education of Russia, starting from the initial organization of this Council to Abroyan's last days.

Abroyan was an outstanding person; he was sincere, decent, and noble and had a high sense of responsibility for all efforts he undertook. He was demanding of his coworkers, but, above all, was demanding of himself and exhibited a fantastic ability for hard work. A democratic attitude and a sense of citizenship were not trivial matters to him. All this, combined with a friendly disposition, a sense of humor, an innovative intellect, and superior erudition, aroused a great fondness for him even among people who were not closely acquainted with him.

The bright memory of Abroyan as an exceptional person, scientist, and teacher will always remain in our hearts.

Colleagues, friends,
and the members of the editorial staff
of this journal

Translated by A. Spitsyn