Diffusion of boron and phosphorus in silicon during high-temperature ion implantation

G. V. Gadiyak

Institute of Computational Technology, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted April 3, 1995; accepted for publication September 2, 1996)
Fiz. Tekh. Poluprovodn. **31**, 385–389 (April 1997)

A theoretical model is proposed for describing the accelerated diffusion of boron and phosphorus impurity in silicon during high-temperature implantation. The model takes into account the production and diffusion of impurities and point defects and the formation and decomposition of defect–impurity pairs. The formation of dislocations and an increase (decrease) of their size during bombardment is taken into account. Dislocations are sinks for defects and impurities. The computed profiles are compared with existing data. It is shown that dislocations play an important role as trapping centers. Oscillatory behavior of the impurity profile was observed near the surface; this behavior is due to impurity implantation and diffusion processes occurring simultaneously. © *1997 American Institute of Physics.* [S1063-7826(97)00104-X]

High-temperature implantation is being developed for the purpose of suppressing the accumulation of defects and decreasing the temperature and duration of subsequent treatments. Under ''hot irradiation'' conditions the formation of defects is determined by the ratio of their rates of introduction and annealing as well as the appearance of structural imperfections in the form of dislocations. The implantation of an impurity under these conditions is accompanied by radiation-accelerated diffusion and precipitation of the impurity on dislocations.

The first experiments on hot irradiation were performed in Refs. 1 and 2 with target temperatures $T=600-700\ {}^\circ\mathrm{C}$. The results for a wide temperature range 600–1100 °C are presented in Refs. 3 and 4. Accurate theoretical calculations of the impurity profiles, reproducing the conditions of hot irradiation, have still not been performed. We note only the empirical model proposed in Ref. 5, where the diffusion coefficient is given phenomenologically as a function of the coordinate.

In this paper we present a theoretical model that takes into account the multicomponent character of the diffusion of an impurity in silicon, chemical reactions between the components, the possibility of trapping of impurities and point defects on dislocations, and the increase in the dislocation density and the dislocation loop sizes with the irradiation dose. The results are compared with experimental data and other models.

1. MODEL

The proposed model is an elaboration of the model in Refs. 6 and 7, where the problem of radiation-accelerated diffusion of boron during high-temperature treatment of a silicon target with protons was solved.[8,9]

According to modern ideas (see, for example, Refs. 10 and 11), the diffusion of impurity boron or phosphorus occurs with the aid of the diffusion of pairs (defect–impurity) and chemical decomposition and formation reactions between components — a point defect d, an impurity atom M_i at a site of the silicon lattice, and a pair P_i (defect–impurity) — according to the scheme

$$M_i+d \underset{k_2}{\overset{k_1}{\rightleftarrows}} P_i, \tag{1}$$

where $i=1$ and 2; M_1 is an impurity boron atom; P_1 is a defect–boron pair; M_2 is an impurity phosphorus atom, and P_2 is a defect–phosphorus pair. The reaction (1) describes the formation and decomposition of pairs. The constants for the forward and backward reactions (1) were estimated according to the formulas

$$k_1(T)=4\pi D_d(T) r_0 \xi_1$$

and

$$k_2(T)=(\tau_1)^{-1}\exp(-e_3/k_BT),$$

where r_0 is the impurity-atom–defect interaction radius, D_d is the diffusion coefficient of defects, e_3 is the binding energy of a pair, and τ_1 is a rate factor (the number of bond-breaking attempts per unit time). We took into consideration in addition the possibility of trapping of an impurity and a point defect on a dislocation

$$d+\text{disl.} \overset{k_3}{\Rightarrow} \text{d.disl.},$$

$$P_i+\text{disl.} \overset{k_4}{\Rightarrow} M_i-\text{disl.} \tag{2}$$

The constants in the reactions (2) were determined by the relations

$$k_3(T,x)=2\pi D_d R_{\text{dis}}(x)\xi_2/r_0,$$

$$k_4(T,x)=2\pi D_p R_{\text{dis}}(x)\xi_3/r_0,$$

where $R_{\text{dis}}(x)$ is the radius of the dislocation loops; D_p is the diffusion coefficient of pairs P_1 with boron atoms or P_2 with phosphorus atoms, respectively; and, ξ_1, ξ_2, and ξ_3 are attachment coefficients. High-temperature ion bombardment gives rise to a high supersaturation of point defects above their equilibrium densities at a given temperature. (This difference between the equilibrium and nonequilibrium densities can reach two to four orders of magnitude for boron and phosphorus, respectively.) They decrease via several channels — recombination of interstices with vacancies, trapping on dislocations, and production of new dislocations. This is indicated by the experimental data of Refs. 3 and 4. As a result of ion bombardment, the impurity atoms are arranged either at the sites of the silicon lattice (the impurity atoms do not diffuse because of the high energy barrier) or they form pairs (which do diffuse). In the most general case it would also be necessary to take into account all charge states in the system, which greatly increases the dimension of the problem but does not lead to any qualitatively new results. For this reason, we can write the following system of equations for the densities C_s and C_p of impurities occupying lattice sites and in pairs, respectively:

$$\frac{\partial C_d}{\partial t}=\frac{\partial}{\partial x}\left(D_d\frac{\partial C_d}{\partial x}\right)+g_d-(C_d-C^0)/\tau_d-k_1C_dC_s+k_2C_p-k_3C_dC_{\mathrm{dis}}, \tag{3}$$

$$\frac{\partial C_p}{\partial t}=\frac{\partial}{\partial x}\left(D_p\frac{\partial C_p}{\partial x}\right)+\xi g_p+k_1C_dC_s-k_2C_p-k_4C_pC_{\mathrm{dis}}, \tag{4}$$

$$\frac{dC_s}{dt}=(1-\xi)g_p-k_1C_dC_s+k_2C_p, \tag{5}$$

$$\frac{\partial C_{\mathrm{dis}}}{\partial t}=-\beta_1(-\beta_1C_{\mathrm{dis}}-g_{1\mathrm{dis}}(x)), \tag{6}$$

$$\frac{\partial C_{\mathrm{com}}}{\partial t}=k_4C_pC_{\mathrm{dis}}, \tag{7}$$

$$\frac{dR_{\mathrm{dis}}}{dt}=D_d\alpha_1(T)(C_d-C^0)-\gamma_1(T)D_{\mathrm{Si}}C_{\mathrm{Si}}, \tag{8}$$

where C_d, C^0, C_p, C_s, C_{com}, C_{dis}, and C_{Si} are the density of point defects and their equilibrium density at the target temperature and the density of pairs (mobile components), concentration of stationary site impurity atoms, impurity atoms trapped on dislocations, the dislocation density, and the density of silicon atoms in an ideal lattice; D_{Si} is the self-diffusion coefficient of silicon atoms; $g_d(x)$, $g_p(x)$, and $g_{\mathrm{dis}}(x)$ are the rates of generation of defects, impurities, and dislocations

$$g_d(x)=(2/\pi)^{1/2}I_0E_i/(E_f\sigma_{rd})\exp[-((x-r_d)/\sigma_{rd})^2],$$

$$g_p(x)=(2/\pi)^{1/2}I_0/\sigma_{rp}\exp[-((x-r_p)/\sigma_{rp})^2],$$

$$g_{\mathrm{dis}}(x)=(2/\pi)^{1/2}I_0/\sigma_{\mathrm{dis}}\exp[-((x-r_d)/\sigma_{\mathrm{dis}})^2],$$

where I_0 is the beam current, r_d is the position of the maximum of the elastic losses, σ_{rd} is the width of the elastic-losses peak, σ_{dis} is the width of the peak in the rate of generation of the dislocation density, E_i are the total inelastic energy losses, E_f is the threshold formation energy of a Frenkel' pair in silicon, r_p is the position of the maximum of the generation rate of impurity atoms, and σ_{rp} is the variance in the generation rate of impurity atoms.

The boundary conditions for the system of equations presented above are

$$C_d(x=0,t)=0,\quad C_d(x=L,t)=C^0,$$

$$\frac{\partial}{\partial x}C_p(x=0,t)=0,$$

$$C_d(x=L,t)=0.$$

The initial conditions for the system of equations are

$$C_d(x,t=0)=C^0,\quad C_p(x,t=0)=0,\quad C_{\mathrm{com}}(x,t=0)=0,$$

$$R_{\mathrm{dis}}(x,t=0)=R_0,\quad C_{\mathrm{dis}}(x,t=0)=C^0_{\mathrm{dis}}.$$

2. METHOD FOR CALCULATING AND ESTIMATING THE PARAMETERS

The system of equations (3)–(8) was solved numerically by an implicit scheme. The time derivatives were approximated to first-order accuracy. The time step was chosen so as to increase in time with increment $\delta t^n/\delta t^{n-1}=1.1$. A conservative scheme on a nonuniform grid of second-order accuracy was used to approximate the spatial derivatives. The nonlinear system of equations was solved successively by the sweep method using the solution from the preceding iteration. The computational accuracy was no worse than 0.1%.

The following considerations were used to estimate the parameters k_1, k_2, τ_d, D_d, D_p, C^0, g_d, g_p, d_{dis}, β_1, α_1, and γ_1. The Arrhenius law was assumed for the diffusion coefficients D_d and D_p. The activation energy was assumed to be 1.7–1.9 eV for interstices[12,13] and 1.2–1.5 eV for vacancies.[11,14,15] For this reason, just as in Ref. 6, the value 1.5 eV was used to estimate D_d. The preexponential factor D_{d0} was calculated from the condition that near the melting point D_d is close to 4×10^{-4} cm^2/s.[16] Since the pair diffusion coefficient D_p is not an impurity diffusion coefficient and it cannot be estimated experimentally, it was assumed that the pair diffusion coefficient must be close in order of magnitude to the defect diffusion coefficient $D_p\sim D_d$. To calculate the pair decay constant k_2, the values $\tau_1^{-1}\simeq10^{-2}f_0$ (where $f_0\simeq10^{13}\ \mathrm{s}^{-1}$ is the characteristic atomic vibrational frequency) were used. The pair binding energies were assumed to be 1.1 eV for boron and 1.4 eV for phosphorus.[15,17] The experimental values[18] $\tau_d\simeq10^{-7}$ s and $\beta_1\simeq10^{-1}\ \mathrm{s}^{-1}$ were used for the defect lifetime τ_d and the dislocation velocity β_1; the interaction radius was assumed to be $r_0\simeq1.2\times10^{-7}$ cm. The equilibrium density of point defects was calculated according to the formula $C_0\simeq5\times10^{25}\cdot e^{-E_f/kT}$, where $E_f=3.6$ eV is the formation energy of point defects.[15,17] The parameters for the functions g_d, g_p, and g_{dis} are presented in Table I. The tables presented in Ref. 19 were used to calculate them.

3. COMPUTATIONAL RESULTS AND DISCUSSION

The calculations were performed for implantation of boron with energy 45 keV and phosphorus with energy 100 keV and beam current density $I_0=10^{-6}$ A/cm^2 in a silicon target heated up to temperature $T=900\ ^\circ\mathrm{C}$.

Several models employed in the literature for predicting the impurity profile in the case of implantation into a heated target are compared in Fig. 1.

TABLE I.

Impurity	r_d, cm	σ_{rd}, cm	r_p, cm	σ_{rp}, cm	E_i, keV	E_f, eV	$\sigma_{\rm dis}$, cm
Boron	1.0666×10^{-5}	5.3×10^{-6}	1.470×10^{-5}	5.070×10^{-3}	13.6	20.0	2×10^{-5}
Phosphorous	0.89×10^{-5}	4.86×10^{-5}	1.245×10^{-5}	4.62×10^{-6}	48.8	20.0	10^{-5}

Model 1 is described by the diffusion equation (4), where the constants are assumed to be zero. The value of the diffusion coefficient was taken from the experimental data in Ref. 3.

Model 2 is described in Ref. 5. It is based on prescribing phenomenologically the diffusion coefficient by an exponential function of depth

$$D_i(x)=D_{i0}\exp(-x/L_{id}), \tag{9}$$

where the index i denotes the boron $(i=1)$ and phosphorus $(i=2)$ impurities. The parameters D_{i0} and L_{id} were taken from Ref. 5.

Model 3 is the above-described model in the complete formulation (3)–(8). Here the smoothed experimental data of Ref. 3 are presented. As one can see from the figure, the behavior of an impurity cannot be described in a wide range of distances by a single diffusion equation. Radiation-accelerated diffusion plays an important role (compare curve *4* for $D=8\times10^{-12}$ cm^2/s and curve *3* for $D=2\times10^{-15}$ cm^2/s in Fig. 1; Ref. 20). The use of the coefficient $D_i(x)$ in the form (9) somewhat improves the agreement with experiment but it does not reproduce the experimental data deep in the substrate. Only the model (1)–(8), proposed above, makes it possible to describe satisfactorily the impurity distribution profile not only near a peak but also deep in the substrate.

We note that the specific weight of phosphorus trapped on dislocation sinks is very small (less than 1%) and therefore dislocations have virtually no effect on the formation of the impurity density profile. In the case boron, taking account of the trapping of boron on dislocations determines the impurity density at large distances from the position of the maximum of the impurity distribution profile. In the case of a high dislocation density the diffusion impurity flux decreases substantially, since some boron atoms are trapped by these dislocations.

The numerical results obtained for boron and phosphorus using the model described above are presented in Figs. 2–6.

After a short time $t=10^{-3}$ s (which is much longer than the lifetime τ_d of point defects) a stationary distribution profile of point defects is established in the system. It is determined by the lifetime τ_d and generation rate $g_d(x)$ of defects. The presence of diffusion results in broadening of the profile by the amount $l_d=(D_d\tau_d)^{1/2}$. This stationary distribution of the defect profile starts to change when a sufficiently high dislocation density accumulates in the system. The defect lifetime will then be determined not by the vacancy–interstice recombination time but rather by the travel time of a defect to the nearest dislocation

$$\tau_{d2}^{-1}=k_3C_{\rm dis}=2\pi D_dR_{\rm dis}(x)\xi_2\cdot C_{\rm dis}/r_0\,.$$

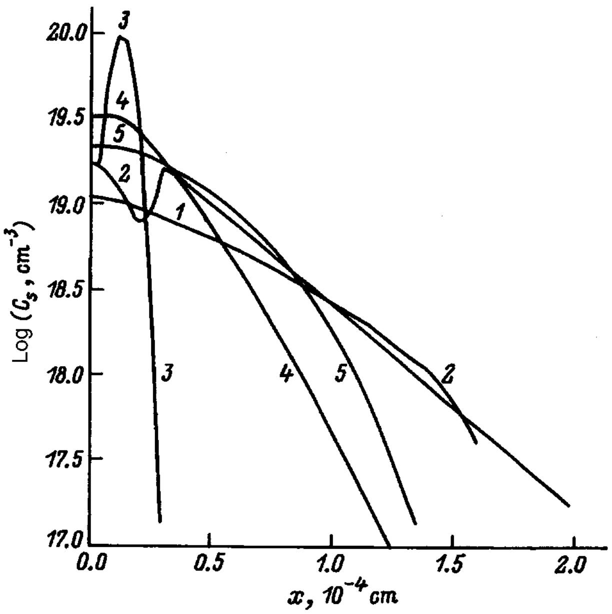

FIG. 1. Phosphorus impurity density profile for different models. Dose: 10^{15} cm^{-2}. *1* — Experimental data; *2* — model 3; *3* — model 1, $D=2\times10^{-15}$ cm^2/s; *4* — model 1, $D=8\times10^{-12}$ cm^2/s; *5* — model 2.

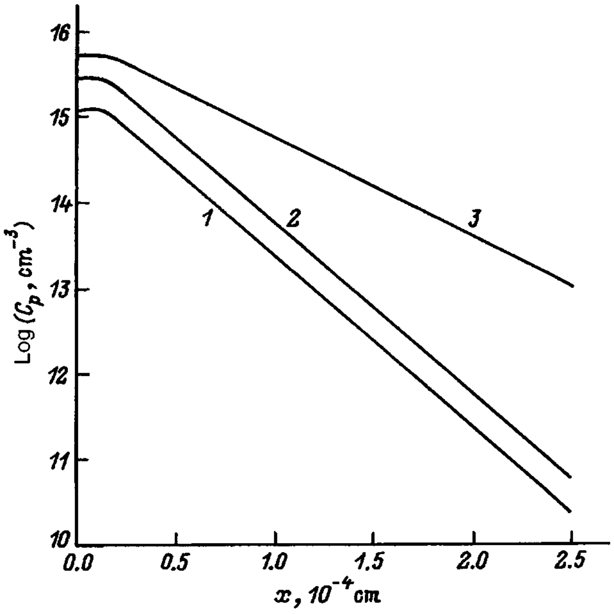

FIG. 2. Phosphorus pair density for different irradiation doses, cm^{-2}: *1* — 10^{13}, *2* — 10^{14}, *3* — 10^{15}.

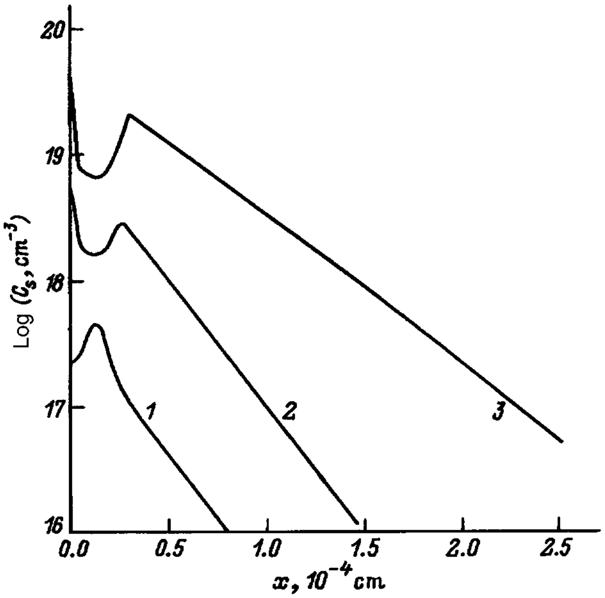

FIG. 3. Site phosphorus impurity density for different irradiation doses, cm^{-2}: *1* — 10^{13}; *2* — 10^{14}; *3* — 10^{15}.

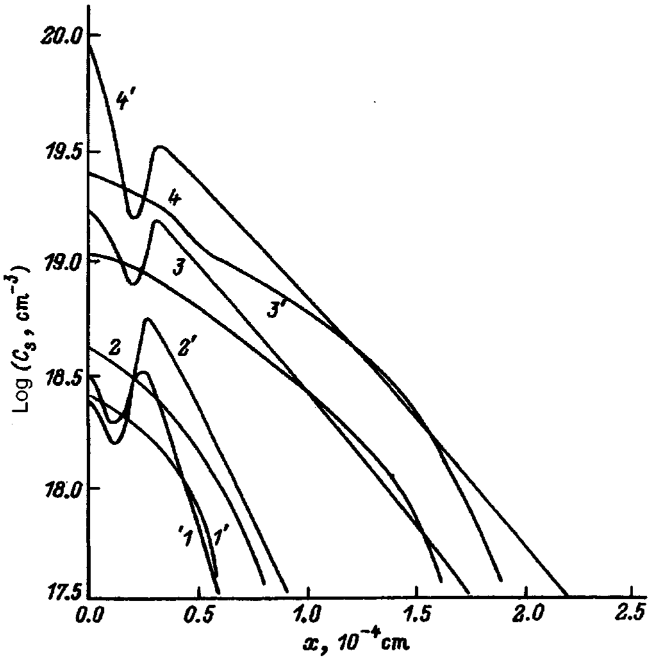

FIG. 5. Phosphorus impurity concentration profile for different irradiation doses, cm^{-2}: *1,1'* — 10^{14}; *2,2'* — 3×10^{14}; *3,3'* — 10^{15}; *4,4'* — 10^{16}; *1–4* — experimental data; *1'–4'* — computational results.

This instant of time occurs when the dislocation density becomes of the order of 10^{10} cm^{-2} and higher, and the radius of the dislocation loops becomes 10^{-5} cm and larger.

The profiles of pairs (defect–impurity) and interstitial impurity phosphorus atoms, respectively, are presented in Figs. 2 and 3 for different values of the irradiation dose. For short irradiation times, the impurity distribution profile reproduces the rate distribution profile $g_p(x)$ of the atoms introduced into the volume. This distribution has a maximum at the point r_p and its amplitude increases with the dose. Next, the position of the peak in the pair profile shifts in the direction of the surface and is broadened because of diffusion. Pair diffusion occurs in both directions away from the point r_p, the pairs decompose, and the density of interstitial atom increases near the surface. Because of the reaction (1), the pair density will then beging to increase. As a result, the peak shifts toward the surface.

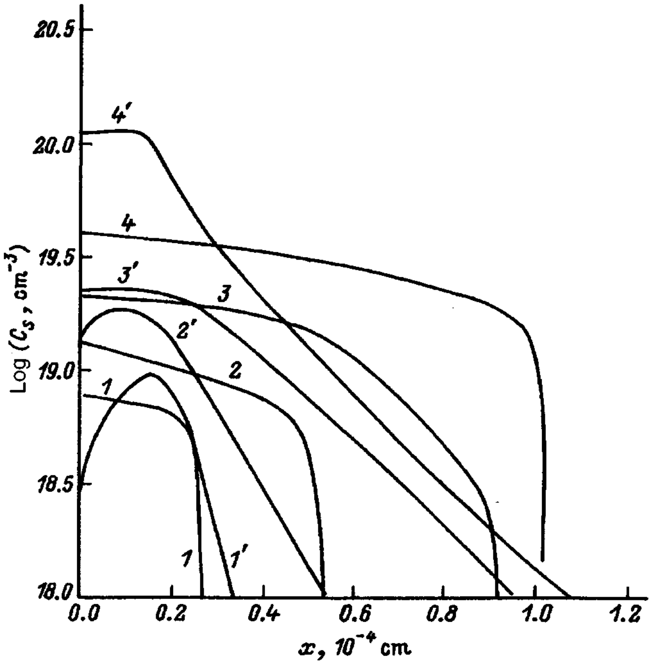

FIG. 4. Boron impurity concentration profile for different irradiation doses, cm^{-2}: *1,1'* — 2×10^{14}; *2,2'* — 5×10^{14}; *3,3'* — 10^{15}; *4,4'* — 3×10^{15}; *1–4* — experimental data; *1'–4'* — computational results.

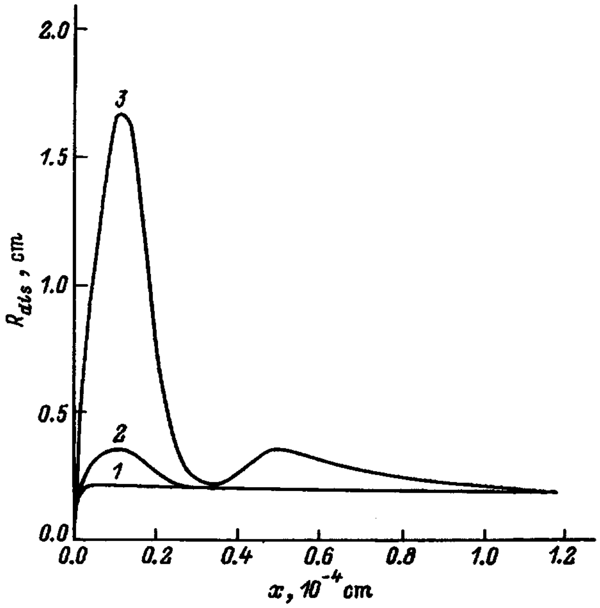

FIG. 6. Radius of dislocation loops in the case of irradiation of silicon with boron for different doses, cm^{-2}: *1* — 10^{13}, *2* — 10^{14}, *3* — 10^{15}.

The surface is a sink for migrating atoms. For this reason, for high doses, the pair distribution profile has a maximum at the surface.

In the case of interstitial impurity (Fig. 3), in addition to a peak near the r_p, a new impurity peak arises in time near the surface. The peak itself shifts in time into the volume of the target to a depth of the order of $(D_p \cdot k_2^{-1})^{1/2}$, where pair decomposition occurs. As a result, the profile can exhibit an oscillatory behavior as a function of the coordinate, similar to the behavior observed in Ref. 4.

The coordinate dependence of the dislocation loop radius is shown in Fig. 6 for different boron irradiation doses. As one can see, for low doses (less than 10^{13} cm^{-2}) the distribution of radii as a function of the coordinate is uniform and corresponds to the initial value of the radius (of the order of 30 Å). As the dose increases further, the peak with a maximum at the point r_d starts to grow, and for a dose of the order of 10^{15} cm^{-2} a second peak appears near $l_d + r_d$. Its appearance is due to the deformation of the point-defect profile. A powerful absorber of point defects appears near the first peak, and the supersaturation with defects on the ''tail'' of the point-defect generation function $g_d(x)$ is different from zero.

We note that to obtain good agreement with experiments,[3,4] we employed an impurity attachment coefficient at a dislocation of the order of 1 in the case of boron and 10^{-5} for phosphorus. Therefore, boron is trapped by structural imperfections more easily than phosphorus. The pair diffusion coefficient for phosphorus was an order of magnitude lower than for boron. The absence of oscillations in the experimental impurity profile[5] is explained by the fact that their amplitude apparently turned out to be small and in Ref. 5 only smoothed curves were presented. The amplitude of the oscillations can be decreased by using the parameters of the problem $g_{\rm dis}, \alpha_1$, and k_4. In this case, however, unrefined basic experimental profiles must be available.

We wish to thank G. A. Kachurin and I. E. Tysshchenko for many useful discussions of the problem and its results. We also thank A. V. Bibik and I. I. Gadiyak for assistance in preparing the manuscript.

[1] B. L. Crowder and J. M. Fairfield, J. Electrochem. Soc. **117**, 363 (1970).

[2] J. W. Mayer, L. Eriksson, and J. A. Davis, *Ion Implantation in Semiconductors*, Academic Press, N. Y. (1970).

[3] G. A. Kachurin, I. E. Tyschenko, and A. P. Mazhirin, Nucl. Instrum. Meth. B **43**, 535 (1989).

[4] G. A. Kachurin, I. E. Tyschenko, and L. I. Fedina, Nuc. 1. Instrum. Meth. B **68**, 323 (1992).

[5] L. N. Aleksandrov, T. V. Bondareva, G. A. Kachurin, and I. E. Tyschenko, Fiz. Tekh. Poluprovodn. **25**, 227 (1991) [Sov. Phys. Semicond. **25**, 137 (1991)].

[6] G. A. Kachurin, G. V. Gadiyak, V. I. Shatrov, and I. E. Tyschenko, Fiz. Tekh. Poluprovodn. **26**, 1978 (1992) [Sov. Phys. Semicond. **26**, 1111 (1992)].

[7] G. V. Gadiyak and D. E. Blaginin, COMPEL **12**, 407 (1993).

[8] P. Baruch, Inst. Phys. Conf. Ser. **31**, 126 (1977).

[9] V. V. Kozlovskiĭ and V. N. Lomasov, Zh. Tekh. Fiz. **54**, 1157 (1984) [Sov. Phys. Tech. Phys. **29**, 658 (1984)].

[10] Hiroyuki Kinoshita and Dim-Lee Kwong, IEDM Tech. Dig, 165 (1992).

[11] D. Mathiot and J. C. Pfister, J. Appl. Phys. **55**, 3518 (1984).

[12] W. Wijarakula, J. Appl. Phys. **67**, 7624 (1990).

[13] H. U. Jager, T. Feudel, and S. Ulbricht, Phys. St. Sol. A **116**, 571 (1989).

[14] B. J. Masters and E. F. Gorvey, J. Appl. Phys. **49**, 2717 (1978).

[15] R. B. Fair in Proc. Mater. Res. Soc. Symp., **35**, *Energy Beam-Solid Interactions and Transient Thermal Processing*, edited by D. K. Biegelsen, G. A. Rozgonyi, C. V. Shank (N. Y., 1985).

[16] V. V. Voronkov, J. Cryst. Growth **59**, 7624 (1982).

[17] C. S. Nichols, C. C. Van de Walle, and G. Pantelides, Phys. Rev. B **40**, 5484 (1989).

[18] I. E. Tychenko, Author's Abstract of Candidate's Dissertation, Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk (1992).

[19] A. F. Burenkov, F. F. Komarov, M. A. Kumakhov, and M. M. Temkin, *Spatial Distribution of the Energy Released in a Cascade of Atomic Collisions in Solids* [in Russian], Énergoatomizdat, Moscow (1985).

[20] A. L. Alexandrov, M. S. Obrecht, and G. V. Gadiyak, Solid−State Electron. **35**, 1549 (1992).

Translated by M. E. Alferieff

Effect of electron and neutron bombardment on the orange luminescence spectra of not specially doped and copper-doped cadmium sulfide single crystals

G. E. Davidyuk, V. S. Manzhara, N. S. Bogdanyuk, A. P. Shavarova,
and V. V. Bulatetskiĭ

Lesi Ukrainki Volynskiĭ State University, 263009 Lutsk, Ukraine

(Submitted March 20, 1996; accepted for publication April 24, 1996)
Fiz. Tekh. Poluprovodn. **31**, 390–392 (April 1997)

CdS single crystals which were not specially doped and which were doped with copper ($N_{\mathrm{Cu}}=10^{18}$ cm^{-3}) have been investigated. It is concluded on the basis of an analysis of the dose dependences of the orange luminescence intensity ($\lambda_M=605$ nm) of ''pure'' and doped samples upon bombardment by electrons with $E=1.2$ MeV and by fast reactor neutrons that the centers responsible for this luminescence are complex in nature. They consist of interstitial cadmium atoms and oxygen atoms. Electron bombardment of CdS:Cu single crystals results in the formation of new centers which are responsible for luminescence with $\lambda_M=570$ and 545 nm. © *1997 American Institute of Physics.* [S1063-7826(97)00204-4]

In Ref. 1 it was shown that interstitial cadmium ions (Cd_i) play a large role in the formation of orange luminescence centers ($\lambda_M=605$ nm) in CdS single crystals. However, being shallow donors, they cannot be luminescence centers.[2] For this reason, different authors[3–6] propose complex models of O-luminescence centers with the participation of Cd_i:

$$(\mathrm{Cd}_i^+-V_{\mathrm{Cd}}^-),\quad(\mathrm{Cd}_i^+-\mathrm{Cu}_{\mathrm{Cd}}^-),\quad(\mathrm{Cd}_i^+-\mathrm{Ag}_{\mathrm{Cd}}^-).$$

In Ref. 7 it is concluded on the basis of an investigation of the symmetry of *O* centers that the photoluminescence centers are trimers which consist of interstitial copper or silver atoms and deep acceptors. Therefore, the nature of the *O*-luminescence centers in CdS has by no means been established, and the methods of radiation physics may be helpful in clarifying it. Such methods make it possible to introduce characteristic lattice defects in a controlled manner at comparatively low temperatures and to study their interaction with one another and with different impurities.

In our work we investigated CdS single crystals which were not specially doped and which were doped with copper ($N_{\mathrm{Cu}}=10^{18}$ cm^{-3}). The synthesis conditions of the material studied are described in Refs. 8–11. Bombardment with 1.2-MeV electrons was performed in a linear accelerator; the irradiation temperature did not exceed 20 °C. Neutron bombardment was performed in a vertical channel of a nuclear reactor in specially evacuated quartz cells with a cadmium filter. The average energy of the fast reactor neutrons was $E=2$ MeV. The temperature during irradiation did not exceed 70 °C.

The spectral characteristics and dose dependences of the *O*-luminescence intensity of cadmium sulfide single crystals were investigated. Before and after irradiation wide luminescence bands (Fig. 1) with $\lambda=605$ nm (peak I) were observed in the ''pure'' and doped single crystals. The form of the luminescence and its excitation spectra, as well as the polarization characteristics did not change after irradiation, which suggests that the *O* centers formed by radiation are identical to the *O* centers before irradiation.

The *O*-luminescence intensity in pure samples and some doped crystals, in which the initial orange luminescence is very weak, increases with the irradiation dose (Fig. 2). In CdS:Cu single crystals in which *O* luminescence dominates, the intensity of the luminescence decreases with irradiation. This case is illustrated in Fig. 1.

The nonlinearity and saturation of the dose dependence of the *O* radiation (Fig. 2) and the decrease in the luminescence intensity for $\Phi>3\times10^{17}$ electrons/cm^2 attest to the complicated nature of the *O* centers. Furthermore, the lower rate of introduction of orange-luminescence centers in CdS:Cu single crystals (Fig. 2, curve *2*), and in some doped samples the decrease in the *O* luminescence with irradiation dose, apparently suggest that Cu atoms are not constituents of *O* centers ($\Lambda_M=605$ nm).

If it is assumed, as done in many works,[1–6] that interstitial cadmium atoms are responsible for the *O* luminescence, then the dose dependence of the *O* radiation can be explained by the formation of complexes of radiation-produced Cd_i with atoms of some uncontrollable impurity, which is an effective sink for Cd_i and is present in both the pure and doped samples. Oxygen atoms can play the role of such an impurity. In many technological works,[12–14] oxygen, which is present in the lattice of CdS single crystals, is considered to be responsible for the *O* luminescence. In pure samples with a very low initial Cd_i density, most oxygen atoms are free and, interacting with Cd_i, they form *O* centers. For this reason, at the initial moment of irradiation the highest rate of introduction of orange-luminescence centers is observed in these crystals (Fig. 2, curve *1*). At high doses, a large fraction of the oxygen atoms is trapped by radiation-produced Cd_i (saturation of the sinks occurs) and the rate of introduction of *O* centers decreases. Further irradiation ($\Phi>3\times10^{17}$ electrons/cm^2) with a nearly constant concentration of O centers results in the formation of other radiative and nonradiative recombination centers, which redistribute among themselves a substantial fraction of the recombination flux,[8] decreasing the intensity of the *O* luminescence.

It is well known[15] that when cadmium sulfide single crystals are doped with copper, the density of Cd_i, which

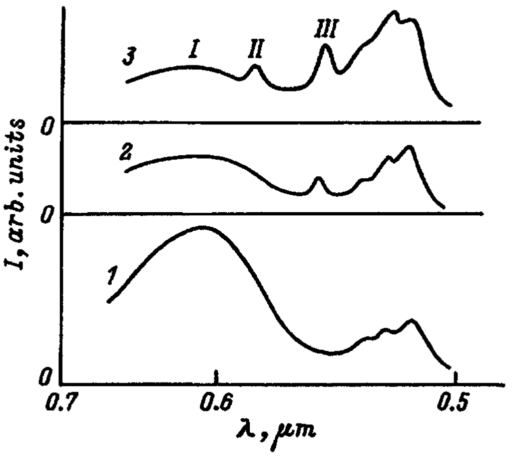

FIG. 1. Luminescence spectra of CdS:Cu single crystals ($T=77$ K) before (*1*) and after irradiation by electrons with $E=1.2$ MeV and dose $\Phi=2\times10^{17}$ cm^{-2} (*2*, *3*), 2 — before annealing, *3* — after annealing at $t=150$ °C (10 min).

can be trapped by oxygen atoms, increases. This explains the high O-luminescence intensity in the initial doped crystals and the low rate of introduction of O centers when the crystals are irradiated as compared with the pure samples (Fig. 2). In some CdS:Cu single crystals with a dominant O emission band, all oxygen atoms are apparently saturated with Cd_i and their irradiation leads at the outset (for the reasons examined above) to a decrease in the intensity of the orange luminescence (Fig. 1).

It should be noted that the radiation-produced O-luminescence centers ($\lambda_M=605$ nm) are thermally stable and essentially do not anneal up to temperatures 250 °C (Fig. 3).

Polarization measurements showed that the O-band luminescence ($\lambda=605$ nm) of CdS single crystals satisfies the condition $|P(0°)|=|P(90°)|$($|P(0°)|$ and $|P(90°)|$ are the degrees of polarization corresponding to the maxima in the polarization diagram), characteristic for isotropic centers which could consist of distributed donor–acceptor pairs. The values $P_{sp}=-0.03$ (P_{sp} is the degree of spontaneous polarization) and $P(0°)=P(90°)=-0.04$ agree with the values obtained in Ref. 7 in an investigation of not specially doped CdS samples. The inequality $|P(0°)| \neq | P(90°)|$ holds on the high-energy section of the orange band of CdS:Cu single crystals. This attests to the anisotropy of the centers responsible for the radiation in this region.

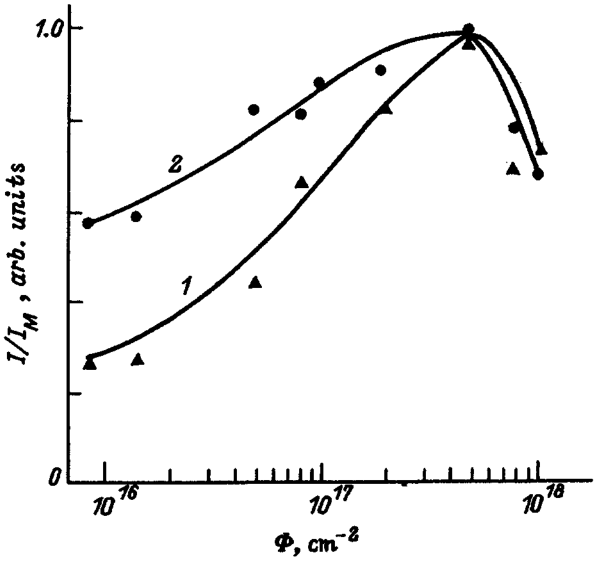

FIG. 2. Relative intensity of the $\lambda_M=605$-nm orange luminescence of CdS (*1*) and CdS:Cu (*2*) single crystals versus the dose of the electronic irradiation with $E=1.2$ MeV ($T=77$ K).

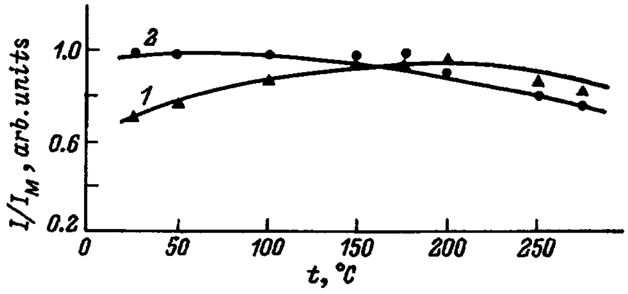

FIG. 3. Effect of isochronous annealing on the intensity of the $\lambda_M=605$-nm orange luminescence of CdS (*1*) and CdS:Cu (*2*) single crystals irradiated by electrons with $E=1.2$ MeV and dose $\Phi=2\times10^{17}$ cm^{-2}.

In doped crystals, besides the $\lambda_M=605$-nm band (peak I), there is probably also a band with $\lambda_M=570$ nm (due to anisotropic centers, peak II), which is more noticeable after irradiation and especially after the irradiated samples are annealed (Fig. 1). Irradiation of the doped crystals also results in the appearance of a weak yellow-green band with $\lambda_M=545$ nm (peak III), whose intensity increases somewhat after annealing (Fig.1). It should be noted that we observed the formation of a band with a maximum located nearby ($\Lambda_M=537$ nm) upon annealing ($t>100$ °C) samples bombarded by electrons without any special doping.[9]

Neutron bombardment of CdS and CdS:Cu single crystals always resulted in almost complete vanishing of the orange luminescence, irrespective of the initial O-radiation intensity, although other luminescence lines were clearly identified.

As is well known,[16] when cadmium sulfide single crystals are irradiated with fast neutrons, clusters of defects, which are good sinks for impurity atoms, are formed. Apparently, they are also good sinks for oxygen atoms, decreasing their density in the volume of the crystal and, correspondingly, the O-luminescence intensity.

In summary, the dose dependences of the O-luminescence intensity in CdS and CdS:Cu single crystals upon electron and neutron bombardment can be explained by assuming that complexes consisting of interstitial cadmium atoms and an uncontrollable impurity, whose role can be played by oxygen atoms, are responsible for the O centers.

[1] B. A. Kulp, Phys. Rev. **125**, 1865 (1962).
[2] V. E. Lashkarev, A. V. Lyubchenko, and M. K. Sheĭnkman, *Nonequilibrium Processes in Photoconductors* [in Russian], Naukova dumka, Kiev (1981), p. 264.
[3] B. Ermolovich, G. I. Matvievskaya, and M. K. Sheĭnkman, Fiz. Tekh. Poluprovodn. **9**, 2290 (1975) [Sov. Phys. Semicond. **9**, 1070 (1975)].

[4] P. L. Kuk, A. Ya. Érm, M. E. Altosaar, Ya. E. Mellikov, and M. D. Moin, Ukr. Fiz. Zh. **26**, 990 (1981).

[5] J. E. Ralph, Phys. Status Solidi A **53**, 611 (1979).

[6] N. Sysa, H. Watanabe, and M. Wada, Jpn. J. Appl. Phys. **15**, 2365 (1976).

[7] Yu. M. Émirov, S. S. Ostapenko, M. A. Rizakhanov, and M. K. Sheĭnkman, Fiz. Tekh. Poluprovodn. **16**, 1371 (1982) [Sov. Phys. Semicond. **16**, 879 (1982)].

[8] G. E. Davidyuk, N. S. Bogdanyuk, and A. P. Shavarova, Fiz. Tekh. Poluprovodn. **28**, 2056 (1994) [Semiconductors **28**, 1132 (1994)].

[9] N. S. Bogdanyuk, G. E. Davidyuk, and A. P. Shavarova, Fiz. Tekh. Poluprovodn. **29**, 201 (1995) [Semiconductors **29**, 99 (1995)].

[10] N. S. Bogdanyuk, G. E. Davidyuk, and A. P. Shavarova, Fiz. Tekh. Poluprovodn. **29**, 357 (1995) [Semiconductors **29**, 181 (1995)].

[11] A. P. Galushka, G. E. Davidyuk, V. T. Mak, V. I. Kuts, and N. S. Bogdanyuk, Izv. Vyssh. Uchebn. Zaved. SSSR, Fiz., No. 10, 128 (1977).

[12] N. A. Vlasenko, N. I. Vitrikhovskiĭ, Z. L. Denisova, and V. F. Pavlenko, Opt. spektrosk. **21**, 466 (1966) [Opt. Spectrosc. (USSR) **21**, 289 (1966)].

[13] G. V. Bushueva, V. I. Reshetov, A. A. Khromov, S. A. Pendyur, S. A. Nosibov, and A. N. Pechenov, Fiz. Tekh. Poluprovodn. **22**, 201 (1988) [Sov. Phys. Semicond. **22**, 124 (1988)].

[14] N. K. Morozova, A. V. Morozova, I. A. Karetnikov, L. D. Nazarova, and N. D. Danilevich,

[15] A. M. Gurevich, *Introduction to Physical Chemistry of Phosphor Crystals* [in Russian], Vyssh. Shkola, Moscow 91982), p. 372.

[16] G. E. Davidyuk, A. P. Glushka, V. S. Manzhara, and N. S. Bogdanyuk, Izv. vyssh. uchebn. zaved. SSSR, Fizika, No. 2, 37 (1980).

Translated by M. E. Alferieff

Effect of lateral transport of photoinduced charge carriers in a heterostructure with a two-dimensional electron gas

V. A. Sablikov[1)] and O. A. Ryabushkin

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 141120 Fryazino, Russia

S. V. Polyakov

Institute of Mathematical Modeling, Russian Academy of Sciences, 125047 Moscow, Russia

(Submitted November 13, 1995; accepted for publication May 20, 1996)
Fiz. Tekh. Poluprovodn. **31**, 393–399 (April 1997)

It is shown that the nonequilibrium charge carriers produced by a local optical disturbance of the heterostructure with a two-dimensional electron gas are transported in the plane of the structure over an extremely large distance from the excitation location, which greatly exceeds the diffusion length in the bulk. The effect is attributable to the fact that the photogenerated electrons and holes are separated by the built-in electric field of the heterojunction to opposite edges of the buffer layer, where they are transported along parallel planes. The distance over which the nonequilibrium carrier density spreads reaches large values because of 1) the high conductivity of two-dimensional electrons, 2) the barrier for electron–hole recombination, and 3) hole drift in the electric field produced by the charge of nonequilibrium carriers in the plane of the structure. © *1997 American Institute of Physics.* [S1063-7826(97)00304-9]

1. The effects produced by optical action on semiconductor structures (such as photoluminescence and photoreflection) are extremely important as a physical basis for contact-free methods of investigation of these structures. Ordinarily, attention is focused on uniform illumination. However, interest in local optical action on a heterostructure, for example, by a focused laser beam, has increased in recent years.[1–6] This interest stems from the development of contact-free methods of diagnostics with spatial scanning and new physical phenomena which appear under local-illumination conditions. One of the most interesting phenomena is that the effect of optical action is observed over extremely large distances from the excitation location in the plane of the heterostructure. For example, in Ref. 7 it was determined that when part of the surface of a selectively doped heterostructure is illuminated, the effect of the illumination, measured according to the reflection of a probe light beam, was observed in a shadow at a distance of up to 3 mm. As it turned out, this effect arises in the presence of a high electron mobility. The effect has still not been explained, but it is the basis for applications in the contact-free determination of the mobility or conductivity of a two-dimensional electron gas (2DEG).

In this paper we propose a mechanism for this phenomenon. The mechanism is that the photogenerated electrons and holes are separated by the built-in electric field of the heterojunction to opposite surfaces of the buffer layer and, being separated, they are transported in the plane of the heterostructure. The nonequilibrium carriers are transported over large distances because of the long lifetime of the separated electrons and holes and because of the high mobility of the 2DEG. The lateral transport length depends on the magnitude of the band bending of the heterostructure, temperature, density of surface states, and rate of surface recombination on the buffer layer surface near which the holes move. It can reach several centimeters in magnitude.

2. Let us consider a GaAs/AlGaAs-type heterostructure illuminated with a beam of light which is absorbed mainly in the buffer layer consisting of a narrow-gap material. The geometry and energy diagram of the structure are shown in Fig. 1. The electrons and holes generated by the light in the buffer layer are separated by the built-in electric field of the heterojunction, as shown in Fig. 1. The electrons enter the 2DEG layer and the holes are compressed by the field against the opposite surface of the buffer layer, which is ordinarily separated from the substrate by a superlattice layer. In this manner, a nonequilibrium potential difference $V(r)$, which depends on the coordinate r in the plane of the heterostructure ($V \to 0$ as $r \to \infty$), arises between the edges of the buffer layer. This gives rise to a photoinduced electric field that possesses a component in the plane of the heterostructure. This tangential field E_t is maximum near the back surface of the buffer layer and minimum in the 2DEG layer because of the high conductivity of the layer. The field E_i causes the holes to drift along the back surface in a direction away from the light spot. If the conductivity of the 2DEG is sufficiently high, hole transport is accompanied by a change in the surface charge density of the 2DEG that locally compensates for the nonequilibrium hole charge. In other words, together with the holes, the charge of their mirror image is also transported. In this case, the lateral transport of the electrons and holes is limited only by recombination, which is impeded by the presence of a potential barrier which increases away from the light spot. Of course, in-plane charge carrier transport arises not only because of the drift in the photoinduced tangential field, but also because of diffusion. The ratio of the drift and diffusion fluxes depends on the density Δp of the nonequilibrium holes, since the drift flux depends nonlinearly on Δp (roughly speaking, quadratically) because the tangential field increases with Δp. For this reason, for a high light intensity (which under real conditions is quite weak in absolute magnitude) the drift flux predomi-

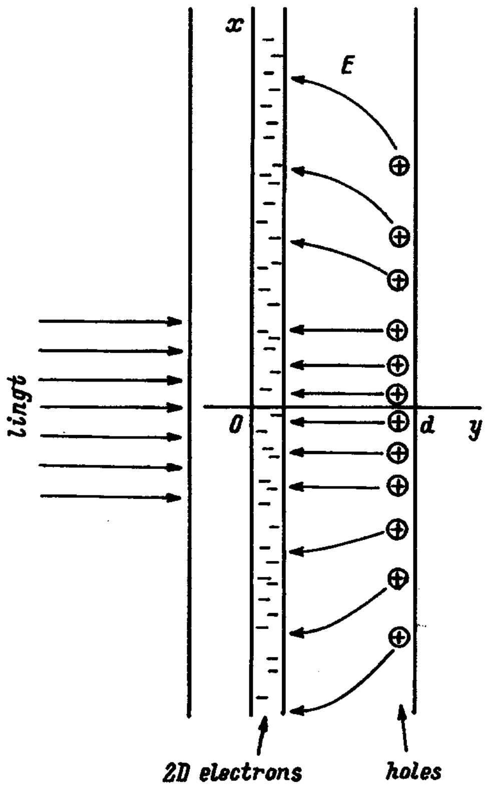

FIG. 1. Geometry of the structure and the charge distribution in it under illumination and the flux lines of the photoinduced electric field.

nates over the diffusion flux. Diffusion becomes dominant at low density Δp.

2.1. A complete quantitative description of the effect is a difficult problem, which includes the equations governing the transport of the photoinduced charge carriers in and across the buffer layer, the Poisson equation for the electric field, and the equation governing the transport of electrons in the 2DEG layer, taking into account, in general, the changes in the shape of the well that localizes the electrons. In the present paper we employ a simplified model, which makes it possible to demonstrate lateral transport under close to experimental conditions and to establish its basic laws.

We assume that the conductivity of the 2DEG is high, so that under illumination of the structure the layer of two-dimensional (2D) electrons remains an equipotential. We also take into account the fact that the thickness d of the buffer layer is small compared with the characteristic lateral transport distances. In this case, it is sufficient to follow only the hole transport, since the electrons are transported in the 2DEG layer and locally compensate for the nonequilibrium hole charge. The disturbance Δn_s of the 2DEG density can be assumed small, even if the disturbance of the hole density is large. Indeed, the nonequilibrium hole density should be assumed large, if it is comparable to the impurity density in the buffer layer. Ordinarily, these are acceptors with density $N_A \simeq 10^{14}-10^{15}$ cm^{-3}. If $\Delta p \simeq N_A$ near the back surface of the buffer layer, where the hole density is highest, then the total number of nonequilibrium holes per unit area of the buffer layer is $\Delta p_s \sim N_A l_D$, where l_D is the distance at which the holes are concentrated at the back boundary of the layer, i.e., the Debye length, which for the indicated densities is of the order of 10^{-5} cm, so that Δp_s is of the order of 10^9-10^{10} cm^{-3}. Because of the electrical neutrality, $\Delta n_s = \Delta p_s$; i.e., Δn_s is much less than the typical 2DEG density, which is of the order of 10^{12} cm^{-2}. In the calculation we assume that the light intensity is moderately high, so that $\Delta p < N_A$ and the perturbation of the 2DEG density is negligible. The photoinduced voltage can then be much greater than $k_B T/e$ (k_B is Boltzmann's constant, and T is the temperature).

A further simplification of the problem involves the thickness d of the buffer layer. In reality, the layer is 0.5–1 μm thick. This is much greater than the thickness of the 2DEG layer but much less than the lateral transport distances of interest to us. Since light absorption and the distribution of nonequilibrium electrons and holes by the built-in field of the heterojunction occurs over the entire thickness d, in studying carrier transport in a direction transverse to the buffer layer we ignore the thickness of the 2DEG layer and treat it as an equipotential surface on which surface recombination occurs: Holes recombine with 2D electrons directly and via surface states. The fact that we disregard the thickness of the 2DEG layer also means that the electron and hole tunneling lengths under the heterojunction barrier are small compared with d. This makes it possible to use local densities to describe carrier transport, though in so doing the interband recombination coefficient can depend on the electric field as a result of tunneling effects. However, this effect is small for the degree of modulation of the electric field considered in the present work. Nonetheless, for generality, we draw a distinction between the interband recombination coefficient B in the volume of the buffer layer, where the field is weak, and the hole–2D-electron interband recombination coefficient B_s.

The fact that the thickness of the buffer layer is small compared with the characteristic lateral transport distances makes it possible to separate the ''fast'' motion of charge carriers in a direction transverse to the layer and the ''slow'' motion in the plane of the layer. In addition, the standard quasiequilibrium approximation can be used to describe carrier transport in a direction transverse to the layer (i.e., it can be assumed that the Fermi quasilevels of the electrons and holes do not depend on the transverse coordinate). The calculations show that this approximation holds well (since there is no straight-through current flowing in the layer), but only at sufficiently high temperatures. The quasiequilibrium condition breaks down at low temperatures.

Our analysis is therefore restricted to fairly high temperatures and low light intensities. Specific estimates are presented below.

2.2. For simplicity, we shall examine a one-dimensional geometry of lateral carrier transport in the x direction in the plane of the structure. This corresponds to a light beam in the form of a strip perpendicular to this direction.

In the stationary state, hole transport in the buffer layer is described by the equation

$$\frac{1}{e}\left(\frac{\partial j_{px}}{\partial x}+\frac{\partial j_{py}}{\partial y}\right)=G(x,y)-B(np-n_i^2). \qquad (1)$$

Here it is assumed that interband transitions are the dominant recombination mechanism. The following notation is adopted: n_i is the intrinsic carrier density, $G(x,y)$ is the rate of optical generation of carriers

$$G(x,y)=\frac{\alpha I_0(x)}{h\nu}e^{-\alpha y},$$

I_0 is the light intensity, $h\nu$ is the photon energy, and α is the absorption coefficient of light. The hole current density $\mathbf{j}_p$ is a sum of the drift and diffusion currents. At the boundaries of the buffer layer j_{py} is determined by surface recombination. At the boundary with the 2DEG

$$(1/e)j_{py}(x,y\approx 0)=-B_s n_s[p(x,0)-p_0(x,0)]$$
$$-S_1[p(x,0)-p_0(x,0)],$$

where B_s is the hole–2D-electron recombination coefficient, S_1 is the rate of recombination via surface states, and $p_0(x,\,0)$ is the equilibrium hole density at the boundary with the 2DEG. The hole current at the opposite boundary of the buffer layer ($y=d$) is determined by recombination via surface states (for simplicity, monoenergetic)

$$(1/e)j_{py}(x,d)=c_p n_t p(x,d)-c_p(N_l-n_l)p_1, \tag{2}$$

where N_t is the density of surface states, n_e is the electron occupancy of the surface states, c_p is the hole trapping coefficient, $p_1\approx N_\nu\ \exp(-\varepsilon_l/k_BT)$ is the Shockley–Read factor for thermal activation of holes into the valence band, ε_l is the energy of traps above the top of the valence band, and N_ν is the effective density of states in the valence band. Assuming for simplicity that under nonequilibrium conditions the surface states are almost completely filled with holes, we rewrite Eq. (2) in the form

$$(1/e)j_{py}(x,d)=S_2\,\frac{np-n_i^2}{p+p_T},$$

where $S_2\ =c_n n_t$ is the surface recombination rate, $p_T=p_1+n_1c_n/c_p$, and c_n and n_1 are the trapping coefficient and Shockley–Read factor for electrons.

The equation (1) can be simplified, if it is assumed that the desired quantities (the carrier density and the field) vary much more rapidly in a direction perpendicular to the layer than along the layer. In this case it can be assumed that a quasiequilibrium state is established in a direction perpendicular to the buffer layer, i.e., the hole and electron Fermi quasilevels do not depend on y. We can then express the hole density in terms of the potential $\phi(x,\,y)$ in the buffer layer, measured from the plane $y=0$,

$$p(x,y)=p(x,d)\exp\{e[\phi(x,d)-\phi(x,y)]/(k_BT)\}, \tag{3}$$

where $p(x,\ d)$ is the hole density at the bottom (back) boundary of the buffer layer $y=d$. Since the electron gas is nondegenerate in virtually the entire buffer layer, we have for the electron density the analogous expression

$$n(x,y)=n(x,0)\exp\{e\phi(x,y)/(k_BT)\} \tag{4}$$

with the difference that the density $n(x,\,0)$ is low and is virtually independent of x. It can be assumed that $n(x,\,0)\approx N_c$, where N_c is the effective density of states in the conduction band. In this approach to the description of $n(x,\,y)$, we neglect the photoinduced potential difference on the 2DEG layer, which is completely justified, since our analysis is limited by low light intensities.

The equation governing hole transport in the plane of the layer can be obtained by integrating Eq. (1) over dy from $y=0$ to $y=d$ and using Eqs. (3) and (4). For the quantities averaged over the thickness of the layer we therefore have

$$\frac{d}{dx}\left(\mu_p\overline{E_x p}-D_p\frac{d\overline{p}}{dx}\right)=\overline{G(x)}-\frac{B}{d}\int_0^d dy(np-n_i^2)$$
$$-\frac{1}{ed}[j_{py}(x,d)-j_{py}(x,0)],$$

where μ_p and D_p are the mobility and diffusion coefficient of the holes. The density $\overline{p(x)}$ averaged over the layer can be expressed in terms of the hole density at the back surface of the buffer layer $p(x,\ d)\equiv p(x)$

$$\overline{p(x)}=\beta p(x),$$

where β depends on the distribution of the potential over the coordinate y near the back surface of the buffer layer. For simplicity, we assume that the spatial charge of the heterojunction is distributed over the entire buffer layer, whose thickness is much greater than the Debye length $l_D=\sqrt{\epsilon k_BT/(4\pi e^2N_A)}$, where ϵ is the permittivity. Then the space charge density near the back boundary equals approximately $-eN_A$ and the solution of Poisson's equation, together with Eq. (3) gives $\beta\approx l_D/d$. The quantity $\overline{E_x p}$ can be replaced by $E_x\overline{p}$, interpreting E_x as the field near the back surface of the buffer layer where $p(y)$ is largest. We find the field E_x as the derivative $-dV/dx$, using the formula for a flat capacitor, one plate of which is the 2DEG and the other is the layer of thickness of the order of l_D at the back surface where the photoinduced positive charge is concentrated, to find $V(x)$. Taking into account the charge due to the free carriers and the carriers localized in surface states, we obtain

$$V(x)\approx\frac{4\pi ed}{\epsilon}\left[\beta p(x)d+N_t\frac{p(x)}{p(x)+p_T}\right]. \tag{5}$$

The potential $\phi(x,d)$ appearing in Eq. (3) is $V_0-V(x)$, where V_0 is the height of the barrier without illumination (see Fig. 2).

As a result, we obtain the following equation for $p(x)$ in the dimensionless form:

$$\frac{d}{d\xi}\left\{\left[1+\tilde{p}\left(1+\frac{N_g}{(\tilde{p}+g)^2}\right)\right]\frac{d\tilde{p}}{d\xi}\right\}$$
$$=-\tilde{G}(\xi)+\left(1+\frac{\sigma}{\tilde{p}+g}\right)\tilde{p}\ \exp\left[\tilde{p}+\frac{N\tilde{p}}{\tilde{p}+g}\right], \tag{6}$$

where $\tilde{p}=p/p^*$, $p^*=N_Al_D^2/(\beta d^2)$, $N=N_l/(\beta dp^*)$, $g=p_T/p^*$, $\xi=x/l$, $l=\sqrt{\beta D_p\tau}$,

$$\sigma=\frac{S_2N_c}{p^*(BN_cd+B_sn_s+S_1)},$$

$$\tau=\frac{d\ \exp(eV_0/k_BT)}{BN_cd+B_sn_s+S_1},\quad \tilde{G}=\frac{I_0(x)\tau}{h\nu p^*d}(1-e^{-\alpha d}).$$

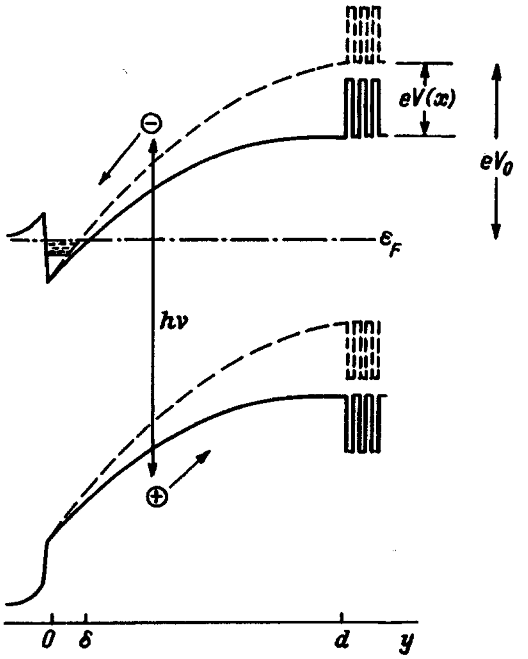

FIG. 2. Energy diagram of a heterostructure under illumination (solid line) and far from the light beam (dashed line).

The equation (6) is presented in a simplified form for the case where the density $p(x)$ is much greater than the thermodynamically equilibrium hole density $p_0(d)$ at the back boundary of the buffer layer.

3. The nonequilibrium hole density distribution in the plane of the heterostructure, obtained by solving Eq. (6) numerically for GaAs at temperature 200 K, is presented in Fig. 3. The figure also shows for comparison the light intensity distribution employed in the calculation and it displays the distribution $p(x)$ in the case where hole drift in the photoinduced field is neglected. As one can see, in-plane carrier transport is connected with both drift and diffusion of the holes.

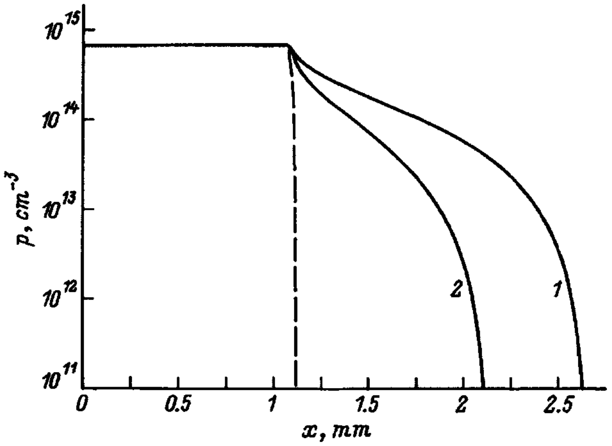

FIG. 3. Nonequilibrium carrier density distribution p in the plane of the heterostructure (*1* — total lateral transport effect and *2* — transport due to diffusion). Dashed line — light intensity distribution. Parameters: $I_0=10^{-1}$ W/cm^2, $T=200$ K, $S_1=S_2=10$ cm/s, $N_t=10^9$ cm^{-3}, $V_0=0.225$ V.

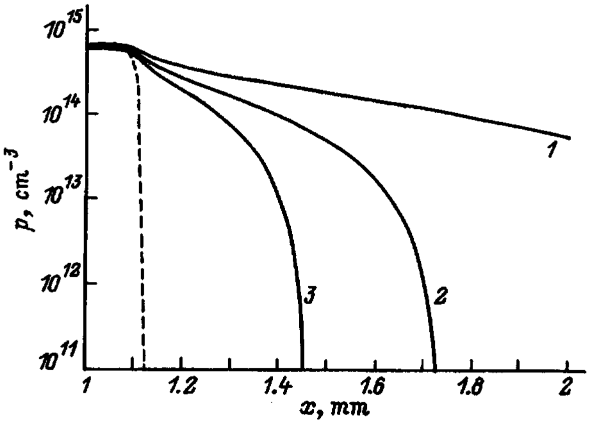

FIG. 4. Nonequilibrium hole density distribution p at temperature T, K: *1* — 200, *2* — 250, *3* — 300. The parameters I_0, S_1,S_2, and N_t are the same as in Fig. 3. Dashed line — light intensity distribution.

The contribution of the drift mechanism as compared with the diffusion mechanism is determined by the quantity $p[1+Ng(\tilde{p}+g)^{-2}]$ on the left-hand side of Eq. (6). This quantity depends on both the nonequilibrium carrier density and the density of surface states and their energy position. As one can see, this quantity can reach large values as $\tilde{p}$ increases. In the limit of low densities p, i.e., for low light intensity or far from the light beam, drift transport becomes unimportant and the asymptotic behavior is reached in the limit $x\rightarrow\infty$: $p\simeq\exp[(-x/e)(1+\sigma/g)^{1/2}]$. Therefore, the characteristic length is of the order of $l(1+\sigma/g)^{-1/2}$. If $\tilde{p}$ is not small (in Fig. 3, as the light beam is approached), the dependence of p on x is determined by the change in the barrier height $e[V_0-V(x)]$, on which the effective lifetime depends, and by the increase in the role of drift.

As the temperature decreases, carrier transport in the plane of the heterostructure strongly intensifies as a result of a decrease in the probability of recombination of electrons and holes separated by the barrier. The distribution of $p(x)$ and of the photoinduced voltage are shown in Figs. 4 and 5 for different temperatures for the case in which the level of the surface states lies 0.5 eV above the valence band. The

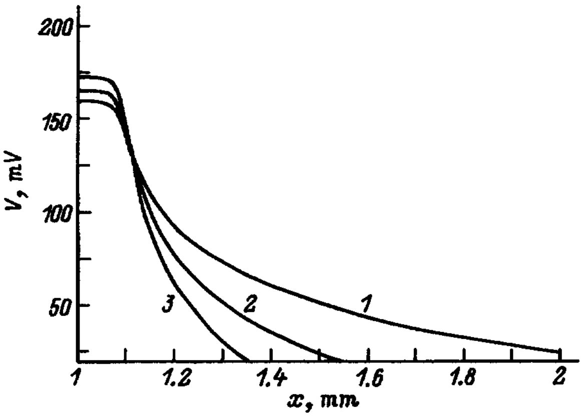

FIG. 5. Distribution of the photoinduced voltage V. The labels *1–3* and the computational parameters are the same as in Fig. 4.

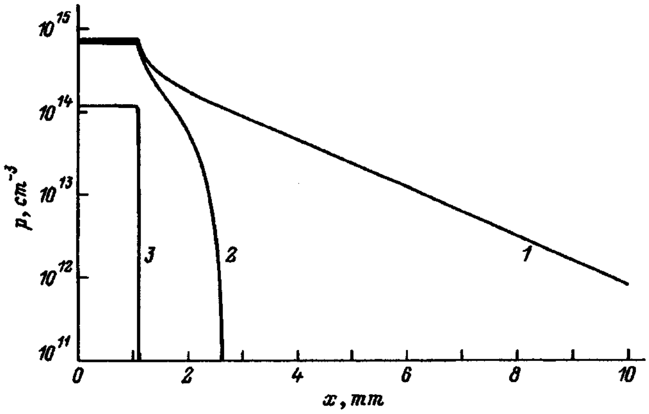

FIG. 6. Nonequilibrium hole density distribution p for different densities of surface states and surface recombination rates on the back surface of the buffer layer with $T=200$ K, $\varepsilon_t=500$ meV, $I_0=10^{-1}$ W/cm^{-2}. *1* — No surface state; the values of the parameters N_t, cm^{-2}, and $S_1=S_2$, cm/s: *2* — 10^9 and 10, 3 — 10^{10} and 100.

lifetime of the photoinduced carriers will then be 4.8, 1.07, and 0.44 ms for $T=200$, 250, and 300 K, respectively.

The effect of the density of surface states and their position in energy on lateral carrier transport is determined by the charge accumulated on them, which decreases the height of the recombination barrier and increases the tangential field, and directly by recombination via surface states. Increasing N_l, ε_l, and the surface recombination rate S_2 decreases the transport length (Figs. 6 and 7). The surface recombination rate S_1 at the heteroboundary with the 2DEG has a much weaker influence on the effect because at this boundary the recombination of holes with the 2DEG is quite strong.

The following values were used in all numerical calculations: $N_A=5\times10^{14}$ cm^{-3}, $d=0.8$ μm, $\alpha=10^4$ cm^{-1}, $B_s=B=2\times10^{-10}(300/T)^{3/2}$ cm^3/s, and $n_s=10^{12}$ cm^{-2}. The carrier mobility and effective density of states were calculated according to Ref. 8. The barrier height eV_0 in the equilibrium state was calculated taking into account the charge of the free electrons and the ionized acceptors (for $p\ll N_A$), assuming that the space charge region occupies the entire buffer layer $(d<l_D(2eV_c/k_BT)^{1/2}$, V_c is the contact potential difference between the strongly doped wide-gap semiconductor and the volume of the buffer semiconductor, provided that it is thick):

$$\frac{eV_0}{k_BT}\approx\frac{1}{2}\left(\frac{d}{l_D}-\sqrt{2}\right)^2+\ln\frac{N_c}{N_A}.$$

As one can see, the height of the recombination barrier depends on the thickness and the doping of the buffer layer. Increasing V_0, for example, by increasing d, results in a strong increase in the lateral transport length.

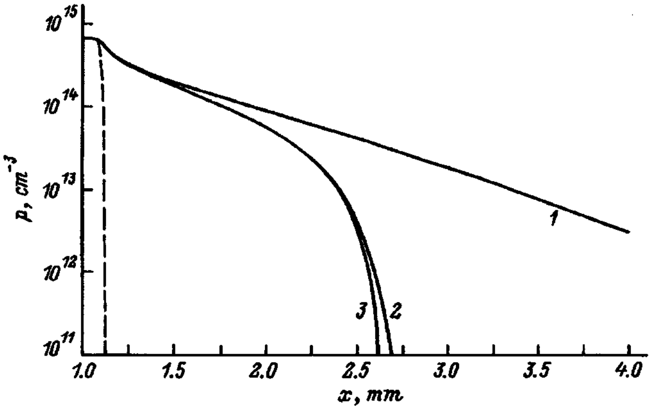

FIG. 7. Nonequilibrium hole density distribution p for different energies ε_t of surface states above the bottom of the valence band. $N_t=10^9$ cm^{-2}, $S_1=S_2=10$ cm/s, $T=200$ K, $I_0=10^{-1}$ W/cm^2; ε_t, meV: *1* — 200, *2* — 300, *3* — 500. Dashed line — light intensity distribution.

4. In the model examined above, the lateral transport length is an exponential function of the quantity eV_0/k_BT, which characterizes the separation of the photoinduced electrons and holes. However, a more detailed analysis shows that this result is valid only if eV_0/k_BT does not exceed a value which depends on the light intensity. The point is that increasing eV_0/k_BT, for example, by decreasing the temperature, at fixed light intensity results in breakdown of quasiequilibrium in the carrier density distribution in a direction transverse to the buffer layer, which made it possible to obtain Eq. (6) without the potential $\phi(x, y)$. To show this and clarify the consequences of a breakdown of quasiequilibrium, let us examine the carrier density distribution in a direction transverse to the buffer layer under uniform illumination. The electron density distribution satisfies the equation

$$\frac{d}{dy}\left(\mu_n n\frac{d\phi}{dy}+D_n\frac{dn}{dy}\right)=R(n,p)-G(y), \tag{7}$$

where μ_n and D_n are the electron mobility and electron diffusion coefficient, and $R(n, p)$ is the recombination rate. Integrating Eq. (7) over y, we obtain

$$n=\left\{n(0)+\frac{1}{D_n}\int_0^y dy' e^{-\frac{e\phi(y')}{k_BT}}\int_0^{y'}dy''[R(n,p)-G(y'')]\right\}e^{\frac{e\phi(y)}{k_BT}}. \tag{8}$$

Here the second term in braces is the deviation from a quasiequilibrium distribution. If it is estimated by using the quasiequilibrium distributions n and p, then the relative contribution of the deviation from quasiequilibrium increases with y, and near the back boundary of the buffer layer it is

$$\frac{\delta n}{n}\simeq\frac{I_0 l_D}{h\nu n(0)D_n d}\exp(eV_B/k_BT), \tag{9}$$

where $V_B=V_0-V$ is the barrier height in the presence of illumination, and V is the photoinduced potential difference, which can be estimated from Eq. (5). The estimate (9) was obtained for the case in which there is no surface recombination. As one can see, as eV_0/k_BT increases, the deviation from quasiequilibrium grows exponentially, and the result of the breakdown of quasiequilibrium is that the density n decreases much more slowly with increasing y. This is shown quantitatively in Fig. 8, which illustrates the results of the

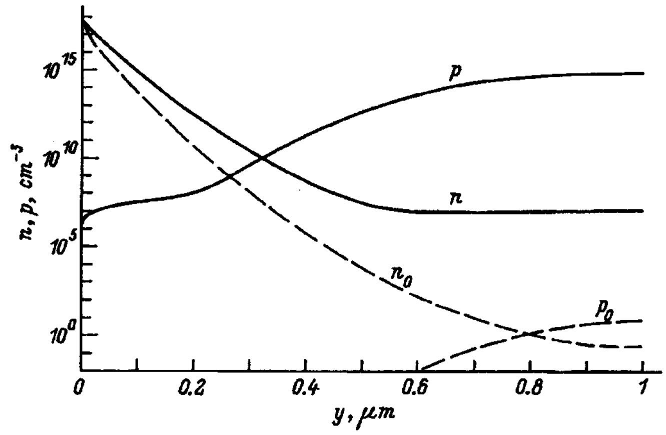

FIG. 8. Electron density distribution n and hole density distribution p in a direction transverse to the buffer layer under uniform illumination. The distribution was obtained by solving simultaneously the carrier transport equations and the Poisson equation without using the assumption of quasiequilibrium. $I_0 = 10^{-3}$ W/cm^{-2}, $N_A = 10^{15}$ cm^{-3}, $T = 200$ K. Dashed lines — nonequilibrium density distributions n_0 and p_0.

numerical calculation of $n(y)$ and $p(y)$, obtained without assuming quasiequilibrium for conditions under which quasiequilibrium breaks down.

Therefore, the model examined in the present paper is valid at sufficiently high temperatures and low light intensities, when $\delta n/n \ll 1$. In addition, the lower the light intensity, the lower the temperatures at which quasiequilibrium breaks down. For example, the estimate (9), calculated into taking account the fact that the photovoltage V depends on I_0, shows that for $I_0 = 10^{-1}$ W/cm^2 quasiequilibrium breaks down at $T \approx 195$ K, and if $I_0 = 10^{-3}$ W/cm^2, then quasiequilibrium breaks down at $T \approx 130$ K (in these calculations it was assumed that $N_A = 5 \times 10^{14}$ cm^{-3} and $d = 0.8$ μm). This conclusion can be generalized by taking into account the surface recombination. This gives more complex formulas, but the qualitative results remain the same.

Let us now examine how the breakdown of quasiequilibrium influences lateral transport. As one can see from Fig. 8, as a result of the breakdown of quasiequilibrium, the electron density increases strongly near the back surface of the buffer layer as compared with the quasiequilibrium distribution. Similarly, the hole density increases near the front surface of the buffer. As a result, the recombination rate increases strongly near the two boundaries of the buffer layer, and this sets a limit on the lateral transport. A more detailed investigation of lateral transport under such conditions will be made in a separate work.

5. In the present paper we have accordingly proposed a mechanism of lateral transport of photoinduced charge carriers in heterostructures with a 2DEG and we have developed a model of this effect which is valid at sufficiently low temperatures and low light intensities, which are close to the conditions under which lateral transport is observed experimentally. The lateral transport intensifies at lower temperatures, but a more complicated calculation is required in order to describe it quantitatively for experimentally acceptable light intensities.

The effect studied is similar to the phenomenon of frozen photoconductivity,[9] and it is especially close to the experiments performed by S. M. Ryvkin and D. V. Tarkhin.[10] Lateral transport of nonequilibrium charge carriers in the surface layers of semiconductors has been known since the 1970s.[11] In Ref. 12 it was investigated in connection with the determination of the maximum achievable spatial contrast of photoinduced etching of semiconductors. In Ref. 13 the lateral transport effect was used to explain the characteristic features of the photoconductivity spectra of Si/GaAs heterojunctions.

The lateral transport in selectively doped heterostructures can be investigated experimentally by performing measurements of the photoreflection response (as in Refs. 7 and 14) with spatial separation of the exciting and probe beams or according to the photoluminescence, also measured locally with a displacement relative to the exciting beam.[14] In this case, the photoreflection response is determined mainly by the electric field in the buffer layer, i.e., the nonequilibrium charge of the free carriers and the carriers localized in surface states, and the photoluminescence makes it possible to follow the nonequilibrium density of free charge carriers. To use this effect for the purpose of contact-free diagnostics of transistor structures (of the NEMT type), additional investigations of the dependence of the transport distance on the conductivity of the 2DEG must be performed.

This work was supported by the Russian Ministry of Science as part of the program ''Physics of solid-state nanostructures'' (Project No. 1-037), the International Science Foundation (Grant RL-7300), and the Russian Fund for Fundamental Research (Grant 96-02-18476a).

1)Electronic mail: vas199@ire216.msk.su

---

[1] G. D. Gillard, D. J. Woklford, T. F. Kuech, and J. A. Bradley, Appl. Phys. Lett. **59**, 216 (1991).

[2] G. D. Gillard, D. J. Wolford, T. F. Kuech *et al.*, J. Appl. Phys. **73**, 8386 (1993).

[3] Z. W. Wang, J. Windscheif, D. J. As, and W. Jantz, J. Appl. Phys. **73**, 1430 (1993).

[4] Q. X. Zhao, B. Monemar, P. O. Holtz *et al.*, Phys. Rev. B **50**, 7514 (1994).

[5] I. A. Avrutskiĭ, O. P. Osaulenko, V. G. Plotchenko, and Yu. N. Pyrkov, Fiz. Tekh. Poluprovodn. **26**, 1907 (1992) [Sov. Phys. Semicond. **26**, 1069 (1992)].

[6] J. S. Massa, G. S. Buller, A. C. Wölker *et al.*, Appl. Phys. Lett. **66**, 1346 (1995).

[7] M. Sydor, J. R. Engholm, M. O. Manasreh *et al.*, Phys. Rev. B **45**, 13 796 (1992).

[8] J. S. Blakemore, J. Appl. Phys. **53**, R123 (1982).

[9] M. K. Sheĭknman and A. Ya. Shik, Fiz. Tekh. Poluprovodn. **10**, 209 (1976) [Sov. Phys. Semicond. **10**, 128 (1976)].

[10] S. M. Ryvkin and D. V. Tarkhin, Fiz. Tekh. Poluprovodn. **7**, 1447 (1973) [Sov. Phys. Semicond. **7**, 970 (1973)].

[11] J. E. Carnes, W. F. Kosonocky, and E. F. Ramberg, IEEE Trans. Electron. Dev. **ED-19**, 798 (1972).

[12] L. V. Belyakov, D. N. Goryachev, S. M. Ryvkin, O. M. Sreseli, and R. A. Suris, Fiz. Tekh. Poluprovodn. **13**, 2173 (1979) [Sov. Phys. Semicond. **13**, 1270 (1979)].

[13] A. G. Denisov, G. S. Dordzhin, Yu. G. Sadof'ev, L. V. Sharonova, A. Ya. Shik, and Yu. V. Shmartsev, Fiz. Tekh. Poluprovodn. **16**, 2152 (1982) [Sov. Phys. Semicond. **16**, 1386 (1982)].

[14] O. A. Ryabushkin, V. A. Sablikov, V. G. Mokerov, and Yu. V. Fyodorov in *Abstracts of Int. Symp. ''Nanostructures: Physics and Technology-95'*, St. Petersturg, Russia (1995) p. 52.

Translated by M. E. Alferieff

The electrical conductivity of polycrystalline SnO_2(Cu) films and their sensitivity to hydrogen sulfide

B. A. Akimov,[1)] A. V. Albul, A. M. Gas'kov, V. Yu. Il'in, M. N. Rumyantseva,
and L. I. Ryabova

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

M. Labeau

National Polytechnical Institute, Grenoble, France

(Submitted February 26, 1996; accepted for publication May 22, 1996)
Fiz. Tekh. Poluprovodn. **31**, 400–404 (April 1997)

The effect of doping with copper on the sensor properties and the electrical conductivity of polycrystalline SnO_2(Cu) films has been investigated. It has been found that at room temperature the residual conductivity is observed after the films are exposed to H_2S. This made it possible to determine the character of the low-temperature conductivity of the films for different degrees of saturation with hydrogen sulfide. A comparison of the obtained data with the results of layerwise elemental analysis suggested a model that explains the mechanism of the gas sensitivity of SnO_2(Cu) to hydrogen sulfide. In contrast to the mechanisms, which are associated with the work done by the surface and which are standard for gas sensors, in the present case the change in the conductivity is due to the chemical reaction of the electrically active copper with sulfur in the entire volume of the film. This reaction determines the selectivity and high sensitivity of SnO_2(Cu) to H_2S. *© 1997 American Institute of Physics.*
[S1063-7826(97)00404-3]

Tin dioxide films are widely used as relatively simple resistance sensors for detecting toxic gases in the atmosphere.[1] However, such gas sensors are quite complicated structures from the standpoint of investigating the mechanism of the processes which are induced by a change in the composition of the gas medium. The sensitive element of such a sensor consists of a polycrystalline semiconductor film. The most important characteristics of the structure as a whole (selectivity, speed of operation, sensitivity, stability, and others) are determined not only by the film thickness, grain size, and type of dopants, but also by the substrate material with transitional layers, as well as contacts which contribute to the change in the electrical conductivity.

One way to modify the properties of a sensor is to dope its sensitive element — a polycrystalline film. As determined in Refs. 2–4, for two-phase ceramic samples $SnO_2+5\%$ CuO and somewhat later for uniphase doped polycrystalline SnO_2(Cu) films,[5] copper increases the sensitivity to hydrogen sulfide, while the sensitivity to other gases, such as CO or ethyl alcohol vapors, is suppressed. In the case of ceramics, the increase in sensitivity is attributed to the removal of p-CuO–n-SnO_2 interblock barriers as a result of formation of a phase of high-conductivity copper sulfide.[3,4] In the uniphase doped films, one of the factors which determine the increase in sensitivity could be a substantial increase in the resistance R of the initial SnO_2 films accompanying the introduction of copper. On the other hand, the selective sensitivity with respect to hydrogen sulfide is evidently a consequence of the special chemical mechanisms of the interaction of doped films with a gaseous medium.

Our objective in the present work was to study the mechanisms of the gas sensitivity and selectivity of SnO_2(Cu) films with respect to H_2S, which is of general interest. We intended to clarify the possibility of applying low-temperature methods, ordinarily used to investigate disordered and polycrystalline semiconductors, to sensor structures, which, as a rule, operate at high temperatures (up to 500–600 K). The implementation of our schemes depended on the extent to which it is possible to saturate a film with hydrogen sulfide at the working temperatures and maintain such a quasiequilibrium state when the structure is cooled. Solid-state memory effects of this kind make it possible to investigate objects in which, by analogy to doping, the content of the sorbed elements can be varied continuously. The results presented below show the extent to which this can be done for SnO_2(Cu).

Since in very high-resistivity SnO_2(Cu) films, the substrate and contacts can have a large effect on the conduction process,[6] we studied films with the same copper content but deposited on different substrates. In choosing the synthesis conditions, we allowed for the fact that the optimal level of doping with copper, from the standpoint of increasing the gas sensitivity, equals 1 at. % Cu.[5]

The investigations of the temperature dependences of the conductivity of structures with different degrees of exposure to H_2S were supplemented by an analysis of the elemental composition of the films.

1. MEASUREMENT METHOD. EXPERIMENTAL SAMPLES

Films with a dopant were synthesized by the method of aerosol pyrolysis.[7,8] A 0.20-M solution of tin dibutyldiacetate in acetylacetone and a 0.05-M solution of copper trifluoroacetylacetonate in butanol were used as precursors. These solutions were mixed in the ratio [Cu]/([Cu]+[Sn]) = 1%. Dried air was used as the carrier gas. Films with

TABLE I. Characteristics of the experimental films.

Sample No.	Cu concentration, at. %	Substrate type	R, $\Omega(T=273$ K)	E_a, meV	Film thickness, μm
1	-	Si+SiO_2	5.4×10^4	-	1.5
2	-	″″	6.9×10^3	-	1.5
3	1	″″	8.8×106	130	1
4	1	Si+Ta_2O_5 (Amorphous)	2.2×10^6	130	1.3
5	1	MgO	1.1×10^7	130	1.5
6	1	Si+Ta_2O_5 (Crystalline)	1.4×10^7	130	1.5

thickness ranging from 1 to 1.5 μm were deposited on a $\langle 100\rangle$ single-crystal Si or MgO substrates. A JSM-35 (JEOL) scanning electron microscope was used to determine the thickness of the films on sheared surfaces. When Si was used as the substrate material, an insulating transitional layer was first formed on its surface. To this end, either the substrate was oxidized in air at a temperature of 1000 °C for 24 h or crystalline or amorphous Ta_2O_5 was deposited. The substrate temperature was maintained constant at $T=773$ K during the growth process. The phase composition and microstructure of the films were studied by the method of x-ray diffraction on a Siemens diffractometer. Only the SnO_2 phase (cassiterite) was found; a copper-containing phase was not found in the films. The elemental composition of the films was determined by local x-ray spectral analysis (Cameca-SX50). The average SnO_2 grain sizes, 5–7 nm, were estimated from the broadenings of the x-ray diffraction reflections.

A quantitative analysis of the composition of the main components and dopant, as well as layerwise analysis over the thickness of the films were performed by secondary-mass spectrometry of neutral particles (SNMS) on a INA-3 spectrometer (Leybold) with a resolution of 5 nm in thickness and analyzed surface area of 4 mm^2.

The measurements of the electrical resistance of the films were performed in the temperature range 77–523 K in a regime with a stabilized voltage $U=1$ V applied to the current contacts of the film.

The films had an area of 4×5 mm^2. Special attention was given to the contacts. The procedural requirements were that the contacts had to operate at high and low temperatures. Four types of contacts were investigated in order to work out the procedures: contacts deposited with the aid of silver paste directly on the film surface; contacts deposited with the same paste on gold deposited beforehand on the contact areas; indium contacts (more accurately, the alloy 95 wt. % In + 4 wt. % Ag + 1 wt. % Au) soldered on the film surface; and indium contacts soldered on the gold layer. The measurements of the current-voltage characteristics and temperature dependences of the conductivity showed that the data obtained for films with contacts of the second type differ substantially from the results for films with contacts obtained for the three other types. This is evidently due to the fact that according to the data of Ref. 9, gold produces a Schottky barrier to n-SnO_2, and soldering of such wires with indium does not damage the gold layer. In this case, contact phenomena make a large contribution to the conductivity of the films with the highest resistivity. The increase, observed in Ref. 5, in the slope of the low-temperature curves $\log R=f(1/T)$ for SnO_2(Cu) samples with increasing copper doping level is often due to the increase in the Schottky barrier on the gold contacts under conditions when the Fermi level decreases as a result of an increase in the degree of compensation. It should be noted that gold contacts are preferable from the standpoint of obtaining the highest gas sensitivity. However, ohmic contacts were required in order to study the processes occurring in the film itself. In this connection, we performed all measurements, with the exception of specially stipulated cases, using ohmic contacts obtained by depositing silver paste. After the paste was deposited, the samples were heated to a temperature of 423 K. Apparently, the heating procedure caused the gold layer to dissolve, and for this reason the data obtained with silver contacts deposited on a gold interlayer and directly on the film are identical.

A special metal chamber, which screened the sample virtually completely from background radiations, was used to perform measurements at temperatures below room temperature. A different chamber was used to determine the gas sensitivity at temperatures of 290–550 K. The sample was mounted on a sapphire substrate, which was secured to a foamed quartz support containing a nichrome wire heater. The chamber could be filled with air or a mixture of air with hydrogen sulfide. This mixture could be produced in a separate vessel and blown through the chamber or produced directly in the chamber by placing in it a Teflon cell containing liquid hydrogen sulfide. The permeability of the cell walls to H_2S is limited. The measurements were performed with H_2S partial pressures up to 1000 ppm.

2. EXPERIMENTAL RESULTS

The temperature dependences of the relative resistance R/R_{273} of the films, where R_{273} is the resistance at $T=273$ K (see Table I), at temperatures $T<300$ K are presented in Fig. 1. For the doped films 3–6, the curves $\log(R/R_{273})=f(1/T)$ virtually coincide (the curves for samples 4 and 5 are shown in the figure), irrespective of the type of substrate. For the film 3, the curve obtained with gold contacts is presented for comparison.

A preliminary estimate of the gas sensitivity $S=(\sigma_G-\sigma_{air})/\sigma_{air}$, where σ_G is the conductivity of the film in the presence of hydrogen sulfide, and σ_{air} is the conductivity in air, was made with the aid of a series of measure-

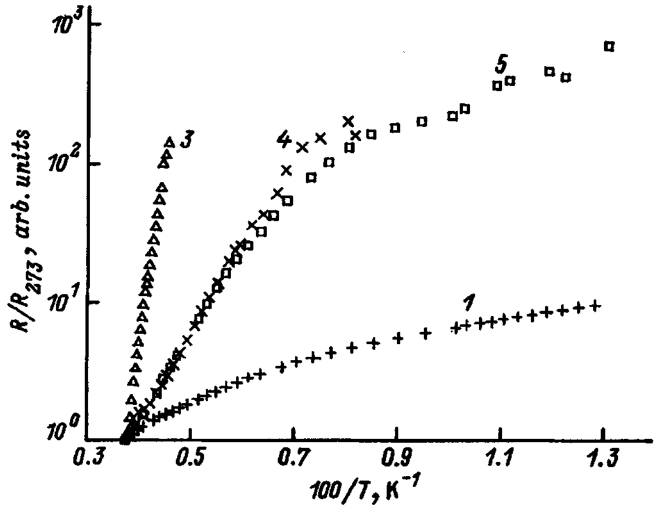

FIG. 1. Temperature dependences of the relative resistance R/R_{273}. The numbers on the curves correspond to the number of the sample in Table I.

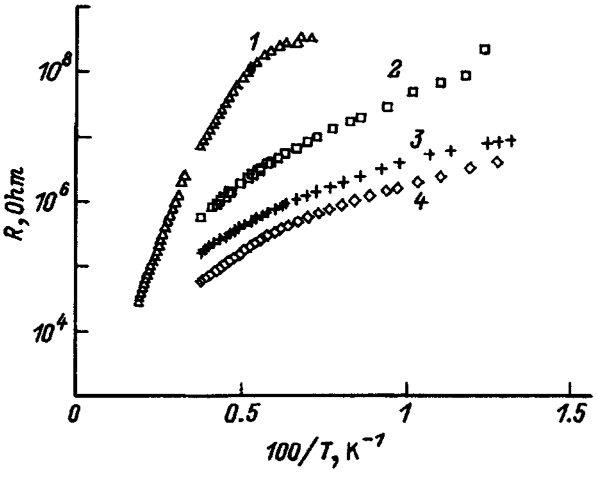

FIG. 3. Temperature dependences of the resistance of film 3 before (*1*) and after (*2, 3*) exposure to a mixture of air and hydrogen sulfide and also for the undoped film 1 (*4*).

ments which are illustrated in Fig. 2. First, the conductivity σ was measured for a sample heated from room temperature to 520 K and then cooled in air (transition between the points *1→2* in Fig. 2). The sample was then placed in a hydrogen sulfide atmosphere at $T=300$ K; in the process the conductivity of the film increased somewhat (point *3*). Next, the sample was heated to the same temperature (transition *3→4*) in the presence of H_2S. The sample was then again placed in air (point 6), heated (transition *6→7*), and cooled (transition *7→8*). The cycles of the first heating and last cooling in air after measurements in the presence of hydrogen sulfide are reproduced. However, all other dependences remain nonequilibrium dependences, since the conductivity varies very slowly under the action of hydrogen sulfide. The time for one heating–cooling cycle was equal to about 40 min. The data for sample 6, in which a maximum difference between the values of σ_{air} and σ_G was observed, are shown in Fig. 2. For all other films, including undoped films, the qualitative change of the conductivity in the cycle described is similar.

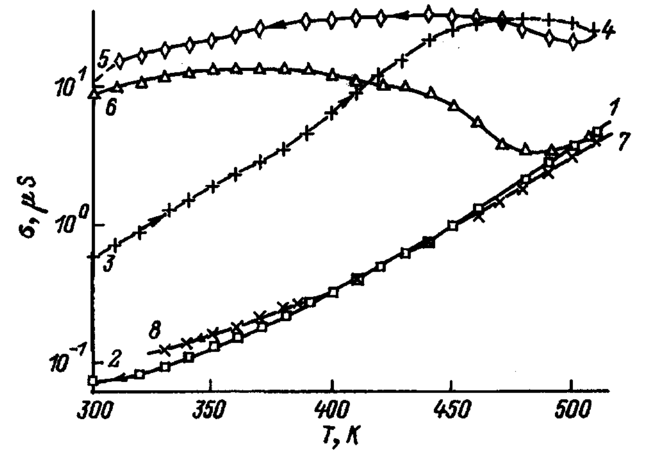

FIG. 2. Variation of the conductivity in successive heating and cooling cycles of film 6 in air and in a mixture of air and hydrogen sulfide (see explanation in the text).

However, the maximum values of S, which can be formally determined from the nonequilibrium data, are found to be more than an order of magnitude lower for the undoped films than for the doped films. It should be noted that the residual conductivity after the films are exposed in an active gaseous medium with a SnO_2-based sensitive element is by no means observed in all cases.[10]

The long-time relaxation of the conductivity to the initial value for $SnO_2(Cu)$ films after exposure in a mixture of air and H_2S makes it possible to monitor the change in the low-temperature dependences $R(T)$ with changing degree of saturation of the sample with hydrogen sulfide. Figure 3 shows the temperature dependences of the resistance for film 3, which were measured before (curve *1*) and after (curves *2* and *3*) the film was exposed in H_2S. Curve *2* was obtained after a heating–cooling cycle in the presence of H_2S in the temperature range from 300 to 500 K and subsequent holding of the film in air at room temperature for one day. Curve *3* was obtained immediately after the samples were held for 35 min in the presence of hydrogen sulfide at room temperature. During the heating–cooling cycling of the films, the curves $R(T)$, which gradually approach curve *2* in form, are obtained immediately after the curve *3* is measured. For comparison, Fig. 3 also shows the temperature dependence of R for the undoped film 1 (curve *4*).

The kinetics of the variation of the sensitivity S in time t for the doped film 6 and the undoped film 2 at different temperatures are shown in Figs. 4 and 5, respectively. During the first 10 min, the sample was held in a mixture of air and hydrogen sulfide, after which the atmosphere was changed to air. We see from Fig. 4 that the kinetics of the process for the $SnO_2(Cu)$ film 6 changes substantially with the temperature. For example, the increase in sensitivity is slowest at $T=373$ K, and as relaxation to the equilibrium state in air occurs, $S(t)$ asymptotically approaches a value much different from zero: The residual conductivity phenomenon is observed. An increase in temperature initially not only accelerates the relaxation process, but also increases the sensitivity at each fixed moment in time. At $T\simeq450$ K

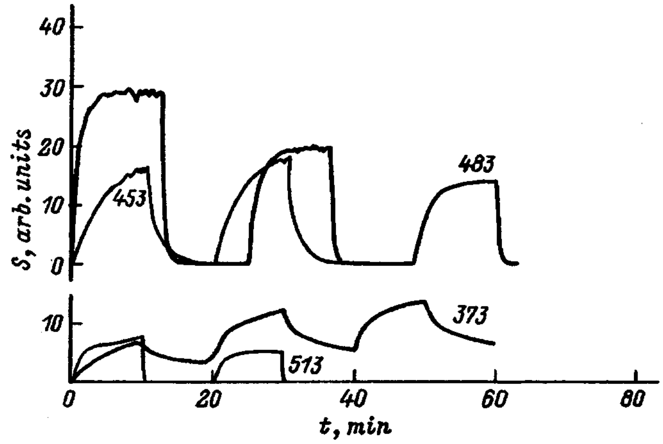

FIG. 4. Kinetics of the variation of the sensitivity S of the $SnO_2(Cu)$ film 6 with exposure to a mixture of air and hydrogen sulfide and to air. The values of the temperature are given in K.

the data for a fixed cycling process are highly reproducible, and the conductivity relaxation becomes nearly exponential. At $T = 483$ K a saturation section is observed in the curve $S(t)$; S reaches its maximum value compared with data for other temperatures. But, in contrast to the kinetic dependences for σ at lower temperatures, the maximum value of σ gradually decreases from cycle to cycle. Each next increase in temperature decreases the sensitivity of the film. The sensitivity is more than an order-of-magnitude lower in an undoped film, and its value decreases monotonically with increasing temperature.

3. DISCUSSION

Analysis of the low-temperature curves $\log R = f(1/T)$ for the films investigated shows that in contrast to the undoped samples, a linear section characterized by an activation energy $E_a = 130$ meV is observed in all doped films with silver contacts. The relation $R \sim \exp(E_a/kT)$ was used to determine E_a. The value of E_a is virtually identical in order of magnitude to the energy level of an oxygen vacancy.[9] In contrast to the type of contact, the type of substrate has no appreciable effect on the character of the temperature dependences of the resistance. Therefore, the introduction of copper into SnO_2 in amounts of ~ 1 at. % evidently accounts for the compensation, which allows us to see the activation of charge carriers from the level of an oxygen vacancy.

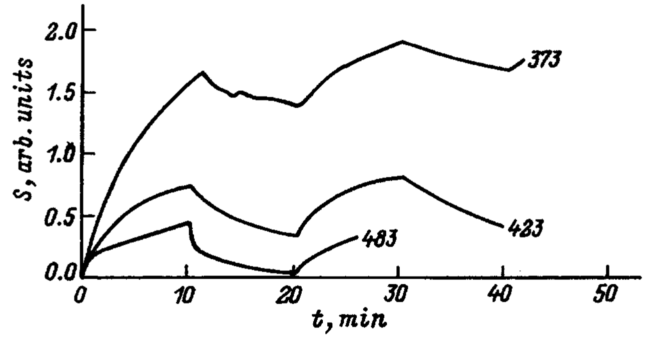

FIG. 5. Kinetics of the variation of SnO_2 film 2 and with exposure to a mixture of air and hydrogen sulfide and to air. The temperature values are given in K.

Exposure to hydrogen sulfide substantially changes the character of the low-temperature dependences of the resistance of doped films. It is justified to attribute these changes to a change in the composition of the experimental films after exposure to H_2S. As determined in Ref. 5, in the case of layerwise SNMS analysis of the same films the copper is distributed uniformly over the volume of the film. After holding in hydrogen sulfide the presence of quite uniformly distributed sulfur, whose concentration increases with increasing exposure time to H_2S, is observed in the entire volume of the copper-doped films. It is important to note that in undoped SnO_2, after holding in an atmosphere containing hydrogen sulfide, the sulfur is observed only at the surface. Therefore, it can be assumed that the presence of sulfur in $SnO_2(Cu)$ is due to the chemical interaction of sulfur with copper accompanied by the formation of a compound whose exact chemical composition cannot yet be identified. This assumption is confirmed by the fact that the temperature dependences of the resistance of the films after exposure to hydrogen sulfide approach the dependences characteristic of undoped SnO_2; i.e., the concentration of electrically active copper in the case of an interaction with hydrogen sulfide decreases and the film is decompensated.

The fact that in $SnO_2(Cu)$ films the sensitivity to hydrogen sulfide is due to processes occurring in the entire volume of the film and in undoped SnO_2 films the sensitivity is due only to processes occurring at the surface of the film could lead to a large difference in the value of S for the structure studied. An attempt to calculate the instantaneous relaxation time $\tau = \Delta\sigma/|\partial\sigma/\partial t|$ in the temperature range 423–453 K, where the kinetic curves are closest to being exponential, was made in order to make quantitative assessments of the kinetic processes. However, even in this temperature range, the calculation of τ is complicated by the fact that to determine $\Delta\sigma = \sigma - \sigma_0$ for the dropoff curve and $\Delta\sigma = \sigma - \sigma_{\max}$ for the growth curve of S in an atmosphere containing hydrogen sulfide, it is necessary to know accurately the initial and final values of the conductivities σ_0 and $\sigma_{\max}$. Even a comparatively small variation of these values results in a large change in the computed values of τ. Nonetheless, it is possible to judge from the data obtained the order of magnitude of τ (5–20 min) and to show that even for kinetic dependences $S(t)$ closest to exponential, τ is not a constant throughout the entire time interval of variation of S. Comparing the characteristic features of the gas sensitivity with respect to H_2S for undoped and copper-doped SnO_2 films suggests that in the first case the change in the conductivity is characteristic of the entire volume of the film, while in the second case all processes proceed on the surface.

1)E-mail: akimov@mig.phys.msu.su

---

[1] W. Gopel, Progr. Surf. Sci. **20**, 9 (1985).
[2] K. Takahata in *Chemical Sensor Technology*, edited by T. Seyama, Elsevier, N. Y., **84**, 441 (1988).
[3] T. Maekawa, J. Tamaki, N. Miura, and N. Yamazoe, Chem. Lett. **4**, 575 (1991).

[4] J. Tamaki, T. Maekawa, N. Miura, and N. Yamazoe, Sensors Actuators **9**, 197 (1992).
[5] A. M. Gas'kov, L. I. Ryabova, M. Labo, Zh. Delabugliz, M. N. Rumyantseva, T. A. Kuznetsov, and M. N. Bulova, Neorg. Mater. **41**, 989 (1996).
[6] K. D. Schierbaum, U. K. Kirner, J. F. Geiger, and W. Gopel, Sensors and Actuators B **4**, 87 (1991).
[7] M. Labeau, B. Gautheron, F. Celier, M. Vallet-Regi, E. Garcia, and J. C. Gonzales-Calbert, J. Solid State Chem. **102**, 379 (1993).
[8] M. Labeau, P. Rey, and J. L. Deschanvres, J. C. Joubert, G. Delabouglise, Thin Solid Films **213**, 94 (1992).
[9] C. J. Fonstad and R. H. Rediker, J. Appl. Phys. **42**, 2911 (1971).
[10] R. M. Voshchilova, D. P. Dimitrov, N. I. Dolotov, A. R. Kuz'min, A. V. Makhin, V. A. Moshnikov, and Yu. M. Tairov, Fiz. Tekh. Poluprovodn. **29**, 1987 (1995) [Semiconductors **29**, 1036 (1995)].

Translated by M. E. Alferieff

Characteristic features of the accumulation of vacancy- and interstitial-type radiation defects in dislocation-free silicon with different oxygen contents

I. I. Kolkovskiĭ

Belorussian State University, 220064 Minsk, Belarus

V. V. Luk'yanitsa

Belorussian State Medical Institute, 220116 Minsk, Belarus

(Submitted March 20, 1996; accepted for publication May 22, 1996)
Fiz. Tekh. Poluprovodn. **31**, 405–409 (April 1997)

The formation of the main radiation defects in silicon (*A* and *E* centers, $C_i - C_s$ and $C_i - O_i$ complexes) in dislocation-free crystals and crystals with a low dislocation density ($N_D \simeq 1\times10^4$ cm^{-2}) have been investigated as a function of the oxygen density N_O. The characteristic features of the accumulation and annealing of radiation defects observed in dislocation-free silicon are interpreted taking into account the presence of interstitial inclusions in the volume of such crystals. It has been determined that the gettering properties of the inclusions depend in a complicated manner on the oxygen concentration and are most obvious when $N_O \simeq 3\times10^{16}$ cm^{-3}. © *1997 American Institute of Physics.* [S1063-7826(97)00504-8]

1. INTRODUCTION

The efficiency of the formation of radiation defects (RDs) in silicon depends on the content of not only dopants, but also background impurities, primarily oxygen, which is a constituent of RDs of the vacancy type (*A* centers) and the interstitial type (complexes $C_i - O_i$), in the crystals.[1–3] In addition to oxygen, the effect of not previously monitored factors on the RD accumulation has been observed recently.[4–7] Analysis of available data shows that in order to gain a deeper understanding of radiation defect formation processes in silicon, it is necessary to study the effect of structural imperfections, which appear as a result of the absence of growth dislocations and the presence of oxygen impurity at definite concentrations.

We have accordingly performed a comparative study of the processes leading to the accumulation of the main RDs in dislocation-free and low-dislocation-density silicon crystals as a function of the oxygen content in the crystals.

2. EXPERIMENTAL PROCEDURE

The study was performed on *n*- and *p*-type silicon single crystals with no dislocations and with a low dislocation density ($N_D \simeq 1\times10^4$ cm^{-2}) with resistivity $\rho \simeq 100\ \Omega\cdot$cm, grown by zone melting method in vacuum or an argon atmosphere, as well as by pulling from melt following Czochralski. The oxygen and carbon background impurity contents were determined from infrared (IR) absorption on ~5-mm-thick plates cut from different parts of the ingots. The degree of structural perfection of the crystals was checked by the metallographic method after selective etching of the surface. Samples with dimensions of 12×2×2.5 mm with approximately the same carbon density ($N_C \simeq 2\times10^{16}$ cm^{-3}) and different oxygen contents ($N_O \leqslant 1\times10^{16} - 1\times10^{18}$ cm^{-3}) were employed for the experiments. The samples were irradiated with ^{60}Co γ rays at temperature $T_{\text{irr}} \simeq 50$ °C and 15-min isochronous annealing was conducted in the temperature range $T_{\text{ann}} = 50 - 450$ °C. The experimental results were obtained from measurements performed at different stages of irradiation and subsequent annealing of the temperature dependences of the conductivity, Hall coefficient, and charge carrier lifetime, measured by the method of conductivity modulation in a point contact. The temperatures dependences of the Hall coefficient were analyzed according to the electroneutrality equation by the least-squares method or differential method.[8] The RD formation rates ($\eta = N/\Phi$) were found on the linear sections of the curves of the defect density N versus the irradiation flux Φ by averaging the data on η obtained on five samples.

3. EXPERIMENTAL RESULTS AND DISCUSSION

It was determined from an analysis of the temperature dependences of the Hall coefficient of the irradiated samples and data on the annealing of the defects formed that in all investigated crystals irradiation produces both vacancy- and interstitial-type radiation defects. Vacancies of the first type are *A* and *E* centers, which contribute levels $E_c - 0.18$ eV and $E_c - 0.42$ eV, respectively, in the band gap and the interstitial-type RDs are the complexes ⟨interstitial carbon⟩–⟨site carbon⟩ ($C_i - C_s$, levels $E_c - 0.16$ eV and $\sim E_v + 0.1$ eV) and ⟨interstitial carbon⟩–⟨interstitial oxygen⟩ ($C_i - O_i$, level $\sim E_v + 0.35$ eV).[1–3] Under the conditions of our experiments the defects mentioned above form in appreciable quantities and the other defects form in negligible small quantities.

Figure 1 shows the rate of formation of vacancy-type (Fig. 1a) and interstitial-type (Fig. 1b) RDs versus the oxygen concentration in crystals with no dislocations and with a low dislocation density. One can see that in all crystals studied the rates of formation of *A* centers and $C_i - O_i$ complexes tend to increase with increasing oxygen concentration and at the same time the rates of formation of *E* centers and $C_i - C_s$complexes decreases. A detailed analysis of the data presented makes it possible to identify the following character-

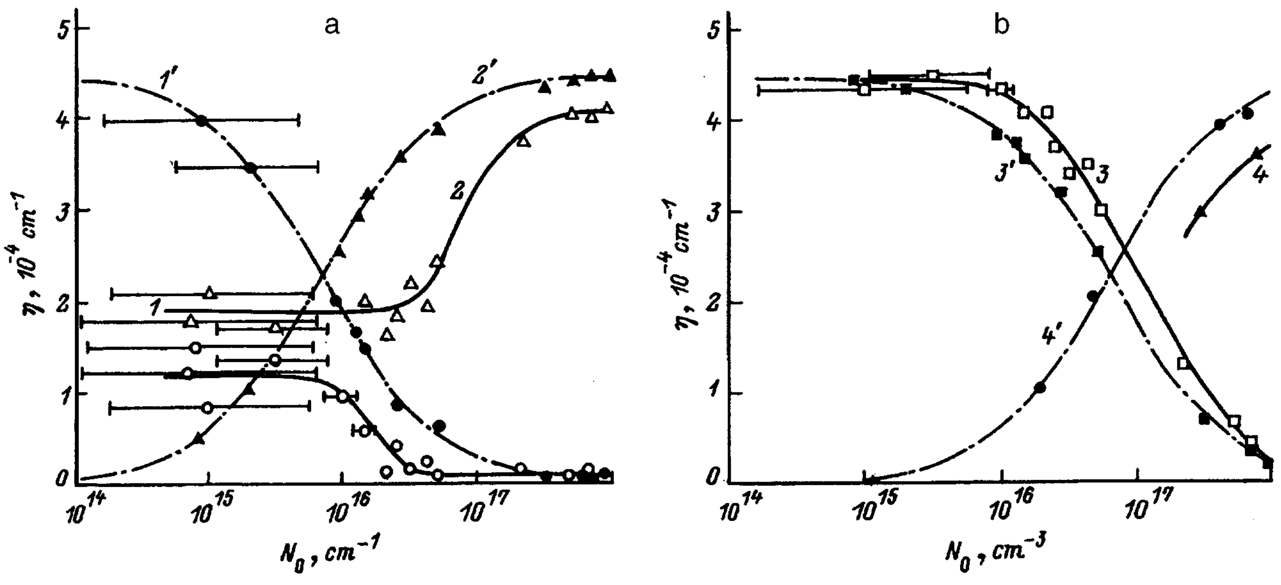

FIG. 1. Formation rates of radiation defects of the vacancy (a) and interstitial (b) types versus oxygen density. Crystals: *1–4* — dislocation free, *1′–4′* — low-dislocation density. Defects: *1, 1′* — *E* centers, *2, 2′* — *A* centers, *3, 3′* — C_i-C_scomplexes, *4, 4′* — C_i-O_i complexes.

istic features of the accumulation of RDs in the dislocation-free material:

1) The values of η are different (especially for low densities N_O) for vacancy RDs in dislocation-free silicon and in silicon with a low density of dislocations;

2) the variance in the values of η from ingot to ingot in the dislocation-free crystals is large compared with the low-dislocation-density crystals;

3) the total rate of formation of RDs does not remain constant in the dislocation-free crystal as the oxygen density changes.

The observed behavior of the dependence $\eta = f(N_O)$ in the materials investigated can be explained qualitatively on the basis of the theory of quasichemical reactions between primary RDs (vacancies and interstitial atoms) generated in the volume of the crystal and uniformly distributed impurities (oxygen, phosphorus, carbon). Since the main sinks for vacancies are phosphorus and oxygen atoms, the relative formation efficiency of *A* and *E* centers is proportional to the ratio of the oxygen and phosphorus densities. For this reason, as the oxygen density in the crystal increases, *A* centers will be introduced in greater numbers and the efficiency of introduction of *E* centers will decrease.

The situation is similar for the interstitial channel of defect formation, when the radiation-generated characteristic interstitial atoms (I) displace carbon atoms from the lattice sites ($C_s + I \rightarrow C_i$). Since O_i also acts as a sink for C_i, together with C_s,[3] the relative efficiency of formation of C_i-O_i and C_i-C_s complexes is determined by the competition with respect to the capture of C_i between interstitial oxygen atoms and site carbon: $O_i + C_i \rightarrow C_i - O_i$, $C_s + C_i \rightarrow C_i - C_s$. As a result, as N_O increases, the efficiency of formation of C_i-O_i complexes increases and that of C_i-C_s complexes, conversely, decreases.

In the quantitative analysis, the experimental data (Fig. 1) were compared with the results of calculations performed on the basis of the system of equations presented below, describing in the stationary approximation the accumulation of *A* and *E* centers at the initial stages of irradiation:[9]

$$\begin{cases} N_A = \sigma_V N_{\mathrm{Si}} \Phi \varkappa_{VO} N_{\mathrm{O}} / (\varkappa_{VO} N_{\mathrm{O}} + \varkappa_{V\mathrm{P}} N_{\mathrm{P}}), \\ N_E = \sigma_V N_{\mathrm{Si}} \Phi \varkappa_{V\mathrm{P}} N_{\mathrm{P}} / (\varkappa_{VO} N_{\mathrm{O}} + \varkappa_{V\mathrm{P}} N_{\mathrm{P}}), \\ N_V = \sigma_V N_{\mathrm{Si}} \Phi / (\varkappa_{VO} N_{\mathrm{O}} + \varkappa_{V\mathrm{P}} N_{\mathrm{P}}), \end{cases} \tag{1}$$

where σ_V is the effective cross section for the formation of vacancies and silicon as a result of γ irradiation, N_{Si} and N_{P} are the densities of silicon and phosphorus atoms, N_A and N_E are the densities of *A* and *E* centers, N_V is the stationary density of vacancies, $\varkappa_{VO}$ and $\varkappa_{V\mathrm{P}}$ are the vacancy–oxygen and vacancy–phosphorous interaction constants, and Φ is the γ-ray flux.

A similar system of equations, in which σ_V, N_V, N_{P}, $\varkappa_{VO}$, and $\varkappa_{V\mathrm{P}}$ were replaced by σ_{C_i}, N_{C_s}, N_{C_i}, $\varkappa_{\mathrm{C}_i\mathrm{O}_i}$, and $\varkappa_{\mathrm{C}_i\mathrm{C}_s}$, respectively (the types of defects are indicated in the subscripts), was used to describe the accumulation of interstitial-type RDs (C_i-O_i; and C_i-C_s).

It was found that the results of calculation (dot-dashed lines) are in satisfactory agreement with the experiment in a wide range of variation of the oxygen concentration for the following values of the phenomenological parameters $\sigma_V = 9\times 10^{-26}$ cm^{-2}, $\varkappa_{V\mathrm{P}}/\varkappa_{VO} = 135$, and $\varkappa_{\mathrm{C}_i\mathrm{O}_i}/\varkappa_{\mathrm{C}_i\mathrm{C}_s} = 3$ and only for slightly dislocated crystals.

We note in this connection that the method, which is well known in the literature[10] and employed in practice, for determining low densities of oxygen (not observed by IR absorption) in *n*-silicon crystals based on measurements of the formation rates of *A* and *E* centers is, in general, correct only for low-dislocation-density crystals. It has not been ruled out, however, that the application of this method without regard for the structural perfection of the crystal can give oxygen densities which are too low [at the level ($\leq 10^{13}$ cm^{-3} (Ref. 11)]. In the different low-dislocation-density crystals which we investigated and which were grown by the method of zone melting in vacuum, the oxygen density determined from the measured values of the formation rate η_E of E centers using as a calibration curve $\eta_E = f(N_{\mathrm{O}})$ calculated with the parameters indicated above was in the range $N_{\mathrm{O}} = 8\times10^{14} - 5\times10^{15}\ \mathrm{cm}^{-3}$, which agrees with the data of Refs. 12 and 13.

A similar approach can also be used to determine low oxygen densities in crystals.

In the case of dislocation-free crystals it is impossible to describe satisfactorily in the model considered here the change in the RD formation rates in the experimental range of oxygen densities. This probably stems from the assumption made in the calculations that the impurity distribution is uniform and (or) that only point trapping centers (oxygen, phosphorus, and carbon) of primary RDs are present. The fact that in the experimental dislocation-free crystals the charge carrier mobility at liquid-nitrogen temperatures is, as a rule, lower than in the low-dislocation-density crystals indicates the presence of complicated structural imperfections in the bulk. These imperfections, as suggested in Refs. 7 and 14, impurity–defect clusters (IDCs) which form during growth of dislocation-free crystals, are not removed by selective etching, and consist of interstitial-type inclusions surrounded by an atmosphere of background impurities. Taking into account this circumstance, the features observed above in the accumulation of RDs in dislocation-free crystals with different oxygen density can be explained on the basis of model-based ideas developed in Refs. 7 and 14–16.

Under the action of the deformation fields produced by the clusters, the irradiation-induced free vacancies migrate toward the clusters. The effect of vacancy redistribution between the crystal matrix and interstitial-type IDCs is determined by the gettering properties of the latter. As the oxygen content in the crystal increases, this effect will weaken, since the oxygen atoms, which are effective sinks for vacancies, decrease their mean-free path length. The presence of IDCs is therefore most obvious in dislocation-free crystals with a low oxygen concentration ($N_{\mathrm{O}} < 10^{15}\ \mathrm{cm}^{-3}$) and it gives rise to a sharp decrease in the rate of formation of E centers in the crystal matrix. In contrast, vacancy-type defects accumulate more efficiently in the region of IDCs. Since the local density of oxygen atoms in IDCs is high, the predominant formation is that of A centers. Their formation rate η_A, just as in the matrix, is determined by the ratio of the oxygen density and the density of other vacancy sinks localized in the IDCs (phosphorus atoms, other impurities, and inclusion–matrix interfaces, where annihilation and ''freezing'' of trapped vacancies are possible[16]). As a result, it is quite sensitive to the structure and impurity composition of IDCs.

Thus, the first of the above-noted features is due to the redistribution of free vacancies between the crystal matrix and IDCs. As a result, the ratio η_E/η_A in dislocation-free crystals changes compared with the low-dislocation density crystals.

Changing the conditions of crystal growth (growth rate, temperature gradient on the axis, growth atmosphere, etc.) changes the density of the IDCs which are formed, changes their gettering properties, and changes the composition and density of the impurity atmosphere. For this reason, the variance of η_E, η_A, and η_E/η_A (second feature), which is observed in the dislocation-free crystals, is a consequence of the differences in their structural perfection as a result of the fact that the growth conditions are not identical.

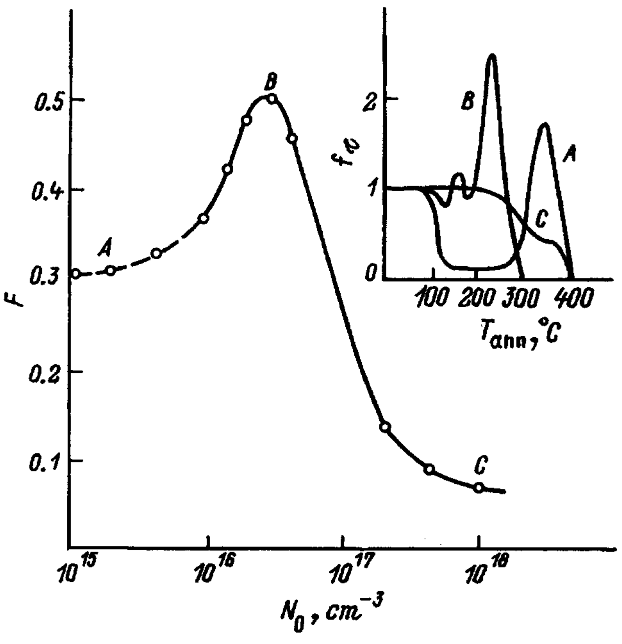

FIG. 2. Fraction F of ''unreacted'' vacancies versus oxygen density. Inset: Annealing of recombination centers.

Since the elastic stress fields produced by the inclusions are anisotropic, a qualitatively similar picture is also possible for interstitial-type defects.

The third of the above-noted features, which is connected with the fact that the total rate of formation of RDs in dislocation-free crystals does not remain constant, attests to the fact that in IDCs not all vacancies participate in the formation of electrically active defects. The N_{O}-dependence of the parameter $F = (\eta^c - \eta^e)/\eta^c$ (where $\eta^c = \eta_A^c + \eta_E^c$ and $\eta^e = \eta_A^e + \eta_E^e$ are, respectively, the computed and experimentally determined total rates of formation of vacancy-type RDs with fixed oxygen density), which reflects the fraction of ''unreacted'' (trapped IDCs) vacancies, is shown in Fig. 2. As one can see, as N_{O} increases from $10^{15}\ \mathrm{cm}^{-3}$ to $10^{18}\ \mathrm{cm}^{-3}$, F at first increases, reaching a maximum value for $N_{\mathrm{O}} = 3\times10^{16}\ \mathrm{cm}^{-3}$ (point B), and then decreases, becoming at $N_{\mathrm{O}} = 10^{18}\ \mathrm{cm}^{-3}$ three times smaller (point C) than at the initial oxygen density $N_{\mathrm{O}} = 10^{15}\ \mathrm{cm}^{-3}$ (point A).

The most likely reason for the increase in F with increasing oxygen density in the interval from $N_{\mathrm{O}} \leqslant 10^{15}\ \mathrm{cm}^{-3}$ to $N_{\mathrm{O}} \simeq 3\times10^{16}\ \mathrm{cm}^{-3}$ is an increase in the volume and (or) density of IDCs, which formed during crystallization, and thereby also the inclusion–crystal matrix interface area, which increases the capability of trapping of free vacancies. Moreover, with an increase in N_{O}, so-called nuclei, whose interaction with the vacancies results in the formation of electrically inactive complexes, can form near the IDCs.

For oxygen densities $N_O > 3 \times 10^{17}$ cm^{-3}, the gradient of the magnitude of the elastic fields which are produced by IDCs in the crystal matrix decreases as a result of the accommodation of clusters and the presence of high densities of point sources (oxygen atoms) whose elastic stress fields are of the same sign as those of IDCs in the crystal matrix. As a result, the effect of vacancy redistribution between the crystal matrix and the IDCs becomes weaker and therefore decreases for vacancies trapped by IDCs.

The data on the annealing of recombination centers in irradiated dislocation-free material (see the inset in Fig. 2, the curves A, B, and C correspond to the points A, B, and C on the curve of F versus N_O) attest to the effect of oxygen on the gettering properties of IDCs. It is obvious that the change in the fraction of unannealed recombination centers $f_\tau = (\tau^{-1} - \tau_0^{-1})/(\tau_\Phi^{-1} - \tau_0^{-1})$ (τ_0, τ_ϕ, and τ are the lifetimes of the nonequilibrium charge carries before irradiation, after irradiation, and at different stages of annealing) with increasing annealed temperature T_{ann} depends on N_O and correlates with the dependence $F = f(N_O)$. The curve A is similar to the curve of f_τ versus T_{ann} for the vacuum, low-dislocation-density, n-silicon,[14] with the exception of the peak near $T_{\text{ann}} = 300 - 400$ °C, which, in our opinion, indicates the presence of IDCs in the crystals. The curve B is characteristic of crystals with oxygen density $N_O \simeq 3 \times 10^{16}$ cm^{-3}, where the gettering properties of IDCs are most pronounced. The extremely nonmonotonic form of this curve attests to the fact that the fraction of vacancies trapped by IDCs, all of which by no means participate near IDCs in the formation of A centers, increases; some of them are ''frozen'' in the composition of less-stable electrically neutral complexes or enter a quasistable state on the increasing inclusion–⟨crystal matrix⟩ boundary. Freeing of vacancies by some method starts at $T_{\text{ann}} > 100$ °C and is accompanied by additional formation (maxima in f_τ) and subsequent decomposition (minima in f_τ) of a number of new recombination-active defects.

Curve C shows that although IDCs are present in the crystal, their gettering properties are weaker than in the preceding case (curve C is more even and passes below the peaks on curves B and A).

In conclusion, we note that the results which we obtained give us a reason to believe that one of the main factors responsible for the variance of the ratio $\varkappa_{VP}/\varkappa_{VO}$ in the range from 10 to 250, observed by different authors,[4,10,17–19] is the presence in dislocation-free crystals, in contrast to low-dislocation-density crystals, of impurity-defect clusters, whose gettering properties are determined by the growth conditions.

4. CONCLUSIONS

Accordingly, the observed characteristic features of the formation of RDs in dislocation-free silicon with different oxygen contents are due to the presence of interstitial-type impurity–defect clusters in the crystals. The elastic fields of these clusters give rise to a redistribution and trapping of the radiation-generated vacancies. It was found that the gettering properties of the impurity-defect clusters depend nonmonotonically on the oxygen density, reaching maximum strength at $N_O \simeq 1 \times 10^{16} - 1 \times 10^{17}$ cm^{-3}.

[1] V. V. Emtsev and T. V. Mashovets, *Impurities and Point Defects in Semiconductors* [in Russian], Moscow (1981).

[2] G. D. Watkins and J. W. Corbett, Phys. Rev. **121**, 1001 (1961).

[3] L. I. Murin, Phys. Status Solidi A **93**, K147 (1986); **101**, K107 (1987).

[4] N. V. Kuznetsov and G. G. Solov'ev, *Radiation Resistance of Silicon* [in Russian], Moscow (1989).

[5] R. F. Vitman, N. A. Vitovskiĭ, A. A. Lebedev, T. V. Mashovets, Fiz. Tekh. Poluprovodn. **24**, 45 (1990) [Sov. Phys. Semicond. **24**, 27 (1990)].

[6] V. V. Emtsev, G. A. Oganesyan, and K. Shmal'ts, Fiz. Tekh. Poluprovodn. **27**, 1549 (1993) [Semiconductors **27**, 856 (1993)].

[7] I. I. Kolkovskii, P. F. Lugakov, V. V. Lukjanitsa, and V. V. Shusha, Phys. Status Solidi A **118**, 65 (1990).

[8] H. I. Hoffman, Appl. Phys. **19**, 307 (1979).

[9] V. V. Vinetskiĭ and G. A. Kholodar', *Radiation Physics of Semiconductors* [in Russian], Kiev (1979).

[10] Z. Su, A. Husain, and J. W. Farmer, J. Appl. Phys. **67**, 1903 (1990).

[11] F. P. Korshunov, V. P. Markevich, I. F. Medvedeva, and L. I. Murin, Fiz. Tekh. Poluprovodn. **26**, 2007 (1992) [Sov. Phys. Semicond. **26**, 1129 (1992)].

[12] Yu. V. Trubitsyn, K. N. Neĭmark, I. F. Chervonyĭ, Vysokochistye veshchestva, No. 3, 129 (1989).

[13] B. O. Kolbsen and A. Muhlbener, Solid–State Electron. **25**, 759 (1982).

[14] I. I. Kolkovskii, P. F. Lugakov, and V. V. Shusha, Phys. Status Solidi A **127**, 103 (1991).

[15] P. F. Lugakov and V. V. Luk'yanitsa, Fiz. Tekh. Poluprovodn. **17**, 1601 (1983) [Sov. Phys. Semicond. **17**, 1019 (1983)].

[16] I. I. Kolkovskiĭ, P. F. Lugakov, and V. V. Shuma, Izv. Vyssh. Uchebn. Zaved., Fiz. **31**, 128 (1988).

[17] I. D. Konozenko, A. K. Semenyuk, and V. I. Khirvich, *Radiation Effects in Silicon* [in Russian], Kiev (1974).

[18] V. S. Vavilov, V. B. Glazman, N. U. Isaev, B. N. Mukashev, Fiz. Tekh. Poluprovodn. **8**, 471 (1974) [Sov. Phys. Semicond. **8**, 303 (1974)].

[19] A. S. Zubrilov and S. V. Koveshnikov, Fiz. Tekh. Poluprovodn. **25**, 1332 (1991) [Sov. Phys. Semicond. **25**, 804 (1991)].

Translated by M. E. Alferieff

Solid solution $In_xGa_{1-x}As_ySb_zP_{1-y-z}$: A new material for infrared optoelectronics. I. Thermodynamic analysis of the conditions for obtaining solid solutions, isoperiodic to InAs and GaSb substrates, by liquid-phase epitaxy

N. A. Charykov, A. M. Litvak, M. P. Mikhaĭlova, K. D. Moiseev, and Yu. P. Yakovlev

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

(Submitted May 14, 1996; accepted for publication May 22, 1996)
Fiz. Tekh. Poluprovodn. **31**, 410–415 (April 1997)

The diagrams of melt–solid solution (fusibility curves) and solid solution (I)–solid solution (II) (surfaces of spinodal decomposition of solid solutions) phase equilibria in the five-component system In–Ga–As–Sb–P (the solid solutions are isoperiodic to the GaSb and InAs substrates) are calculated. The concentration ranges of the isovalent substitution solid solutions $In_xGa_{1-x}As_ySb_zP_{1-y-z}$, which are accessible for synthesis by liquid-phase epitaxy, are calculated. © *1997 American Institute of Physics.* [S1063-7826(97)00604-2]

1. FORMULATION OF THE PROBLEM

InAs- and GaSb-based multicomponent III–V solid solutions, ternary (InAsSb, InGaAs) and quaternary (GaInAsSb, InAsSbP, GaAlAsSb), are now widely used for producing optoelectronic devices (lasers,[1–4] light-emitting diodes[5] and photodiodes[6,7]) for the spectral range 2–5 μm at 300 K. The quaternary solid solutions, which are isoperiodic to InAs and GaSb substrates, have an important advantage over the ternary solid solutions, since they make it possible to vary the band gap of the material while maintaining the lattice period.[8] At the same time, these solutions have the drawback that a large range of compositions cannot be obtained by liquid-phase epitaxy (LPE) because of limitations with respect to the region of immiscibility of the solid solutions (i.e., region of their spinodal decomposition) and the condition of molecularity of the melt.[9] For example, the solid solution $In_xGa_{1-x}As_ySb_{1-y}$ can be grown on a GaSb substrate, which is strictly isoperiodic to the latter, only in the range $0<x<0.29$; the composition is close to GaSb (which corresponds to E_g in the range 0.72–0.50 eV) or to InAs ($1>x>0.74$, $E_g \cong 0.30-0.24$ eV).[10,11] The solid solutions $In_xAs_ySb_zP_{1-x-y}$/InAs, which are widely used as cover layers in long-wavelength lasers, also have a limited composition range accessible for synthesis by LPE $0<1-x-y<0.38$.[12] These fundamental thermodynamic limitations prevent the production of optoelectronic devices for discrete wavelengths in the entire range 2–5 μm on the basis of only one III–V material by the LPE method. The methods of molecular-beam and gas-phase epitaxy, widely employed in the last few years, make it possible to level out appreciably the effect of the two limitations and to expand the accessible composition range of the solid solutions. However, the simplicity and relative low cost of LPE are very attractive for fabricating optoelectronic devices. The need to integrate lasers and photodiodes into GaAs- and Si-based electronic schemes has led to the appearance of an entire series of studies on the growth of multicomponent III–V solutions on these substrates. However, as a result of the large difference in the lattice periods of the GaAs substrate and solid solutions based on GaSb, InAs, InSb, etc., synthesis by the LPE method has still not led to substantial positive results. These drawbacks can be overcome by producing five-component solid solution $In_xGa_{1-x}As_ySb_zP_{1-y-z}$, which makes it possible to grow isoperiodic materials of the same qualitative composition on four substrates immediately: GaSb, InAs, InP, and GaAs. In this paper we report the results of a thermodynamic analysis of the conditions of LPE synthesis of solid solutions isoperiodic to the first two substrates.

Our objective in this study was to determine the following:

1) The regions of existence of solid solutions corresponding to the five-component system, which are isoperiodic to InAs and GaSb substrates and are stable with respect to spinodal decomposition;

2) the regions of possible preparation of solid solutions by the LPE method, bounded by the so-called ‘‘condition of molecularity of the melt,’’ which corresponds to equal total concentrations of III and V components in the melt;[13] and

3) the effective region of one of the above-listed limitations, and thus the composition range which is accessible for synthesis by the LPE method, and the two-phase equilibrium temperatures.

To solve such problems it is first necessary to have a thermodynamic description of the coexisting phases — melt and isovalent substitution solid solution. We shall discuss these phases in greater detail.

Model of the melt. To describe the standard chemical potentials of the components of the melt In, Ga, As, Sb, P and the melted compounds InAs, GaAs, InSb, GaSb, InP, and GaP (as well as the solid compounds — see below), we employed the published data on the standard heats and entropies of formation ΔH^0_{298} and S^0, isobaric specific heats in the liquid and solid states C^l_p and C^s_p, and the melting temperatures and entropies T^f and ΔS^f. These data are presented in Table I. Here it should be noted that the data on the standard thermodynamic functions are completely reliable for all components and compounds (except GaP).

According to the few published data for GaP, mainly on the values of ΔS^f and other functions, the latter differ by amounts of up to 10–20 rel. %.

We employed to the EFLCP model to describe the ex-

TABLE I. Thermodynamic functions of the components and compounds in the system In–Ga–As–Sb–P.

Substance	ΔH^0_{298}, kJ	S^0, J/K	C^l_p, J/K	C^s_p=a+b·T, J/K	T^f, K	ΔS^f, J/K
In	−242.8	57.9	31.4	$24.3+0.0105\cdot T$	430	7.6
Ga	−288.8	42.7	28.2	$23.9+0.0105\cdot T$	303	18.5
As	−288.8	35.2	23.0	$21.9+0.0093\cdot T$	1090	18.4
P	−317.0	25.0	20.7	$19.7+0.0151\cdot T$	870	8.0
Sb	−285.6	45.7	29.8	$23.1+0.0071\cdot T$	903	22.7
InSb	−537.0	86.5	67.2	$50.0+0.0110\cdot T$	797	60.1
InAs	−589.4	75.8	67.0	$44.8+0.0148\cdot T$	1210	60.8
InP	598.0	62.4	62.8	$44.9+0.0169\cdot T$	1333	58.8
GaSb	−596.4	76.4	63.0	$49.0+0.0125\cdot T$	985	66.4
GaAs	−658.6	64.2	67.0	$47.4+0.0070\cdot T$	1513	69.7
GaP	−681.8	52.3	56.1	$43.9+0.0070\cdot T$	1738	74.2

cess thermodynamic functions of the melt components (γ_i — activity coefficients). This model was previously used successfully to describe III–V systems (see, for example, Refs. 14–16).

The basic equation of the EFLCP model for systems with an arbitrary number n of components (in our case $n=5$) has the following form:[16]

$$RT\ \ln\ \gamma_i=\sum_{j\neq i}^{n}\alpha_{ij}x_j^2+\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}(\alpha_{ij}+\alpha_{ik}-\alpha_{jk})x_jx_k$$
$$+(2x_i-1/2)\sum_{j\neq i}^{n}\beta_{ij}x_j^2+\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}x_jx_k$$
$$\times[(x_j-x_k)\beta_{kj}+x_j(\beta_{ij}+\beta_{ik})+(1-2x_j)\Psi_{ijk}. \tag{1}$$

Here R is the gas constant, T is the temperature (in K), x_i is the molar fraction of the ith component of the melt, and α_{ij},β_{ij}, and Ψ_{ijk} are model parameters. The parameter α_{ij} is twice the change in the Gibbs free energy which accompanies the formation of one mole of the liquid compound ij from one mole of the liquid components i and j, without regard for the contribution of the entropy of ideal mixing. It is calculated from the data in Table I (Ref. 16):

$$\alpha_{ij}=2(\mu_{ij}^{0l}-\mu_i^{0l}-\mu_j^{0l}+RT\ \ln 4), \tag{2}$$

where μ_i^{0l} and μ_{ij}^{0l} are the standard chemical potentials of the components (i) and compounds (ij). The parameters β_{ij} and Ψ_{ijk} are the differences in the free energies of formation of the compounds iij, ijj, and ijk, respectively. Since such compounds are not formed in III–V systems, we found them from the data on the fusibility curves of the corresponding binary, quasibinary, and ternary subsystems of the In–Ga–As–Sb–P system (just as the parameters $\alpha_{\rm InGa},\beta_{\rm AsP},\alpha_{\rm AsP}$, and $\alpha_{\rm SbP}$). These parameters are presented in Table II.

We can calculate the chemical potentials μ_{ij}^l of all six III–V compounds in the melt (InAs, GaAs, InP, GaP, InSb, and GasB) from the data on the standard and excess thermodynamic functions of the melt components:

$$\mu_{ij}^l=\mu_{ij}^{0l}+RT[\ln(x_ix_j)+\ln(\gamma_i\gamma_j)]. \tag{3}$$

Model of the solid solution. As components of the quaternary common solid solution $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}_y\mathrm{Sb}_z\mathrm{P}_{1-y-z}$, we shall study the III–V compounds InAs, GaAs, InP, GaP, GaSb, and InSb. The concentrations of two components depend on the concentrations of the other four, since the equation of chemical equilibrium for a phase of variable composition — solid solution — holds:[2]

$$\begin{cases}\mu^s_{\rm InAs}+\mu^s_{\rm GaP}=\mu^s_{\rm GaAs}+\mu^s_{\rm InP},\\ \mu^s_{\rm InSb}+\mu^s_{\rm GaP}=\mu^s_{\rm InP}+\mu^s_{\rm GaSb},\end{cases} \tag{4}$$

where μ^s_{ij} is the chemical potential of the compound ij in the solid phase. Equations (4) are a consequence of the establishment of equilibrium in reactions proceeding in the solid phase:

$$\begin{cases}\mathrm{InAs}+\mathrm{GaP}\Leftrightarrow\mathrm{GaAs}+\mathrm{InP},\\ \mathrm{InSb}+\mathrm{GaP}\Leftrightarrow\mathrm{GaSb}+\mathrm{InP}.\end{cases} \tag{5}$$

We consider first the formation of a solid solution which is strictly isoperiodic to the chosen substrate. In other words, the lattice periods of the solid solution and substrate are the same or, according to Vegard's rule,

TABLE II. Nonideality parameters of coexisting liquid and solid phases in the system In–Ga–As–P. $P=a+b\cdot T$.

Parameter	a, kJ	b, J/K
$\alpha_{\rm InGa}$	−14.0	20.9
$\alpha_{\rm AsP}$	−9.0	-
$\alpha_{\rm AsSb}$	24.6	21.0
$\alpha_{\rm SbP}$	73.5	-
$\beta_{\rm InAs}$	39.2	−41.0
$\beta_{\rm GaAs}$	39.0	−39.0
$\beta_{\rm GaSb}$	21.0	−20.0
$\beta_{\rm InSb}$	56.0	−100.0
$\beta_{\rm InP}$	27.6	−80.0
$\beta_{\rm GaP}$	-10^5	-
$\Psi_{\rm InGaAs}=\Psi_{\rm InGaP}$	33.6	−50.0
$\Psi_{\rm InAsP}=\Psi_{\rm GaAsP}$	60.0	-
$\Psi_{\rm InAsSb}=\Psi_{\rm GaAsSb}$	−127.0	100.0
$\Psi_{\rm InSbP}=\Psi_{\rm GaSbP}$	−63.0	-
$2\alpha^s_{\rm InAsGaAs}$	5.7	20.9
$2\alpha^s_{\rm InPGaP}$	18.8	-
$2\alpha^s_{\rm InAsInP}$	2.5	-
$2\alpha^s_{\rm GaAsGaP}$	3.8	-
$2\alpha^s_{\rm GaAsGaSb}$	1.4	27.3
$2\alpha^s_{\rm InAsInSb}$	21.0	-
$2\alpha^s_{\rm InSbInP}$	47.0	-
$2\alpha^s_{\rm GaSbGaP}$	73.9	-

$$a=a_{\mathrm{InAs}}y_{\mathrm{In}}y_{\mathrm{As}}+a_{\mathrm{GaAs}}y_{\mathrm{Ga}}y_{\mathrm{As}}+a_{\mathrm{InP}}y_{\mathrm{In}}y_{\mathrm{P}}+a_{\mathrm{GaP}}y_{\mathrm{Ga}}y_{\mathrm{P}}$$

$$+a_{\mathrm{InSb}}y_{\mathrm{In}}y_{\mathrm{Sb}}+a_{\mathrm{GaSb}}y_{\mathrm{Ga}}y_{\mathrm{Sb}}, \tag{6}$$

where a is the lattice constant of the substrate at 300 K, a_{ij} is the lattice constant of the compound ij at 300 K (6.058 for InAs, 5.653 for GaAs, 5.451 for GaP) and y_i is the atomic fraction in the III(V) sublattice:

$$y_{\mathrm{In}}+y_{\mathrm{Ga}}=y_{\mathrm{As}}+y_{\mathrm{P}}+y_{\mathrm{Sb}}=1. \tag{7}$$

To describe the excess thermodynamic functions of the components of the solid solutions — the activity coefficients γ_{ij}^{is} of the compounds ij in the molar fraction scale, we employed the well-known model of quasiregular solid solutions. For example, the activity coefficients of InAs satisfy

$$RT\ \ln\ \gamma_{\mathrm{InAs}}^{is}=y_{\mathrm{Ga}}^{2}\alpha_{\mathrm{InGa}}^{s}+y_{\mathrm{P}}^{2}\alpha_{\mathrm{AsP}}^{s}+y_{\mathrm{Sb}}^{2}\alpha_{\mathrm{AsAb}}^{s}+y_{\mathrm{P}}y_{\mathrm{Sb}}(\alpha_{\mathrm{AsP}}^{s}$$

$$+\alpha_{\mathrm{AsSb}}^{s}-\alpha_{\mathrm{PSb}}^{s}), \tag{8}$$

where

$$\left[\begin{array}{l}\alpha_{\mathrm{InGa}}^{s}=\alpha_{\mathrm{InPGaP}}^{s}y_{\mathrm{P}}+\alpha_{\mathrm{InAsGaAs}}^{s}y_{\mathrm{As}}+\alpha_{\mathrm{InSbGaSb}}^{s}y_{\mathrm{Sb}},\\ \alpha_{\mathrm{AsP}}^{s}=\alpha_{\mathrm{InAsInP}}^{s}y_{\mathrm{In}}+\alpha_{\mathrm{GaAsGaP}}^{s}y_{\mathrm{Ga}},\\ \alpha_{\mathrm{AsSb}}^{s}=\alpha_{\mathrm{InAsInSb}}^{s}y_{\mathrm{In}}+\alpha_{\mathrm{GaAsGaSb}}^{s}y_{\mathrm{Ga}},\\ \alpha_{\mathrm{PSb}}^{s}=\alpha_{\mathrm{InSbInP}}^{s}y_{\mathrm{In}}+\alpha_{\mathrm{GaSbGaP}}^{s}y_{\mathrm{Ga}}.\end{array}\right. \tag{9}$$

In using Eqs. (8) and (9) to estimate the lattice nonideality parameters α_{ijkj}^{s} (α_{ijik}^{s}), we employed an approximation formula obtained in Ref. 17; the solid phase parameters which we used are presented in Table II.

In calculating the phase equilibria with the formation of solid solutions which are not isoperiodic to the substrate, we took into account the contribution of the energies of the elastic stresses associated with the compression–stretching of the III–V bonds in the solid solution to the chemical potentials:

$$RT\ \ln\ \gamma_{ij}=RT\ \ln\ \gamma_{ij}^{is}+E_{ij}^{\mathrm{def}}, \tag{10}$$

where

$$E_{ij}^{\mathrm{def}}=C_{0r}[(a-a_{\mathrm{sub}})^{2}+2(a-a_{\mathrm{sub}})(a_{ij}-a)]/a_{\mathrm{sub}}^{2}. \tag{11}$$

Here a_{sub} is the lattice constant of the substrate, and C_{0r} is a constant which, in general, depends on the temperature, the type of III–V compound (ij), and the substrate orientation. Disregarding the relatively weak dependence on the first two parameters, we assume C_{0r} = 4200 kJ/mole for the (100) orientation and 3000 kJ/mole for the (111) orientation.[17]

Therefore, to calculate the chemical potentials μ_{ij}^{s} of the III–V compounds in the solid solution, we have

$$\mu_{ij}^{s}=\mu_{ij}^{0s}+RT\ \ln\ y_{i}y_{j}+RT\ \ln\ \gamma_{ij}+\delta_{ij}^{\mathrm{mut}}, \tag{12}$$

where μ_{ij}^{0s} is the standard chemical potential of the compound ij in the solid solution, and $\delta_{ij}^{\mathrm{mut}}$ is the contribution of a standard variation of the Gibbs free energy of the reaction (5) to the chemical potential of the compound ij. We introduce the functions

$$\left[\begin{array}{l}\delta_{1}^{\mathrm{reac}}=\mu_{\mathrm{GaAs}}^{0s}+\mu_{\mathrm{InP}}^{0s}-\mu_{\mathrm{GaP}}^{0s}-\mu_{\mathrm{InAs}}^{0s},\\ \delta_{2}^{\mathrm{reac}}=\mu_{\mathrm{GaSb}}^{0s}+\mu_{\mathrm{InP}}^{0s}-\mu_{\mathrm{GaP}}^{0s}-\mu_{\mathrm{InSb}}^{0s},\\ \delta_{3}^{\mathrm{reac}}=\delta_{1}^{\mathrm{reac}}-\delta_{2}^{\mathrm{reac}},\end{array}\right. \tag{13}$$

$$\left[\begin{array}{l}\delta_{1}^{\mathrm{def}}=E_{\mathrm{InP}}^{\mathrm{def}}+E_{\mathrm{GaAs}}^{\mathrm{def}}-E_{\mathrm{InAs}}^{\mathrm{def}}-E_{\mathrm{GaP}}^{\mathrm{def}},\\ \delta_{2}^{\mathrm{def}}=E_{\mathrm{InP}}^{\mathrm{def}}+E_{\mathrm{GaSb}}^{\mathrm{def}}-E_{\mathrm{InSb}}^{\mathrm{def}}-E_{\mathrm{GaP}}^{\mathrm{def}},\\ \delta_{3}^{\mathrm{def}}=\delta_{1}^{\mathrm{def}}-\delta_{2}^{\mathrm{def}}.\end{array}\right. \tag{14}$$

For the solid solution InGaAsSbP we then have

$$\left[\begin{array}{l}\delta_{\mathrm{InSb}}^{\mathrm{mut}}=y_{\mathrm{Ga}}y_{\mathrm{P}}(-\delta_{2}^{\mathrm{reac}}+\delta_{2}^{\mathrm{def}})+y_{\mathrm{Ga}}y_{\mathrm{As}}(-\delta_{1}^{\mathrm{reac}}+\delta_{1}^{\mathrm{def}}),\\ \delta_{\mathrm{GaP}}^{\mathrm{mut}}=y_{\mathrm{In}}y_{\mathrm{Sb}}(-\delta_{2}^{\mathrm{reac}}-\delta_{2}^{\mathrm{def}})+y_{\mathrm{In}}y_{\mathrm{As}}(-\delta_{3}^{\mathrm{reac}}+\delta_{3}^{\mathrm{def}}),\\ \delta_{\mathrm{InAs}}^{\mathrm{mut}}=y_{\mathrm{Ga}}y_{\mathrm{Sb}}(\delta_{1}^{\mathrm{reac}}-\delta_{1}^{\mathrm{def}})+y_{\mathrm{Ga}}y_{\mathrm{P}}(\delta_{3}^{\mathrm{reac}}-\delta_{3}^{\mathrm{def}}),\end{array}\right. \tag{15}$$

etc.

Conditions of phase equilibrium. Therefore, to determine the composition of the coexisting liquid and solid phases it is sufficient to solve any four (or six) equations of phase equilibrium:

$$\left[\begin{array}{l}\mu_{\mathrm{InAs}}^{l}=\mu_{\mathrm{InAs}}^{s},\\ \mu_{\mathrm{InP}}^{l}=\mu_{\mathrm{InP}}^{s},\\ \mu_{\mathrm{InSb}}^{l}=\mu_{\mathrm{InSb}}^{s},\\ \mu_{\mathrm{GaSb}}^{l}=\mu_{\mathrm{GaSb}}^{s},\\ \mu_{\mathrm{GaAs}}^{l}=\mu_{\mathrm{GaAs}}^{s},\\ \mu_{\mathrm{GaP}}^{l}=\mu_{\mathrm{GaP}}^{s}.\end{array}\right. \tag{16}$$

Any fifth equation will be a linear combination of the first four equations and can be eliminated from the system. The system (16) can be solved for four unknown variables from, for example, the following set: x_{Ga}, x_{As}, x_{P}, x_{Sb}, y_{Ga}, y_{As}, y_{Sb}, T ($x_{\mathrm{In}}=1-x_{\mathrm{Ga}}-x_{\mathrm{As}}-x_{\mathrm{P}}-x_{\mathrm{Sb}}$, $y_{\mathrm{In}}=1-y_{\mathrm{Ga}}$, $y_{\mathrm{P}}=1-y_{\mathrm{As}}-y_{\mathrm{Sb}}$). Therefore we have the following possibility: 1) Fix T, y_{Ga}, y_{As} and the condition of isoperiodicity of some substrate, for example, InAs, and find y_{Sb} and y_{P} in accordance with Eq. (6) and the composition of the liquid phase x_{Ga}, x_{As}, x_{P}, and x_{Sb} in accordance with the system of equations (16) [the functions E_{ij}^{def} and δ^{def} will then be equal to zero in accordance with Eqs. (11) and (14)]; 2) fix x_{Ga}, x_{As}, x_{P}, and x_{Sb} and find in accordance with the system (6) the liquidus temperature T and the composition of the solid solution y_{Ga}, y_{As}, and y_{Sb} (the functions F_{ij}^{def} and $\delta_{i}^{\mathrm{ref}}$ become nonvanishing).

Limitations on the preparation of solid solutions. Let us examine first the limitations on obtaining the solid solutions InGaAsSbP/GaSb(InAs). According to the analysis in Ref. 13, there are two independent limitations on the region of existence of such solid solutions: with respect to the ‘‘condition of molecularity of the melt,’’ which becomes wider with increasing temperature, and with respect to spinodal decomposition of the solid solutions, which becomes narrower with increasing temperature.

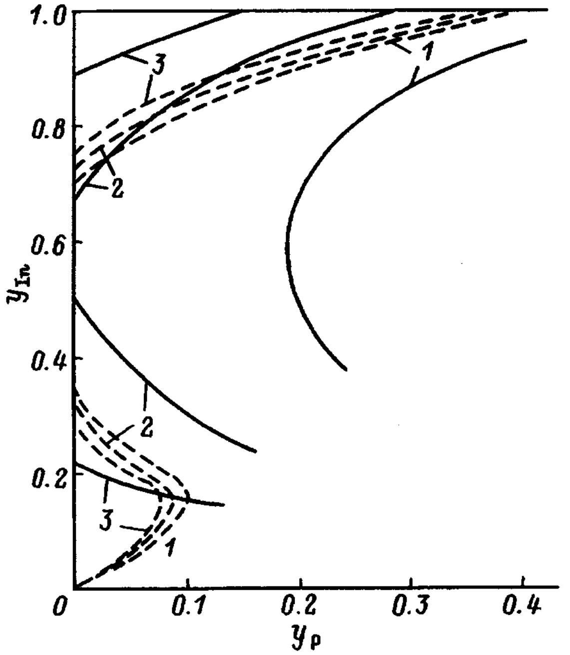

FIG. 1. Concentration dependences of the limitations with respect to the condition of molecularity of the melt (solid line) and with respect to spinodal decomposition of the solid solution (dashed lines) for the solid solutions $In_xGa_{1-x}As_ySb_zP_{1-y-z}$, isoperiodic to an InAs substrate, for the following temperatures, °C: *1* — 500, *2* — 550, *3* — 600.

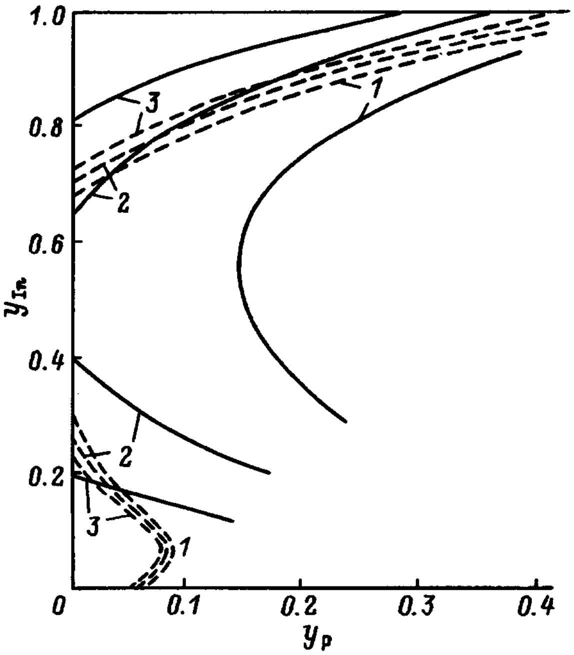

FIG. 2. Concentration dependences of the limitations with respect to the condition of molecularity of the melt (solid line) and with respect to spinodal decomposition of the solid solution (dashed lines) for the solid solutions $In_xGa_{1-x}As_ySb_zP_{1-y-z}$, isoperiodic to a GaSb substrate, for the following temperatures, °C: *1* — 500, *2* — 550, *3* — 600.

Limitation with respect to the ''condition of molecularity of the melt.'' According to Ref. 13, this condition comes into play when a melt can be prepared exclusively from the III–V compounds (InAs, InP, InSb, GaAs, GaSb, and GaP) without additions of the components (In, Ga, As, Sb, P):

$$x_{\mathrm{In}}+x_{\mathrm{Ga}}=x_{\mathrm{As}}+x_{\mathrm{P}}+x_{\mathrm{Sb}}=1/2. \tag{17}$$

The computational data for the compositions of the melts, in equilibrium with the solid solutions which are isoperiodic to GaSb(InAs) substrates, are presented in Figs. 1 and 2 for the temperatures 500, 550, and 600 °C. The data (here and below) are presented in the coordinates $y_{\mathrm{In}}-y_{\mathrm{P}}$. These two independent variables of the composition of the solid solution are sufficient to determine it completely, since the condition of isoperiodicity to some substrate [Eq. (6)] and the normalization equations [Eqs. (7)] hold. It is evident from the figures that the region of limitation shrinks very strongly as the temperature decreases and ceases to be determining at temperatures below 550 °C. We note that the same set of solid solutions can also be obtained from a melt with $x_{\mathrm{As}}+x_{\mathrm{P}}+x_{\mathrm{Sb}}>1/2$ at the same temperatures. This fact, however, is apparently of no practical importance, since P and As have a large partial pressure even for values such that $x_{\mathrm{As}}+x_{\mathrm{P}}+x_{\mathrm{Sb}}>0.5-0.6$, which makes epitaxial synthesis virtually uncontrollable.

Limitation with respect to the spinodal decomposition of solid solutions. This limitation arises in connection with the fact that the phase of the homogeneous solid solution can be unstable. The boundary of stability of the solid solution is determined by the following equation (the solid solution is not formed and does not decompose on the substrate):^{2}

$$\begin{vmatrix} G^*_{\mathrm{GaGa}} & G^*_{\mathrm{GaAs}} & G^*_{\mathrm{GaP}} \\ G^*_{\mathrm{AsGa}} & G^*_{\mathrm{AsAs}} & G^*_{\mathrm{AsP}} \\ G^*_{\mathrm{PGa}} & G^*_{\mathrm{PAs}} & G^*_{\mathrm{PP}} \end{vmatrix}=0, \tag{18}$$

where $G^*_{ij}=\partial^2G^*/\partial y_i\partial y_j$, G^* is the part of the average molar Gibbs free energy of the solid solution, whose second derivatives with respect to the variable compositions are nonvanishing:

$$\begin{aligned} G^*=RT[&y_{\mathrm{In}}y_{\mathrm{As}}\ln(y_{\mathrm{In}}y_{\mathrm{As}})+y_{\mathrm{Ga}}y_{\mathrm{As}}\ln(y_{\mathrm{Ga}}y_{\mathrm{As}}) \\ &+y_{\mathrm{In}}y_{\mathrm{P}}\ln(y_{\mathrm{In}}y_{\mathrm{P}})+y_{\mathrm{Ga}}y_{\mathrm{P}}\ln(y_{\mathrm{Ga}}y_{\mathrm{P}}) \\ &+y_{\mathrm{In}}y_{\mathrm{Sb}}\ln(y_{\mathrm{In}}y_{\mathrm{Sb}})+y_{\mathrm{Ga}}y_{\mathrm{Sb}}\ln(y_{\mathrm{Ga}}y_{\mathrm{Sb}})] \\ &+[y_{\mathrm{In}}y_{\mathrm{As}}\mu^{0s}_{\mathrm{InAs}}+y_{\mathrm{In}}y_{\mathrm{P}}\mu^{0s}_{\mathrm{InP}}+y_{\mathrm{In}}y_{\mathrm{Sb}}\mu^{0s}_{\mathrm{InSb}} \\ &+y_{\mathrm{Ga}}y_{\mathrm{As}}\mu^{0s}_{\mathrm{GaAs}}+y_{\mathrm{Ga}}y_{\mathrm{P}}\mu^{0s}_{\mathrm{GaP}}+y_{\mathrm{Ga}}y_{\mathrm{Sb}}\mu^{0s}_{\mathrm{GaSb}}] \\ &+y_{\mathrm{In}}y_{\mathrm{Ga}}[\alpha^s_{\mathrm{InAsGaAs}}y_{\mathrm{As}}+\alpha^s_{\mathrm{InPGaP}}y_{\mathrm{P}}+\alpha^s_{\mathrm{InSbGaSb}}y_{\mathrm{Sb}}] \\ &+y_{\mathrm{As}}y_{\mathrm{P}}[\alpha^s_{\mathrm{InAsInP}}y_{\mathrm{In}}+\alpha^s_{\mathrm{GaAsGaP}}y_{\mathrm{Ga}}] \\ &+y_{\mathrm{As}}y_{\mathrm{Sb}}[\alpha^s_{\mathrm{InAsInSb}}y_{\mathrm{In}}+\alpha^s_{\mathrm{GaAsGaSb}}y_{\mathrm{Ga}}] \\ &+y_{\mathrm{Sb}}y_{\mathrm{P}}[\alpha^s_{\mathrm{InSbInP}}y_{\mathrm{In}}+\alpha^s_{\mathrm{GaSbGaP}}y_{\mathrm{Ga}}]. \end{aligned} \tag{19}$$

If the solid solution decomposes on the substrate, then in Eq. (18) G^* should be replaced by G^{**}, which is determined by Eqs. (10–(12):

$$\begin{aligned} G^{**}=G^*&+y_{\mathrm{In}}y_{\mathrm{As}}E^{\mathrm{def}}_{\mathrm{InAs}}+y_{\mathrm{Ga}}y_{\mathrm{As}}E^{\mathrm{def}}_{\mathrm{GaAs}}+y_{\mathrm{In}}y_{\mathrm{P}}E^{\mathrm{def}}_{\mathrm{InP}} \\ &+y_{\mathrm{Ga}}y_{\mathrm{P}}E^{\mathrm{def}}_{\mathrm{GaP}}+y_{\mathrm{In}}y_{\mathrm{Sb}}E^{\mathrm{def}}_{\mathrm{InSb}}+y_{\mathrm{Ga}}y_{\mathrm{Sb}}E^{\mathrm{def}}_{\mathrm{GaSb}}. \end{aligned} \tag{20}$$

The solutions are unstable if the determinant in Eq. (18) and its minors belonging to the main diagonal are nonpositive.[2] Then, for the solid solution to be stable, it is necessary that [in addition to the requirement that the determinant (18) be positive] the following system of six inequalities be satisfied simultaneously:

$$\begin{vmatrix} G^*_{ii} & G^*_{ij} \\ G^*_{ji} & G^*_{jj} \end{vmatrix} > 0 \left[\begin{array}{ll} i=\mathrm{Ga}, & j=\mathrm{As}, \\ i=\mathrm{Ga}, & j=\mathrm{P}, \\ i=\mathrm{As}, & j=\mathrm{P}. \end{array} \right. \tag{21}$$

$$G^*_{ii} > 0 \left[\begin{array}{l} i=\mathrm{Ga}, \\ i=\mathrm{As}, \\ i=\mathrm{P}. \end{array} \right. \tag{22}$$

It should be noted that the region of instability (or decay) of solid solutions when E^{def}_{ij} is taken into account is narrower than when it is disregarded; i.e., the substrate ''stabilizes'' the decaying solid solutions.

Instead of calculating the elements G^*_{ij} in an explicit form, we employed a finite-difference approximation, for example,

$$\begin{aligned} G^*_{\mathrm{GaAs}} \cong [& G^*(y_{\mathrm{Ga}}+\Delta, y_{\mathrm{As}}+\Delta, y_{\mathrm{P}}) \\ & + G^*(y_{\mathrm{Ga}}, y_{\mathrm{As}}, y_{\mathrm{P}}) - G^*(y_{\mathrm{Ga}}+\Delta, y_{\mathrm{As}}, y_{\mathrm{P}}) \\ & - G^*(y_{\mathrm{Ga}}, y_{\mathrm{As}}+\Delta, y_{\mathrm{P}})]/\Delta^2 \end{aligned} \tag{23}$$

etc. Here Δ is the numerical differentiation step, corresponding to $\Delta = 10^{-9}$ arb. units.

The computational results for the limitation with respect to the spinodal decomposition of the solid solutions InGaAsSbP/GaSb(InAs) and LPE temperatures of 500, 550, and 600 °C are presented in Figs. 1 and 2. It is obvious from the figure that this region of limitations deforms with increasing temperature (shrinks), though much more slowly than the region of the limitation with respect to the condition of molecularity of the melt.

General limitation on obtaining solid solutions. We showed elsewhere[18] that in studying quaternary systems there arises a critical point corresponding to the intersection of both limitations. This point shows the extremely high (low) concentration of a component that can in principle be obtained in the solid solution isoperiodic with a given substrate for any temperature, and it also determines the temperature of the two-phase equilibrium. In our case the system possesses one more degree of freedom, and the collection of points of intersection of the two limitations determines in the most general case a monovariant polythermal curve. The different extremal compositions of the solid solutions therefore correspond to different LPE temperatures (see Figs. 1 and 2). The general composition regions accessible to synthesis by the LPE method are presented in Fig. 3. Specifically, it is obvious from the figure that at relatively low temperatures the region of accessible compositions is determined by the limitation with respect to spinodal decomposition of solid solutions, and for relatively high temperatures it is determined by both limitations. Therefore, we have determined the range of compositions of five-component solid solutions in the system In–Ga–As–Sb–P, which are isoperiodic to InAs and GaSb substrates and which can in principle be obtained by the LPE method.

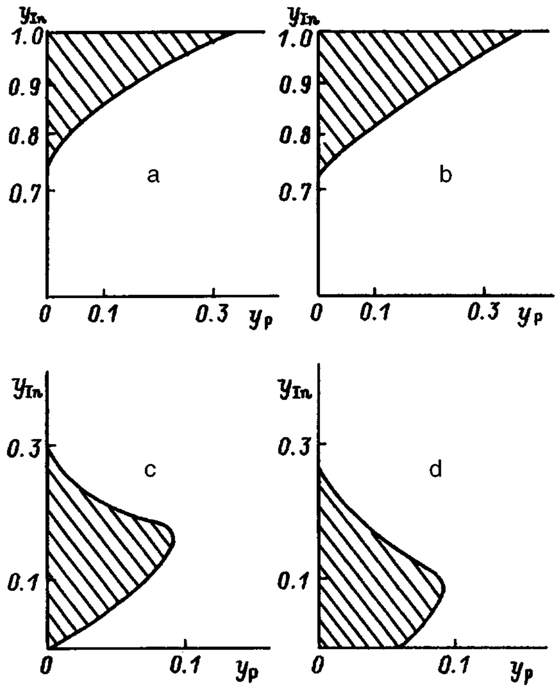

FIG. 3. General limitations on obtaining solid solutions $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}_y\mathrm{Sb}_z\mathrm{P}_{1-y-z}$ isoperiodic to the substrates GaSb (a, c) and InAs (b, d).

In summary, in the present paper we have 1) examined the conditions for obtaining five-component solid solutions InGaAsSbP, which are isoperiodic to InAs and GaSb, by LPE in the entire range of compositions; 2) calculated the fusibility curves of the system In–Ga–As–Sb–P in the crystallization field of the solid solutions and the curves of the solid (I)–solid(II) phase equilibria; and 3) determined the effective regions of the limitations on LPE synthesis of the indicated solutions at a given temperature and determined the ranges of the compositions of the solutions which are accessible in principle for synthesis by the LPE method.

In a future paper we shall present computational results for a series of electrical parameters of the five-component solutions studied in the system In–Ga–As–Sb–P: band gaps, electron affinity, and spin-orbit splitting energy as well as the possible energy band diagrams of the solid solution/substrate.

This work was supported, in part, by the Russian Fund for Fundamental Research (Project 93-03-5492).

[1] J. W. Gibbs, *Thermodynamics. Statistical Mechanics* [Russian trans.], Nauka, Moscow (1982), p. 584.
[2] A. V. Storonkin, *Thermodynamics of Heterogeneous Systems* [in Russian], Leningrad State University, Leningrad (1967), Part 1, p. 467.
[3] M. P. Mikhaĭlova, K. D. Moiseev, O. G. Ershov, and Yu. P. Yakovlev, Fiz. Tekh. Poluprovodn. **30**, 399 (1996) [Semiconductors **30**, 223 (1996)].
[4] M. Mebarki, A. Cardi, and H. Mani, J. Appl. Phys. **73**, 2360 (1993).
[5] T. N. Danilova, A. N. Imenkov, K. D. Moiseev, I. N. Timchenko, and Yu. P. Yakovlev, Pis'ma Zh. Tekh. Fiz. **20**, 20 (1994) [*sic*].
[6] A. I. Andreev, A. N. Baranov, M. P. Mikhaĭfikova, K. D. Moiseev, A. V.

Pentsov, Yu. P. Smorchkova, V. V. Sherstnev, and Yu. P. Yakovlev, Pis'ma Zh. Tekh. Fiz. **18**, 50 (1992) [*sic*].

[7] Yu. Yakovlev, A. N. Baranov, A. N. Imenkov, and V. V. Sherstnev, J. de Phys. **65**, 3219 (1984).

[8] A. N. Baranov, A. N. Imenkov, V. V. Sherstnev, and Yu. Yakovlev, Appl. Phys. Lett. **64**, 3219 (1994).

[9] A. N. Baranov, A. N. Imenkov, V. V. Sherstnev, and Yu. Yakovlev, Pis'ma Zh. Tekh. Fiz. **16**, 19 (1990) [*sic*].

[10] A. N. Baranov, N. A. Charykov, A. M. Litvak, V. V. Sherstnev, and Yu. P. Yakovlev, *First International Conference on Epit. Cryst. Growth*, Budapest (1990), p. 390.

[11] M. Panish and M. Ilegems, *Progress in Solid State Chemistry*, New York (1972) p. 390.

[12] O. Madelung, *Physics of Group III and V Semiconductor Elements* [Russian trans.], Mir, Moscow (1967), p. 480.

[13] A. M. Litvak and N. A. Charykov, Zh. Fiz. Khim. **66**, 923 (1992).

[14] A. M. Litvak and N. A. Charykov, Fiz. Tekh. Poluprovodn. **24**, 2106 (1990) [Sov. Phys. Semicond. **24**, 1307 (1990)].

[15] A. M. Litvak and N. A. Charykov, Zh. Fiz. Khim. **64**, 2331 (1990).

[16] A. M. Litvak and N. A. Charykov, Izv. Akad. Nauk SSSR, Neor. Mater. **27**, 225 (1991).

[17] A. M. Litvak and N. A. Charykov, Zh. Neorg. Khim. **35**, 3008 (1990).

[18] A. N. Baranov, A. M. Litvak, K. D. Moiseev, V. V. Sherstnev, and Yu. P. Yakovlev, Zh. Prik. Khim. **63**, 976 (1990).

Translated by M. E. Alferieff

Interband magnetooptic absorption line shape in bismuth

S. V. Brovko, A. A. Zaĭtsev, K. G. Ivanov, and O. V. Kondakov

St. Petersburg State University of Technology and Design, 191065 St. Petersburg, Russia

(Submitted May 13, 1996; accepted for publication June 3, 1996)
Fiz. Tekh. Poluprovodn. **31**, 416–419 (April 1997)

The results of an investigation of the transmission of a symmetric bismuth stripline at $T=80$ K at the laser wavelength $\lambda=10.6$ μm in magnetic fields up to $B=8$ T are reported. A set of parameters is obtained for the energy spectrum of the L electrons of bismuth by modeling the shape of the experimental curve on the basis of a modified Baraff model. The values of the parameters in the McClure–Choi model are found from an analysis of the field positions of the maxima of the mageto-optic oscillations. [S1063-7826(97)00704-7]

We have investigated the transmission of 10.6-μm radiation by a waveguide line, assembled from two square halves of a bismuth single crystal, in pulsed magnetic fields. Such a system[1] is called a symmetric stripline (SSL). The optical surfaces of the single crystals were prepared by chemical polishing. The ratio of the resistivity of the experimental samples at room temperature to the resistivity at 4.2 K was greater than 400. The transmission of the SSL in a magnetic field at $T=80$ K is oscillatory; the oscillations are associated with optical interband transitions on the Landau levels of the valence and conduction bands. To extract maximum information from the data obtained and to analyze in greater detail the physics of the observed effects, a calculation of the lineshape of the magneto-optic spectrum was performed.

The modeling was performed in a modified Baraff model.[2] The transmission coefficient

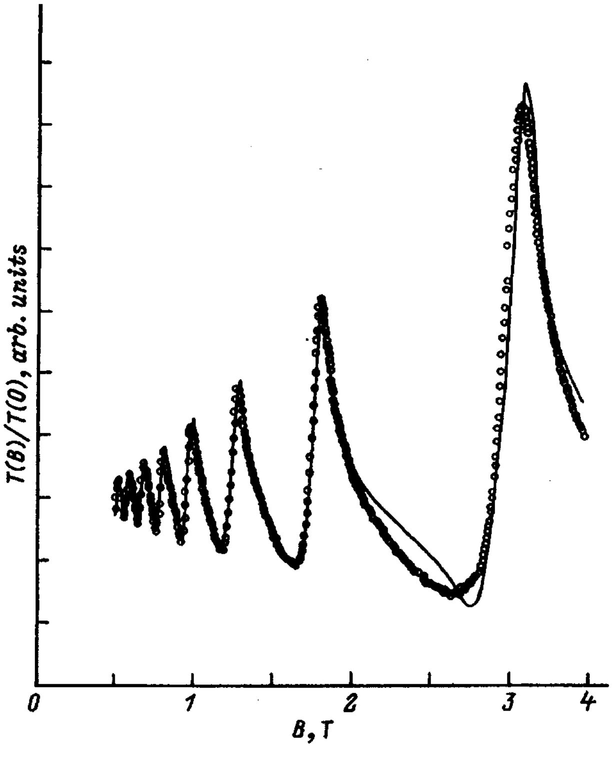

FIG. 1. Transmission of a stripline as a function of the magnitude of the magnetic field **B** directed along the binary axis of the crystal lattice of bismuth. The distance between the two mirror-symmetric halves of the bismuth single crystal is $D_z=25$ μm; the relaxation time associated with the width of a Landau level is $\tau=4\times10^{-3}$ s; the dimensionless wave vector in the direction of the external magnetic field is $\xi=hk_z/(\varepsilon m_B^*)^{1/2}=2$, where $\hbar$ is Planck's constant, k_z is the wave vector parallel to the magnetic field, $\varepsilon=E_g/2$, and m_B^* is the effective mass along the magnetic field. Solid line — calculation, dots — experiment.

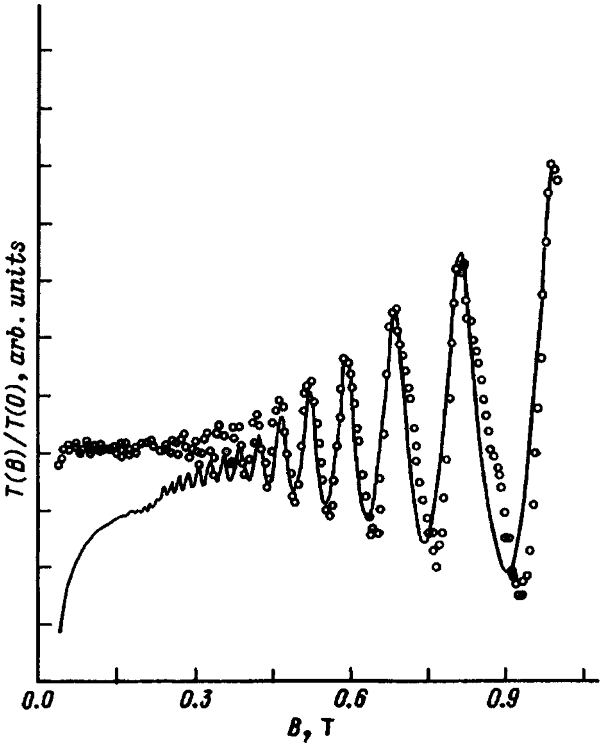

FIG. 2. Transmission of a stripline as a function of the magnitude of the magnetic field **B** directed along the binary axis of the crystal lattice of bismuth in fields up to 1 T. Solid line — calculation, dots — experiment. The values of the parameters D_z, τ, and ξ are the same as in Fig. 1.

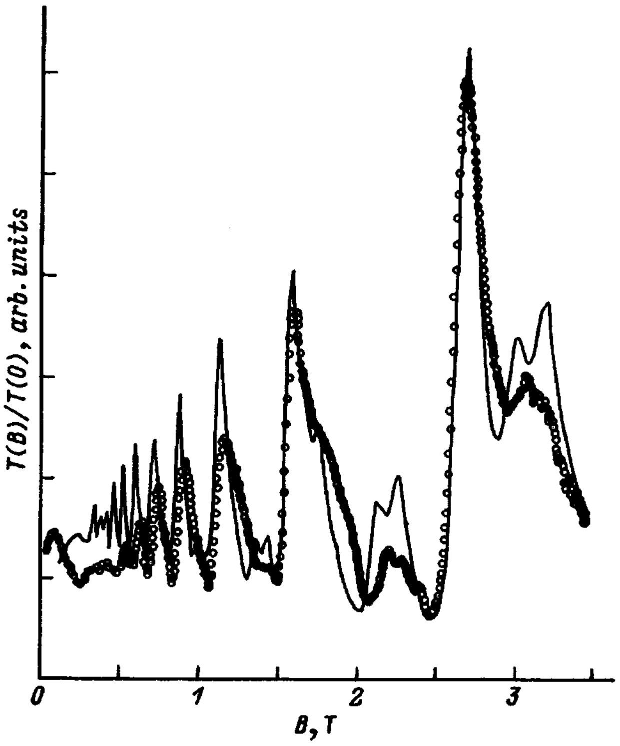

FIG. 3. Transmission of a stripline as a function of the magnitude of the magnetic field **B** directed at an angle of 1° from the bisectrix axis in the basal plane. Solid line — calculation, dots — experiment. The values of the parameters D_z, τ, and ξ are the same as in Figs. 1 and 2.

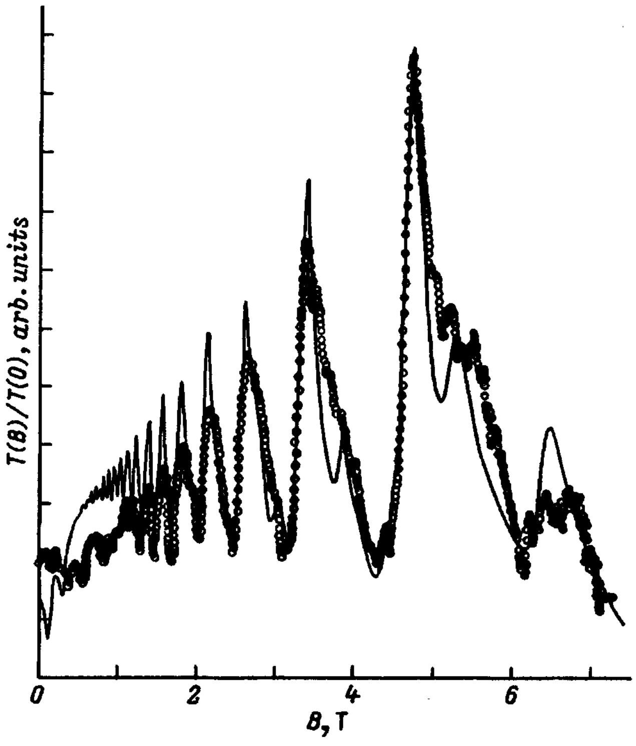

FIG. 4. Transmission of a stripline as a function of the magnitude of the magnetic field **B** directed at an angle of 11.5° from the trigonal axis in bisectrix–trigonal plane. Solid line — calculation, dots — experiment. The values of the parameters D_z, τ, and ξ are the same as in Figs. 1—3.

$$\frac{T(B)}{T(0)}=\frac{\{\sinh[a_0'(B)D]/a_0'(B)+\sin[a_0''(B)D]/a_0''(B)\}}{\{\sinh[a_0'(0)D]/a_0'(0)+\sin[a_0''(0)D]/a_0''(0)\}}\times\left\{\left[\frac{[a_0'(0)]^2+a_0''(0)}{q_y'+q_y''}\right]+1\right\}\times\frac{q_y'}{\mathrm{Re}[k_0^2/q_y(0)]}e^{2L[q_y''(B)-q_y''(0)]}, \tag{1}$$

where D is the gap width between two mirror-symmetric halves of the single crystal, $a_0'(B)$ and $a_0''(B)$ are the real and imaginary parts of a_0 for some value of the magnetic field B, $a_0=\pm\sqrt{q_y^2-k_0^2}$, $q_y=(q_y', q_y'')$ is the complex wave vector of the electromagnetic wave, $k_0=\omega/c$, ω is the angular frequency of the electromagnetic wave, c is the speed of light in vacuum, and L is the length of the experimental crystal, was found by solving the dispersion relation numerically for a wave propagating in the gap between two mirror-symmetric halves of the single crystals,

$$C_2\left(a_2+a_0\tanh\frac{a_0D}{2}\right)\left(a_0L_1\tanh\frac{a_0D}{2}-k_0^2B_1\right)-C_1\left(a_1+a_0\tanh\frac{a_0D}{2}\right)\left(a_0L_2\tanh\frac{a_0D}{2}-k_0^2B_2\right)=0, \tag{2}$$

where $C_n=k_0^2[\varepsilon_{xx}T_n-\varepsilon_{xy}(k_0^2\varepsilon_{xy}-ia_nq_y)]$, $T_n=k_0^2-\varepsilon_{yy}+a_n^2$, $L_n=a_nB_n+iq_yA_n$, $B_n=(k_0^2\varepsilon_{yz}-ia_nq_y)S_n-k_0^4\varepsilon_{xx}\varepsilon_{yx}$, $S_n=k_0^2\varepsilon_{xx}-q_y^2a_n^2$, and $A_n=S_nT_n-k_0^4\varepsilon_{xy}\varepsilon_{yx}$ $(n=1,2)$.

Equation (2) was obtained by solving a system of Maxwell's equations with boundary conditions. The complex roots with positive real parts a_1 and a_2 are found from Eq. (2).

The experimental and computed values of the transmission coefficient of the SSL for the magnetic field directed parallel to the binary axis of the crystal lattice of bismuth are shown in Figs. 1 and 2. The spectra were normalized to the transmission intensity in a zero magnetic field. Up to 15 oscillations corresponding to interband transitions on Landau levels with number $j\geqslant 1$ were observed with the orientation employed. To fit the computed curve to the experimental curve, it was necessary to introduce constant values of the complex lattice permittivity in directions parallel (ε_{yy}) and perpendicular ($\varepsilon_{xx}=\varepsilon_{zz}$) to the magnetic field. The gap width was assumed to be[2] $E_g=15.6$ meV.

Agreement between the experimental and theoretical curves was achieved by using visual monitoring and an automatic parametric Nelder–Mead method. The small increase in the effectiveness of lineshape fit to the experimental lineshape by this method is explained by the difference in the behavior of the experimental and theoretical curves on the initial section of the magnetic fields. Transitions from levels $j\geqslant 19$, which were not observed in the experiment, make a large contribution to the low-frequency permittivity in fields up to 0.3 T. When the number of working Landau levels is increased up to $N=150$ (Fig. 1 and 2), the model curve shifts upward and approaches the experimental curve.

In the model considered it was found to be impossible to eliminate the discrepancy in the behavior of the curves for

TABLE I. Parameters of the energy spectrum of the L electrons of bismuth at 80 K, calculated from the lineshape in a modified Baraff model.

<table>
<tr><td>Effective-mass tensor components</td><td>$m_1=0.00174$</td><td>$m_2=0.1330$</td><td>$m_3=0.0041$</td><td>$m_4=0.0147$</td></tr>
<tr><td colspan="5">Orientation of the magnetic field relative to the axes.</td></tr>
<tr><td>Parameter</td><td>C_1</td><td colspan="2">1° from the bisectrix axis in the binary-bisectrix plane</td><td>11.5° from the trigonal axis in the trigonal-bisectrix plane</td></tr>
<tr><td>m_{cA}^*</td><td>1.814×10^{-2}</td><td colspan="2">2.099×10^{-3}</td><td>6.292×10^{-3}</td></tr>
<tr><td>E_g/m_{cA}^*</td><td>0.86</td><td colspan="2">7.43</td><td>2.49</td></tr>
<tr><td>m_{cB}^*</td><td>2.391×10^{-3}</td><td colspan="2">4.002×10^{-3}</td><td>1.558×10^{-2}</td></tr>
<tr><td>E_g/m_{cB}^*</td><td>6.52</td><td colspan="2">3.90</td><td>1.00</td></tr>
<tr><td>m_{cC}^*</td><td>2.391×10^{-3}</td><td colspan="2">4.238×10^{-3}</td><td>1.558×10^{-2}</td></tr>
<tr><td>E_g/m_{cC}^*</td><td>6.52</td><td colspan="2">3.68</td><td>1.00</td></tr>
</table>

Note: The values of the cyclotron masses m_{cA}^*, m_{cB}^*, and m_{cC}^* corresponding to sections *A, B,* and *C* of the electronic ellipsoids and the components m_i of the effective-mass tensor are given in units of the free-electron mass; E_g is in eV; C_1 is the binary axis.

fields ranging from 2 to 2.8 T (Fig. 1), as well as the somewhat different behavior of the rise and fall of the magneto-optic oscillations in fields below 1 T. The agreement between the experiment and the calculations is improved when the misorientation of the crystal from the magnetic field of about 0.5°, which is unavoidable in the experiment, in the binary-bisectrix plane is taken into account.

The experimental and computed magneto-optic spectra with the magnetic field tilted by 1° from the bisectrix in the binary-bisectrix plane are shown in Fig. 3. This causes the appearance of three series of oscillations with comparable amplitudes, which leads to beats. In the binary direction there are no beats, because the difference in the periods of the oscillations due to the light and heavy electrons.

Modeling of the shape of the experimental curve leads to equal values of the relaxation time in the binary and bisectrix directions and to different values of the lattice permittivity. This indicates a relatively independent existence of electronic subsystems associated with each constant-energy quasiellipsoid.

The results of the modeling of the lineshape of the magneto-optic spectrum, when the magnetic field made in the bisectrix-trigonal plane an angle of 11.5° with the trigonal axis, are shown in Fig. 4. In accordance with the ideas about the structure of isoenergetic surfaces, in the experiment intense oscillations are observed from the light cyclotron mass and less intense and split oscillations are observed from the two heavier cyclotron masses. The splitting could be due either to the small (of the order of 1°) disorientation of the crystal relative to the bisectrix–trigonal plane or the presence of spin splitting of the Landau levels. Quantitative estimates show that the contributions of both effects are approximately equal. A final conclusion can be drawn from an analysis of the lineshape of the experimental curve on the basis of the McClure–Choi model of the energy spectrum, in which the nonmirror nature of the bands is taken into account and the spin-parameter tensor of the valence and conduction bands is introduced.

The parameters of the energy spectrum of the L electrons were obtained from a description of the line shape (Table I). Parameters which are similar and, as one can see from Table II, close in value were determined from the experimental data on the positions with respect to the field of the maxima of the oscillations on the basis of the McClure–Choi dispersion relation. Comparison with the data obtained from the oscillation effects[3] at liquid-helium temperature

TABLE II. Parameters of the energy spectrum of the L electrons of bismuth at 80 K, calculated from the positions of the maxima of the oscillations in a magnetic field in the McClure–Choi model and according to Ref. 3 at 4.2 K.

<table>
<tr><td>Temperature</td><td>Q_{11}</td><td>Q_{22}</td><td>Q_{33}</td><td>a22v</td><td>a_{22}^c</td></tr>
<tr><td>80 K</td><td>0.429</td><td>0.030</td><td>0.327</td><td>1.1</td><td>0.60</td></tr>
<tr><td>4.2 K (Ref. 3)</td><td>0.454</td><td>0.034</td><td>0.034</td><td>1.0</td><td>0.63</td></tr>
<tr><td colspan="6">Orientation of the magnetic field relative to the axes.</td></tr>
<tr><td>Parameter</td><td>C_1</td><td colspan="2">1° from the bisectrix axis in the binary-bisectrix plane</td><td colspan="2">11.5° from the trigonal axis in the trigonal-bisectrix plane</td></tr>
<tr><td>E_g/m_{cA}^*</td><td>–</td><td colspan="2">7.59</td><td colspan="2">2.45</td></tr>
<tr><td>E_g/m_{cB}^*</td><td>6.52</td><td colspan="2">3.93</td><td colspan="2">–</td></tr>
<tr><td>E_g/m_{cC}^*</td><td>6.52</td><td colspan="2">3.71</td><td colspan="2">–</td></tr>
</table>

Note: The values of the parameters Q_{ii} and $a_{ii}^{v,c}$ are given in the atomic system of units, the cyclotron masses m_{cA}^*, m_{cB}^*, and m_{cC}^* are given in units of the free-electron mass, E_g is in eV, and C_1 is the binary axis.

(Table II) shows that the parameters are temperature-dependent.

The application of the SSL method and the use of a comparatively high quantum energy and pulsed magnetic fields made it possible to obtain oscillations associated with interband optical transitions on Landau levels. A number of experimentally observed features are explained on the basis of the description of the lineshape. The band parameters of bismuth at 80 K were calculated. Comparison with existing data obtained from oscillation effects at 4.2 K shows that the parameters are temperature-dependent. The charactistic features of the experiment which were listed above make it possible to investigate not only bismuth but also the solid solutions Bi_{1-x}–Sb_x and narrow-gap semiconductors.

[1] O. V. Konadkov and K. G. Ivanov, Fiz. Tverd. Tela (Leningrad) **32**, 290 (1990) [Sov. Phys. Solid State **32**, 164 (1990)].

[2] M. P. Vecchi, J. R. Pereira, and M. S. Dresselhaus, Phys. Rev. B **14**, 298 (1976).

[3] N. B. Brant, G. I. Golysheva, Nguen Min Thu, M. V. Sudakov, K. N. Kashirin, and Ya. G. Ponamarev, Fiz. Nizk. Temp. **13**, 1209 (1987) [Sov. J. Low Temp. Phys. **13**, 683 (1987)].

Translated by M. E. Alferieff

Intense photoluminescence of porous layers of SiC films grown on silicon substrates

A. M. Danishevskiĭ, V. B. Shuman, E. G. Guk, and A. Yu. Rogachev

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

(Submitted June 3, 1996; accepted for publication June 5, 1996)
Fiz. Tekh. Poluprovodn. **31**, 420–424 (April 1997)

Luminescing porous layers have been prepared on SiC films grown on silicon substrates. The intensity of the photoluminescence increases very strongly as a result of electrolytic oxidation of porous layers. The spectrum of the pulsed photoluminescence consists of a series of overlapping bands from 1.8 to 3.3 eV. Investigations of the initial SiC films showed that they are nonstoichiometric and strongly disordered. Nonetheless, the intensity of the photoluminescence of the oxidized porous layers is much higher than can be obtained from correspondingly treated SiC crystals or crystalline films. © *1997 American Institute of Physics.* [S1063-7826(97)00804-1]

Different electrochemical methods for preparing porous SiC layers based on SiC crystals and crystalline films grown on 6H-SiC substrates and investigations of their structure, phonon spectra and photoluminescence spectra (PL) were reported in Refs. 1–8. Specifically, in Refs. 7 and 8 it was hypothesized that the main luminescence arises from surface layers of shallow voids, where electrochemical etching causes breaking of interlayer bonds in the structure of the 6H-SiC crystal and thin shells of voids, consisting of finely dispersed cubic phase SiC, form. The conclusion that a cubic phase is present in the porous layer was drawn on the basis of the Raman scattering spectra, and it was later confirmed by direct electron-microscopic observations. The character of the PL spectrum in this case is determined by quantum-size effects in a finely dispersed cubic SiC phase. The intensity of the PL of the porous layer increased by up to a factor of 200–300 compared with the initial intensity at 300 K, but the maximum in the continuous PL spectra was located, as a rule,[2–6] in the region 2.4–2.5 eV, which for the 6H-SiC crystals on which the porous layers were formed corresponds to a region of PL with participation of defects, which is why in Ref. 5 the intense luminescence was explained as luminescence associated with a sharply increased concentration of defects in the porous layer. In the pulsed time-resolved PL spectra,[8] the maximum of the wide PL band shifted to 2.8 eV, and a structure of the band was clearly observed. Together with other factors, this confirms the mechanism proposed in Refs. 7 and 8.

This paper reports the results of experiments performed on SiC films obtained by chemical deposition from the gas phase (CVD) on a 76-mm-diam silicon substrate.[9] X-ray measurements showed that these films were not single-crystalline and that they consisted of disordered regions less than 10 μm in size. The initial films exhibited virtually no PL.

To determine the band gap and hence the polytype on the given SiC films (their silicon substrate was etched off), the transmission spectra were measured near the band edge. The curves $(\alpha h\nu)^{0.5}=f(h\nu)$, where α is the absorption coefficient, which were obtained on two films of different thickness ($d=3.5$ and 2.15 μm), are shown in Fig. 1. It should be noted, however, that the films differed not only in thickness but also with respect to the deposition conditions. It is evident from the figure that these curves can be approximated quite well by two straight lines with different slopes, from which the characteristic energy gaps near the indirect absorption edge can be easily determined. They turned out to be ~1.88 eV for sample Sh49 and ~1.76 eV for sample Sh65. Such small, even for cubic SiC, energy gaps can be explained only by the fact that the films not only are strongly disordered, but they are also nonstoichiometric. This conclusion is confirmed to some extent by the fact that with electrochemical etching and oxidation, performed in order to prepare porous luminescing layer, the etching and oxidation processes were much more rapid on the films than on the SiC crystals or crystalline films.

The Raman scattering measurements (the silicon substrate of the films was etched off) confirmed the presence of silicon nanocrystallites (513 and 517 cm^{-1} bands) in the film studied and also the presence of silicon–carbon bonds which do not correspond to stoichiometry (532 cm^{-1} band).[10] The weak 791 and 955-cm^{-1} bands and the 971-cm^{-1} band, which are relatively close to the LO photon energy in 6C-SiC, were observed in the Raman spectra for the initial Sh65 film. For sample Sh49 three partially overlapping weak bands were also observed near the indicated energy — 908, 928, and 955 cm^{-1}. A very weak TO phonon band at 795 cm^{-1} was also recorded.

The infrared (IR) reflection spectra measured on the initial films (Figs. 2 and 3) nonetheless showed the presence of a phonon structure characteristic of silicon carbide. For the same sample Sh49, the dropoff of the reflection band near the TO phonon is not rapid enough and has an additional weak band near 860 cm^{-1}. Figures 2 and 3 also show the IR reflection spectra from porous and specially oxidized porous layers. We see that when the porous layer is formed, the IR reflection spectra of the films change relatively little. However, the oxidation of the surface results in very large changes in the structure of these spectra. The reflection band edge corresponding to the LO phonon frequency seemingly ‘‘collapses’’ and bands appear near 1080–1100 cm^{-1}. Figure 3 shows for comparison the IR reflection spectrum (curve *4*) from a thick (0.5 μm) thermal (obtained at 1100 °C) oxide on silicon. A band is present there at ~1130 cm^{-1}.

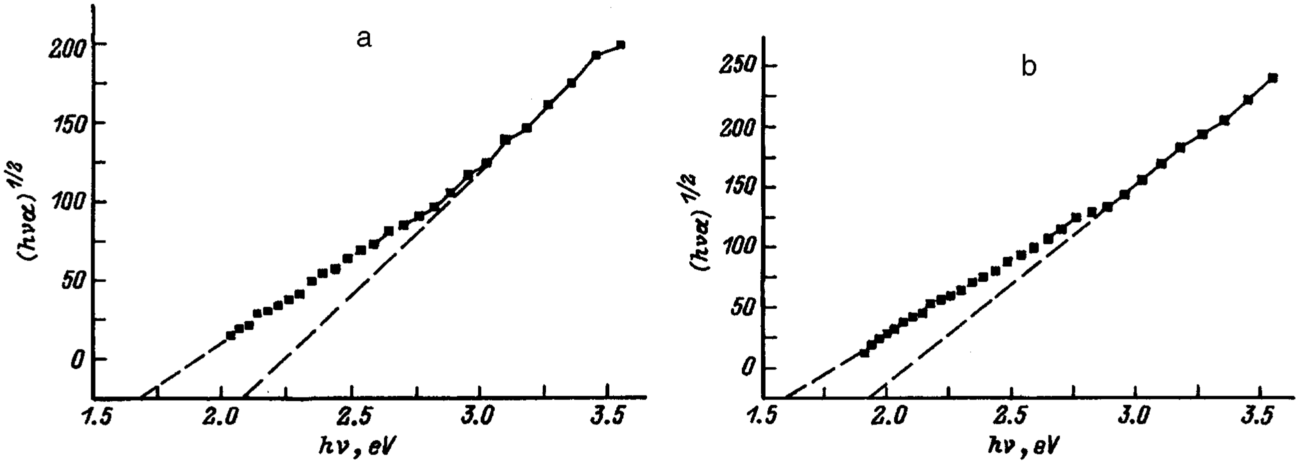

FIG. 1. Edge absorption spectra of the initial experimental SiC films freed of silicon substrates: a — Sample Sh49, $d = 3.5$ μm; b — sample Sh65, $d = 2.15$ μm.

It thus follows that the bands 1080–1100 cm^{-1} noted above on the experimental SiC films (spectra *3* and *4* in Fig. 2 and spectrum *3* in Fig. 3) probably correspond to a nonstoichiometric oxide, often denoted as SiO_x.

After a porous layer was produced on the Sh49 and Sh65 films (anodic etching in a standard water-alcohol mixture with HF), virtually no luminescence of the samples at 300 K was observed in an investigation in a luminescence micro-

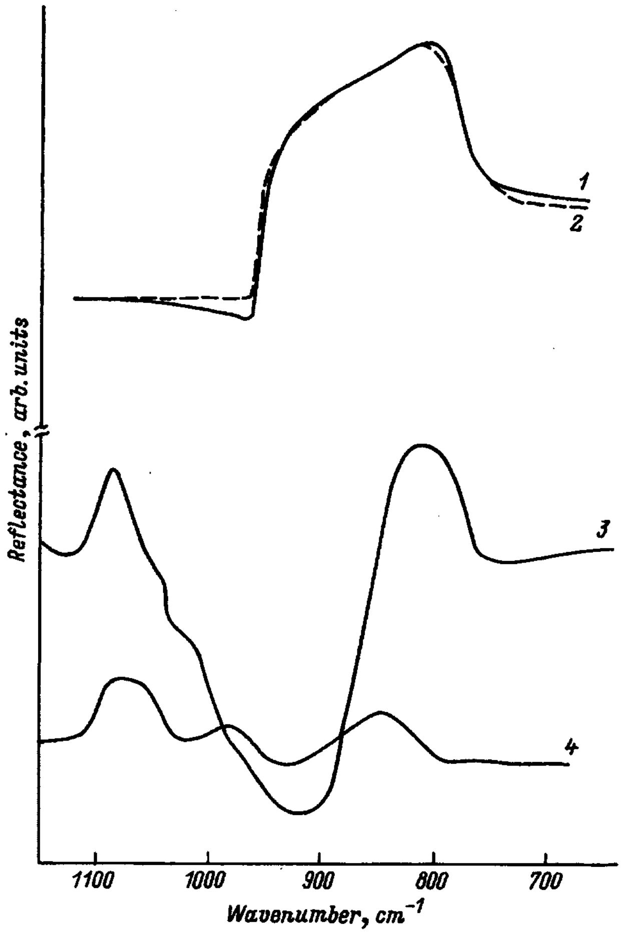

FIG. 2. IR reflection spectra of Sh65 films: *1* — Initial sample; *2* — film after anodic etching; *3* — porous film oxidized electrochemically in a water solution of HCl for 4 min; *4* — porous film oxidized for 8 min.

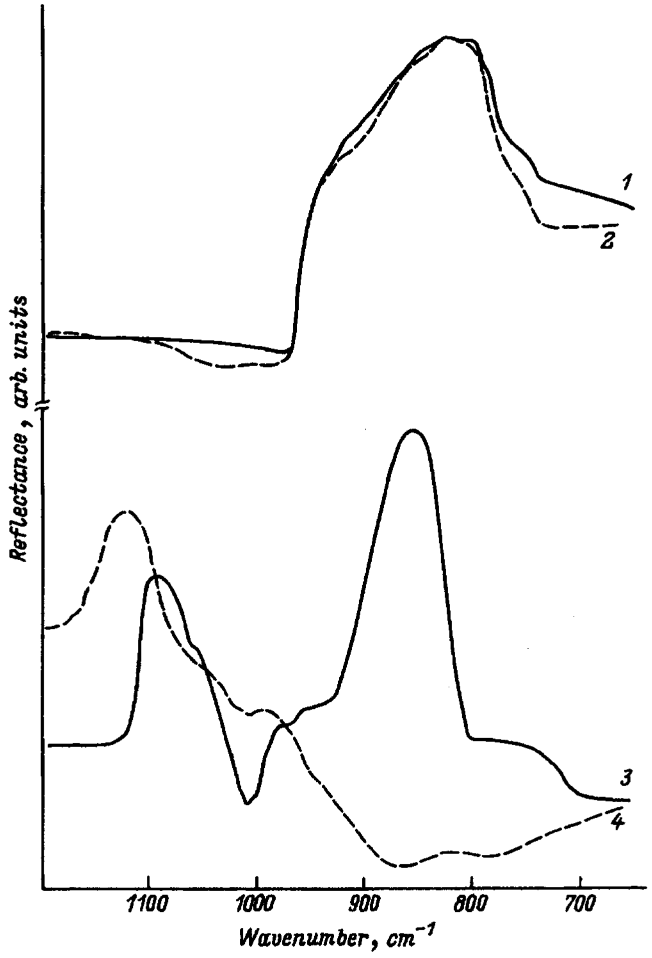

FIG. 3. IR reflection spectra of Sh49 films: *1* — Initial sample; *2* — film after production of a porous layer; *3* — porous film oxidized for 6 min; *4* — reflection spectrum of thermal oxide SiO_2 on silicon.

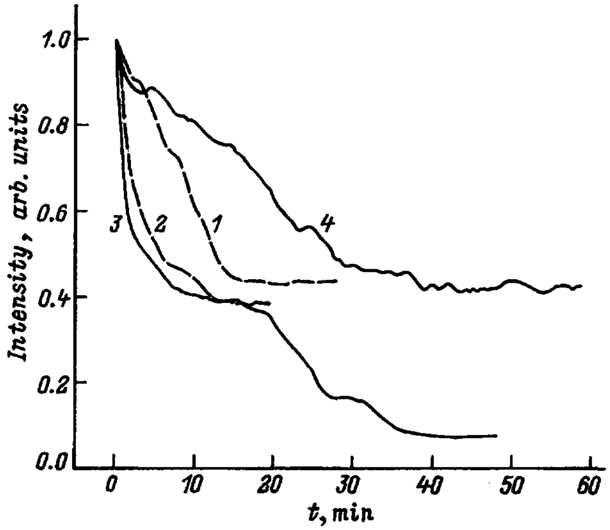

FIG. 4. PL degradation characteristics. *1, 2* — Sh65 films, oxidation for 8 min (*1*) and 4 min (*2*); *3, 4* — Sh49 films, oxidation for 6 min (*3*) and 7 min (*4*).

scope (Lyumam R-2). However, it did arise after even a brief (1–2 min) electrolytic oxidation of the samples (in a water solution of hydrochloric acid) was performed in order to passivate the surface of the porous layer, and it intensified appreciably with increasing oxidation time.1) It was noted that the surface of the porous layer has a granular structure, and the most intense luminescence was observed from grain boundaries (twins) of size in the range 10–15 μm. Its intensity was much higher than that of the porous layers obtained on SiC crystals or crystalline films, and its brightness approached that of the best samples of porous silicon. In contrast to the luminescence observed for silicon, the luminescence was of a white-blue color.

Under the action of intense ultraviolet (UV) radiation from a pulsed laser, the PL of the obtained layers degraded very slowly to a definite value, if a sufficiently dense oxide could be obtained. The quantitative measurements of the PL degradation rate were performed at temperature $T = 90$ K and excitation by radiation from a pulsed nitrogen laser with photon energy $h\nu = 3.68$ eV, pulse duration $t_i = 10$ ns, pulse repetition frequency $f = 50$ s^{-1}, and intensity $\sim 10^6$ W/cm^2. Figure 4 shows the curves of the PL intensity near the spectral maximum of the band measured at the moment corresponding to the maximum of the PL pulse. The maximum PL intensities are scaled here to a single value in order to compare the PL degradation rates with different durations of the oxidation of the samples.

As the electrochemical oxidation time increased, the PL intensity increased substantially within certain limits and the rate of degradation of the PL under the action of the UV laser radiation decreased substantially. For sample Sh49 with a large energy gap (a more stoichiometric film), the rate of degradation under the same oxidation conditions was lower and the PL was weaker than for the Sh65 sample.

The increase in the intensity of the PL from the porous layer with a denser and more stoichiometric oxide is due, naturally, to the decrease in the surface recombination rate. The PL degradation rate is determined by the rate of increase of the density of states in the band gap of the semiconductor at its boundary with the oxide under the action of the UV laser radiation. Processes of this kind have been studied many times for a Si–SiO$_2$ boundary.[11,12] Different models of defect formation were proposed. They all attribute this process in one way or another to the presence of hydrogen near the boundary and the destruction of the silicon–hydrogen bonds by the UV radiation. On all films which we investigated, after the production of the porous layer the bands of the silicon–hydrogen bonds (~ 2070 cm^{-1}) were also observed in the Raman spectra. It can be assumed, therefore, that the processes leading to the production of defects near the interface in our samples are largely similar to those investigated on Si–SiO$_2$ structures.

Figure 5 shows the pulsed time-resolved PL spectra obtained at $T = 90$ K with excitation by nitrogen laser pulses at times corresponding to maximum luminescence. The excitation intensity is $\sim 10^6$ W/cm^2. To decrease the manifestation of nonstationary degradation effects in the spectra, before the measurements were preformed the samples were irradiated with laser radiation for a period of time required to reach an

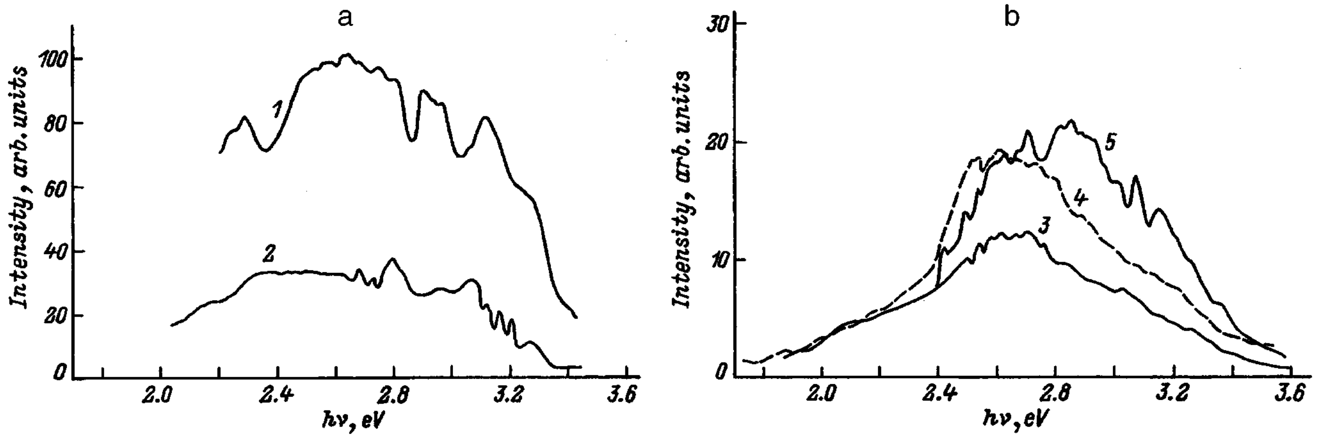

FIG. 5. Pulsed time-resolved PL spectra. a — Sh65 films, oxidation time 8 min (*1*) and 4 min (*2*). b — Sh49 films, oxidation times 6 min (*3*), 7 min (*4*), and 9 min (*5*); the spectrum 5 was obtained with a three times lower excitation intensity than the spectra *1–4*.

approximately stationary section of the degradation characteristic. This time was equal to 40 min for sample Sh49 and 15 min for sample Sh65.

Figure 5a shows two PL spectra obtained from porous layers (Sh65) with different oxidation times (4 and 8 min). A number of wide, partially overlapping bands, encompassing the region from red to near-UV radiation ($\sim$3.5 eV), is observed in the spectra. Figure 5b shows two PL spectra (*3, 4*) for the Sh49 layers with different oxidation times (6 and 7 min); the spectra were obtained under conditions similar to those of spectrum 5 measured with a three times lower excitation intensity. In the latter case, the oxidation time was equal to 9 min.[2)] The radiation of these layers is weaker, and its maxima are localized near 2.5–2.95 eV. The UV luminescence ''tails'' are also present here, but compared with the maximum values they are weaker than for spectra *1* and *2* (Fig. 5a). It is interesting that despite the fact that the spectra shown here are different, most features appearing in them coincide in energy. This fact and the presence of violet and UV sections of the PL spectrum attest quite convincingly to the fact that quantum-well effects have a determining influence on the character of the radiative recombination in these porous layers.

The possible assumption that the UV radiation of the oxide, which coats the porous surface of the layers, contributes to the UV section of the spectrum is refuted by the fact that as the oxidation time increases, the intensity of the radiation of the layer increases in almost the entire spectral range. The small asymmetry of the curves *3, 4,* and *5* and the rise of the short-wavelength wing of the PL band with increasing oxidation time in Fig. 5b are, in our opinion, due to the fact that when the thicker oxide is formed, the ''bridges'' between the pores become smaller and the observed behavior of the spectra can be explained by the influence of quantum-well effects.

The different broadening of the spectra in the short-wavelength direction and the different PL intensities for samples Sh65 and Sh49 are explained in our opinion primarily by the fact that when the porous layer is produced, the etch processes in these samples do not occur in the same manner. The more nonstoichiometric sample Sh65 probably dissolves more strongly. It acquires a more extended surface and smaller bridges between the voids, which is what determines the parameters of the PL. Apparently, the oxidation processes on samples Sh65 and Sh49 also occur at a different rate. As a result, a denser oxide can arise on sample Sh65 under the same conditions and with the same oxidation times.

Figure 6 shows the characteristics of the dropoff of the PL pulses measured under the same conditions as in the spectra presented above. The figure also presents for comparison a similar curve for the scattering laser light, measured under the same conditions, in order to show the temporal resolution of the recording system. One can see that in the measured time interval these curves of the PL and the scattered light are virtually identical. Such short dropoff times of the PL pulses and high luminescence intensity indicate the twin character of the radiative recombination, which also confirms the model of volume recombination in small isolated crystallites or bridges between voids, but is hardly characteristic of the oxide radiation.

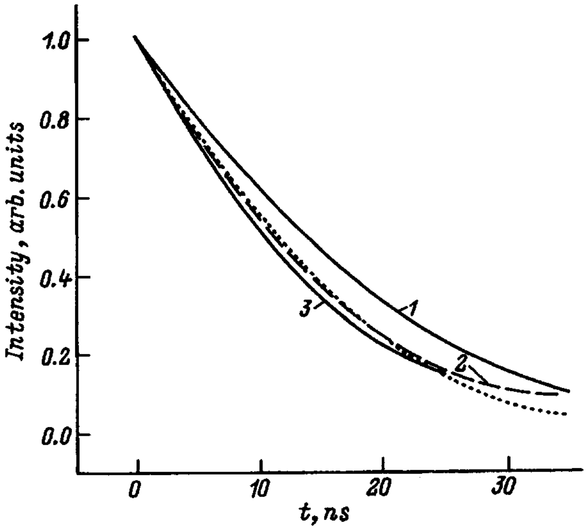

FIG. 6. Temporal characteristics of the dropoff of the PL pulses. *1, 2* — Sh65 films, oxidation times 8 min (*1*) and 4 min (*2*). *3* — Sh49 film, oxidation time 7 min. *4* — characteristics of the dropoff of the scattered laser radiation pulse.

In conclusion, we note that in our study we found a promising and technologically simple system, in which very intense PL with a very wide spectrum was obtained. It was shown that from the standpoint of the radiative characteristics it has advantages over both porous silicon and porous silicon carbide obtained on single-crystalline SiC and can be prepared by different technological techniques. Since the films are deposited on standard silicon substrates, in the case where efficient electroluminescence is obtained we will have sources formed in the regimes of planar silicon technology.

We thank M. P. Shcheglov for performing the x-ray measurements on one of the initial samples, and V. E. Chelnokov for a discussion and support.

This work was supported by the Russian Fund for Fundamental Research grant 95-02-04115 and a grant on the theme ''Nanostructures in semiconductors,'' as well as by the University of Arizona (USA).

[1)] Oxidation was also conducted on porous layers prepared on the basis of 6*H*-SiC single-crystalline films, but in this case it was much less effective.

[2)] The presence of an oxide leads to stresses in the film, and under longer oxidation periods the film separates from the substrate and crumbles because of its brittleness.

---

[1] J. S. Shor, I. Grimberg, B. Z. Weiss, and B. D. Kurtz, Appl. Phys. Lett. **62**, 2836 (1993).

[2] T. Matsumoto, J. Takahashi, T. Tamaki, T. Futaki, H. Mimura, and Y. Kanemitsu, Appl. Phys. Lett. **64**, 226 (1994).

[3] A. O. Konstantinov, C. I. Harris, and E. Yanzen, Appl. Phys. Lett. **65**, 2699 (1994).

[4] J. S. Shor, L. Bemis, A. D. Kurtz, M. Mcmillan, W. J. Choyke, I. Grimberg, and B. Z. Weiss in *Silicon Carbide and Related Materials,* ed. by M. G. Spencer, R. P. Devaty, J. A. Edmond, M. Asifkhan, R. Kaplan, M. Rahman [Inst. Phys. Conf. Ser, N 137 (IOP, Bristol, 1993) p. 193].

[5] A. A. Lebedev, A. A. Lebedev, and Yu. V. Rud', Pis'ma Zh. Tekh. Fiz. **21**, 64 (1995) [Tech. Phys. Lett. **21**, 117 (1995)].

[6] A. O. Konstantinov, A. Henry, C. I. Harris, and E. Janzen, Appl. Phys. Lett. **66**, 2250 (1995).

[7] A. A. Danishevskiĭ, V. B. Shuman, A. Yu. Rogachev, and P. A. Ivanov, Zh. Tekh. Fiz. **29**, 2122 (1995) [*sic*].
[8] A. M. Danishevskiĭ, V. B. Sjuman, A. Yu. Rogachev, E. G. Guk, P. A. Ivanov, and A. A. Mal'tsev, Zh. Tekh. Fiz. **30**, 1064 (1996) [*sic*].
[9] I. M. Baranov, N. A. Belov, V. A. Dmitirev, N. G. Ivanova, T. S. Kondrat'ev, I. P. Nikitina, V. E. Chelnokov, V. F. Shatalov, and R. N. Erlich, Pis'ma Zh. Tekh. Fiz. **15**(12), 50 (1989) [*sic*].
[10] Pham. V. Huong, Diamond and Related Mater. **1**, 33 (1991).
[11] K. G. Druijf, J. M. M. de Nijs, E. van der Drift, E. H. A. Granneman, and P. Bulk, J. Appl. Phys. **78**, 306 (1995).
[12] K. G. Druijf, J. M. M. de Nijs, E. van der Drift, E. H. A. Granneman, and P. Bulk, J. Appl. Phys. **79**, 1505 (1996).

Translated by M. E. Alferieff

Injection enhancement of photocurrent in polycrystalline silicon p^+-n-n^+ structures

R. Aliev

Institute of Electronics, Uzbekistan Academy of Sciences, 700143 Tashkent, Uzbekistan

(Submitted February 20, 1996; accepted for publication June 7, 1996)
Fiz. Tekh. Poluprovodn. **31**, 425–426 (April 1997)

Injection enhancement of photocurrent in structures of the p^+-n-n^+ type fabricated on the basis of films of polycrystalline silicon grown on commercial silicon substrates has been investigated. The questions of obtaining injection photodetectors and other bistable elements with the aid of α-irradiation and heat treatment are discussed. *© 1997 American Institute of Physics.* [S1063-7826(97)00904-6]

In contrast to single-crystalline semiconductors, in polycrystalline semiconductors the electrical properties can be additionally controlled by varying the grain orientation, grain size, and the electronic states at intergrain boundaries (IGBs).[1] The changes produced in the electronic states at IGBs during high-temperature operations give rise to the appearance of a S-shaped current-voltage characteristic (IVC) under forward biases ($U>0$) in polycrystalline silicon structures with a $p-n$ junction.[2] In Ref. 2 it was assumed that it is possible to fabricate injection devices with a S-shaped IVC on the basis of such structures.

The objective of this work is to check this possibility experimentally for the example of polycrystalline silicon p^+-n-n^+ based photodetectors with internal current enhancement. The p^+-n-n^+ structures were obtained by the technology described in Ref. 2, which proposes the use of diffusion or ion implantation.

Structures with a diffused $p-n$ junction, which have a switch-on voltage $U_{d0}\leqslant 1.2$ V in the forward direction, were illuminated with white light on the junction side and the IVC were measured in the forward and reverse directions at ambient temperature ~30 °C. The samples were photosensitive on the section of the IVC $U<U_{d0}$ in the forward direction. Virtually no injection enhancement of the photocurrent was observed.

The experiments performed on structures with a $p-n$ junction obtained by ion doping showed similar results only after heat-treatment at 800 °C for $\geqslant 60$ min. Although the switch-on voltage in such structures is almost quadrupled by heat treatment,[3] the photocurrent gain does not exceed 1.5–2. The maximum photosensitivity in the forward bias direction is achieved with heat treatments of 90–120 min.

These results confirm that the appearance of a S-shaped IVC is associated with the shunting of the $p-n$ junction by an n^+channel along a IGB in accordance with the model proposed in Refs. 2 and 3 . However, stable S-shaped IVCs are not a sufficient condition for injection enhancement of current. According to Ref. 4 , this also requires modulation of the conductivity of the n base of the structure, which is possible only if the resistance of the base is high and the thickness d is much greater than the diffusion length L of the current carriers, $d\gg L$.

Another way to achieve injection enhancement of current is to dope the base with impurities, such as Au, Zn, and Hg, which give deep energy levels in the band gap of silicon. This ensures a τ mechanism of injection enhancement of current (see Ref. 4).

We employed irradiation of the experimental structures with a continuous flux of α particles with energy $E=5$ MeV and flux density $j=10^7-10^8$ $\text{cm}^{-2}\cdot\text{s}^{-1}$.

As the experiments show, the IVC of the structures changes sharply as the the α-irradiation dose D increases. A gradual decrease of the dark forward current in the region $U<U_{d0}$ is accompanied by a rapid increase of the dark switch-on voltage U_{d0} (see Table I). This attests to the fact that compensation conditions are reached in the base. Figure 1 shows the forward branches of the IVC of a 5×5-mm^2 sample irradiated with α-particles with a dose of 10^{12} cm^{-3}. The branches were measured in the absence of illumination (curve *1*) and with illumination with power density $P_0\simeq 100$ mW/cm^2 (curve *2*). The photosensitivity of the structure increases with U_{d0} (see Table I). Therefore, the photocurrent gain, defined as $M=I_{\text{ph}}/I_0$ — the ratio of the photocurrent $I_{\text{ph}}=I-I_d$ (I is the total current, and I_d is the dark current) for $0<U<U_{d0}$ to the photocurrent I_0 with a small reverse voltage $U\simeq 0$, increases. The increase in M with an increase in U can be expressed by the approximate Miller relation for avalanche multiplication of photocurrent:[5]

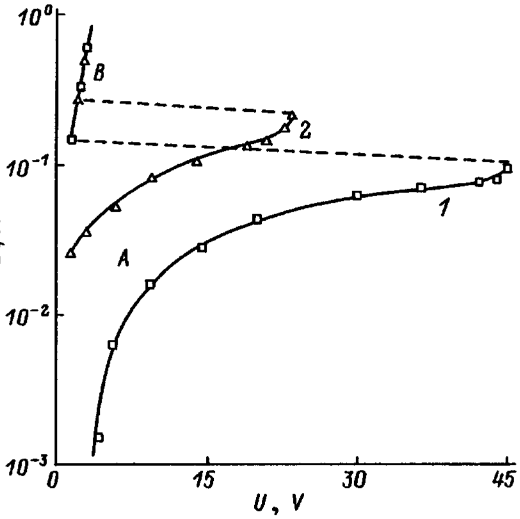

FIG. 1. I–V characteristic of an unilluminated (*1*) and white light-illuminated (*2*) polycrystalline silicon pnn structure. The sections *A* and *B* correspond to nonswitched-on and switched-on states of a bistable IVC.

TABLE I. Basic parameters of one of the experimental p^+–n–n^+ structures and their variation under α irradiation.

D, cm^{-3}	l_0, mA ($U<0$)	M (maximum values)	I_0, mA ($U=0.3$ V)	I_{d0}, mA ($U=U_{d0}$)	U_{d0}, V	U_{ph0}/U_{d0},%
0	3.0	1	0.06	–	–	–
10^9	2.9	2	0.04	0.45	0.9	10
10^{10}	2.4	12	0.14	200	–	–
10^{11}	1.7	80	0.035	5.8	6	15
10^{12}	0.7	230	0.01	80	45	45
10^{13}	0.34	235	0.005	88	58	48

$$M=[1-(U/U_{d0})^n]^{-1}.$$

The exponent n in the experimental structures is relatively small and equals 0.01–0.02, which distinguishes the structures sharply from avalanche photodiodes, where $n \geqslant 1$.

Illumination of the structure decreases the switch-on voltage from U_{d0} to some value U_{ph0}; here the difference $U_{d0}-U_{\text{ph0}}$ increases with increasing illumination intensity. For the structure whose IVC is shown in Fig. 1, $U_{d0}=45.5$ V and $U_{\text{ph0}}=23.8$ V. There exists a critical illumination density, i.e., a photogeneration level G_{cr}, at which a S photodiode transforms from the initial state A into a switched-on state B. When the illumination is removed, the structure remains stable in the switched-on regime. The structure can be switched back into the state A only by decreasing the voltage to some threshold value or lower, $U \leqslant U_{\text{th}}$. For the structures investigated the threshold voltage is 0.6–0.8 V. The critical illumination density is 200–250 mW/cm^2 for α-irradiated structures with $U_{d0} \simeq 40-50$ V. In other words, these structures are so-called ‘‘bistable elements’’ which are controlled by either voltage or illumination.

To confirm the mechanism of the process which leads to the effects described above and which is based on the shunting of a p–n junction along IGBs[2] and on the compensation in the volume of the base, we subjected the structures to thermal annealing in regimes which remove the α-irradiation-induced radiation defects. Thermal annealing (TA) was conducted at $T=100-800\,^\circ$C for 60 min in vacuum. It was shown that as the TA temperature increases from 100 to 400 °C, U_{d0} gradually decreases on the average from 45 to 24 V almost without any change in the dark switch-on current I_{d0}. A further increase of the TA temperature in the range 400–600 °C results in a sharp, almost an order of magnitude, decrease of I_{d0} with a nonuniform variation of U_{d0}, right up to an increase in U_{d0} to 33 V at $T=600$ °C. Thermal annealing at $T=800$ °C completely restores the initial IVC with $U_{d0} \simeq 1.2$ V. In other words, the defects produced by α irradiation, which put the n base into a compensated state, can be destroyed completely by TA. On the one hand, this confirms the mechanism suggested for the process and, on the other, it illustrates the high reproducibility of the technology for obtaining structures with a S-shaped IVC.

In summary, there exists a real possibility of fabricating injection photodetectors with a gain of $\simeq 250$ on the basis of a polycrystalline p^+–n–n^+ structure on low-resistance metallurgical-silicon substrates. These structures can be used as bistable elements, a photoswitch, and a photoresistor, as well as an integral part of a phototransistor with one injection junction.

[1] V. K. Georgiev, L. I. Popova, L. E. Polyak, and L. K. Fionov, Noverkhnost’. Fizika, Khimiya, Mekhanika, No. 9, 5 (1990).

[2] M. S. Saidov, B. M. Abdurakhmanov, R. Aliev, and A. S. Saidov, Fiz. Tekh. Poluprovodn. **30**, 128 (1996) [Semiconductors **30**, 74 (1996)].

[3] R. Aliev, B. M. Abdurakhmanov, and R. R. Bilylov, Interface Sci. (1996).

[4] I. M. Vikulin and V. I. Stafeev, *Physics of Semiconductor Devices* [in Russian], Sov. radio, Moscow (1980).

[5] S. M. Szy, *Physics of Semicondcutor Devices*, Wiley, N. Y., 1969.

Translated by M. E. Alferieff

Particle scattering times on one-dimensional potential barriers

N. L. Chuprikov

Siberian Physicotechnical Institute at Tomsk State University, 634050 Tomsk, Russia

(Submitted March 21, 1996; accepted for publication June 13, 1996)
Fiz. Tekh. Poluprovodn. **31**, 427–431 (April 1997)

The tunneling times for particles whose initial state is described by wave packets of a general form is determined on the basis of a packet analysis. It is shown that the unphysical nature of the results obtained previously by numerical modeling of the motion of wave packets in one-dimensional structures is due to the incorrect interpretation of the packet formalism. Wave-packet tunneling times observed in a numerical experiment are not at all the particle-tunneling times. © *1997 American Institute of Physics.* [S1063-7826(97)01004-1]

1. INTRODUCTION

In this paper we estimate the time-dependent parameters of electron tunneling in semiconductor heterostructures. At present, there is no clear unequivocal solution of this problem even in its simplest formulation — for tunneling of one particle in one-dimensional systems. A study of the dynamics of Gaussian wave packets (GWPs)[1–7] gave such unusual results that their interpretation still elicits serious disputes.[7] For example, it was found for square opaque barriers that the maximum of the transmitted packet can move away from the right-hand boundary of the barrier before the maximum of the incident packet reaches the left-hand boundary of the barrier.

Therefore, in packet analysis there is a serious problem of correctly transferring the properties of wave packets to a particle. It must be acknowledged that this problem for wave packets of a general form remains unsolved. The results of Refs. 1 and 2 are valid only for wave packets which are infinitely narrow in k space. In the general case, as noted in the review in Ref. 7, packet analysis leaves much room for disagreements. Alternative approaches, whose authors reject packet analysis for studying particle tunneling, are therefore widely used at present. Without going into the details of these approaches (see Ref. 8 and also the reviews in Refs. 6 and 7), we note only that their development also encounters serious difficulties.

In the present paper we solve the problem of determining the temporal parameters of particle tunneling on the basis of the transfer matrix method.[9,10] It is shown that the unphysical nature of the results of packet analysis occurs as a result of an incorrect interpretation of the quantum mechanics: The wave-packet tunneling times observed in numerical modeling are by no means the particle-tunneling times. In this paper, analytical expressions are obtained for the tunneling parameters of particles whose initial states are described by arbitrary wave packets. In the concluding section we compare our approach to the method of Ref. 3 where wave packets of a general form are also studied.

2. WAVE PACKETS

Let an electron be incident from the left side on a static potential barrier located in the interval $[a,\ b]$, $a>0$. The particle is free to the left and right of the barrier. We write the matched particular solution of the corresponding time-dependent Schrödinger equation on the basis of the transfer matrix method.[9,10] To the left of the barrier the solution is a superposition of the incident and reflected waves

$$\Psi_{\text{left}}=[\exp(ikx)+\varphi_{\text{ref}}(k)\exp(-ikx)]\exp[-iE(k)t/\hbar], \tag{1a}$$

where

$$\varphi_{\text{ref}}(k)=\sqrt{R(k)}\exp[i(J-F+2ka-\pi/2)].$$

For $x\geqslant b$ the solution is the transmitted wave

$$\Psi_{\text{right}}=\varphi_{\text{tr}}(k)\exp[i(kx-E(k)t/\hbar)], \tag{1b}$$

where

$$\varphi_{\text{tr}}(k)=\sqrt{T(k)}\exp[i(J-kd)].$$

Here $R=1-T$, $d=b-a$, $E(k)=\hbar^2k^2/2m$, m is the particle mass, t is the time, and the transmission coefficient T and the phase parameters J and F are functions of k. Recurrence relations for calculating these quantities and their derivatives with respect to energy are given in Refs. 9 and 10.

Together with Eqs. (1a) and (1b), the solutions of the Schrödinger equation are also the functions

$$\Psi_{\text{left}}(x,t)=\frac{1}{2\pi}\int_{\infty}^{\infty}[\Psi_{\text{inc}}(k,t)+\Psi_{\text{ref}}(k,t)]\times\exp(ikx)dk \quad \text{for} \quad x\leqslant a, \tag{2a}$$

where

$$\Psi_{\text{inc}}(k,t)=cA(k)\exp[-iE(k)t/\hbar],$$

$$\Psi_{\text{ref}}(k,t)=cA(-k)\varphi_{\text{ref}}(-k)\exp[-iE(k)t/\hbar];$$

$$\Psi_{\text{right}}(x,t)=\frac{1}{2\pi}\int_{\infty}^{\infty}\Psi_{\text{tr}}(k,t)\exp(ikx)dk \quad \text{for} \quad x\geqslant b. \tag{2b}$$

Here

$$\Psi_{\text{tr}}(k,t)=cA(k)\varphi_{\text{tr}}(k)\exp[-iE(k)t/\hbar].$$

Here it should be kept in mind that $T(k)$ is an even function of k, $J(k)$ and $F(k)$ are odd functions of k, $A(k)$ is a complex-valued square-integrable function, and c is a normalization constant.

The solutions (2a) and (2b) evidently describe the motion of wave packets in the extrabarrier regions: (2a) is a superposition of the incident wave packet (index ‘‘inc’’), whose center (‘‘of mass’’) at time $t=0$ is located at the point $x=0$, and a reflected packet (index ‘‘ref’’) and (2b) describes the transmitted wave packet (index ‘‘tr’’).

Since the tunneling parameters of the individual waves depend on k, it is more convenient to analyze the dynamics of the wave packets in the k representation. In this case, it is impossible to obtain analytically the wave function in the k representation. However, it is easy to show that the asymptotic behavior of the packets (i.e., at sufficiently early and sufficiently late times, when the packets have not yet interacted or have already interacted with the barrier) is described in the k representation by the functions $\Psi_{\text{inc}}(k,t),\Psi_{\text{ref}}(k,t)$, and $\Psi_{\text{tr}}(k,\ t)$ [see Eqs. (2a) and (2b)]. Of course, we must assume in this case that initially the distance a from the wave packet to the left boundary of the barrier is quite large. We write the conditions that each of the three packets lies in its ‘‘own’’ x region in the form

$$\int_{-\infty}^{a}|\Psi_{\text{left}}(x,t)|^2dx=\int_{\infty}^{\infty}|\Psi_{\text{inc}}(k,t)|^2dk=1$$

for sufficiently early times and

$$\int_{-\infty}^{a}|\Psi_{\text{left}}(x,t)|^2dx=\int_{\infty}^{\infty}|\Psi_{\text{ref}}(k,t)|^2dk=\overline{R},$$

$$\int_{b}^{\infty}|\Psi_{\text{right}}(x,t)|^2dx=\int_{\infty}^{\infty}|\Psi_{\text{tr}}(k,t)|^2dk=\overline{T}$$

for sufficiently long times.

It is obvious that $\overline{T}$ and $\overline{R}$ are simply the transmission and reflection coefficients of the wave packet, respectively. Using the expressions (2a) and (2b), it is easy to show that

$$\overline{T}=\langle T(k)\rangle_{\text{inc}};\quad \overline{R}=\langle R(k)\rangle_{\text{inc}}.$$

Here and below, the average value $\langle\hat{Q}(k)\rangle_{\text{inc}}$ for an operator $\hat{Q}$ referring to the incident wave packet is determined by the expression

$$\langle\hat{Q}\rangle_{\text{inc}}=\frac{\langle\Psi_{\text{inc}}|\hat{Q}|\Psi_{\text{inc}}\rangle}{\langle\Psi_{\text{inc}}|\Psi_{\text{inc}}\rangle}$$

(the average values of an operator $\hat{Q}$ for the transmitted and reflected wave packets are determined similarly).

3. TRANSMISSION AND REFLECTION TIMES OF WAVE PACKETS

To determine the temporal parameters of the tunneling of wave packets, we shall determine the positions of the centers of the wave packets for all scattering channels. Since the coordinate operator $x=i\cdot d/dk$, we have

$$\langle\hat{x}\rangle_{\text{inc}}(t)=m^{-1}\hbar k_0t, \tag{3}$$

$$\langle\hat{x}\rangle_{\text{tr}}(t)=m^{-1}\hbar\langle k\rangle_{\text{tr}}t+d-\langle J'\rangle_{\text{tr}}, \tag{4}$$

$$\langle\hat{x}\rangle_{\text{ref}}(t)=-m^{-1}\hbar\langle -k\rangle_{\text{ref}}t+2a+\langle J'-F'\rangle_{\text{ref}}. \tag{5}$$

Here and below, the prime means a derivative with respect to k and $k_0=\langle k\rangle_{\text{inc}}$.

Let the point Z_i be located at a distance L_1 ($L_1<a$) from the left boundary of the barrier and let the point Z_2 be located at a distance L_2 from the right boundary. Since our approach is of an asymptotic character, we must assume that the distances L_1 and L_2 are large. We shall find the passage time for the segment Z_1Z_2 as well as the time required for the packet center to pass from the point Z_1 to the barrier and return (reflection time).

Let t_1 and t_2 be times such that

$$\langle\hat{x}\rangle_{\text{inc}}(t_1)=a-L_1\quad;\langle\hat{x}\rangle_{\text{tr}}(t_2)=b+L_2. \tag{6}$$

Using Eqs. (3), (4), and (6), we can then write the passage time Δt_{tun} ($\Delta t_{\text{tun}}=t_2-t_1$) for the section of interest in the form

$$\Delta t_{\text{tun}}=\frac{m}{\hbar}\left[\frac{\langle J'\rangle_{\text{tr}}+L_2}{\langle k\rangle_{\text{tr}}}+\frac{L_1}{k_0}+a\left(\frac{1}{\langle k\rangle_{\text{tr}}}-\frac{1}{k_0}\right)\right]. \tag{7}$$

For the reflected packet, let t_1 and t_3 be times such that

$$\langle\hat{x}\rangle_{\text{inc}}(t_1)=\langle\hat{x}\rangle_{\text{ref}}(t_3)=a-L_1. \tag{8}$$

It follows from Eqs. (3), (5), and (8) that the reflection time Δt_{ref} ($\Delta t_{\text{ref}}=t_3-t_1$) is given by

$$\Delta t_{\text{ref}}=\frac{m}{\hbar}\left[\frac{\langle J'-F'\rangle_{\text{ref}}+L_1}{\langle -k\rangle_{\text{ref}}}+\frac{L_1}{k_0}+a\left(\frac{1}{\langle -k\rangle_{\text{ref}}}-\frac{1}{k_0}\right)\right]. \tag{9}$$

We note that the relations (7) and (9) contain terms which are proportional to a, on account of which the temporal characteristics of the wave packet have unusual properties. For example, if $\langle k\rangle_{\text{tr}}>k_0$, then by increasing a the time Δt_{tun} can be made negative and arbitrarily large in modulus. For the reflected packet this happens when $\langle -k\rangle_{\text{tr}}>k_0$ (as will be clear from what follows [see Eq. (12)], for each particle only one of these inequalities is satisfied). In both cases the centers of the transmitted and reflected packets must move away from the barrier before the maximum of the incident packet reaches it. In addition, this lead time can be arbitrarily large, depending on the value of a.

Here it is important to underscore that this effect does not vanish for wave packets with arbitrarily small but finite width in k space (such packets were studied in Ref. 1). By increasing the parameter a, the expression $a(\langle k\rangle_{\text{tr}}^{-1}-k_0^{-1})$ in this case can be made arbitrarily large in modulus.

In this connection, one could cast doubt on the validity of the expressions (7) and (9). However, as has already been noted, this unusual behavior of the wave packets is indeed observed in the numerical modeling.[1–7] The questions arise not with the behavior itself of the wave packets. (With time a wave packet spreads, and whatever the value of a is, the center of the transmitted packet will appear on the other side of the barrier no earlier than the moment at which the leading edge of the incident packet reaches the left boundary of the barrier. ‘‘Extra particles’’ cannot appear beyond the barrier, since the normalization of the wave function holds at each moment in time.) The problem lies in how to relate a given property of the wave packets with the motion of a tunneling particle. Since the question is of fundamental importance, we shall discuss it in greater detail.

4. CHANGE IN THE AVERAGE MOMENTUM OF A WAVE PACKET AND IN THE MOMENTUM OF A PARTICLE DURING TUNNELING

We note that the unusual behavior of wave packets stems from the fact that the average values of k in the two scattering channels are individually different from k_0. For example, for the GWP $A(k)=\exp[-l_0^2(k-k_0)^2]$, where l_0 is the width of the GWP in x space at zero time, and k_0 is the average value of the wave number, it is easy to show that

$$\langle k\rangle_{\mathrm{tr}}=k_0+\frac{\langle T'\rangle_{\mathrm{inc}}}{4l_0^2\langle T\rangle_{\mathrm{inc}}}, \tag{10}$$

$$\langle -k\rangle_{\mathrm{ref}}=k_0+\frac{\langle R'\rangle_{\mathrm{inc}}}{4l_0^2\langle R\rangle_{\mathrm{inc}}}. \tag{11}$$

In the case of wide GWPs the expressions (10) and (11) are identical (asymptotically) to the expressions (4.13) and (4.14) of Ref. 6, respectively.

Setting

$$\langle k\rangle_{\mathrm{tr}}=k_0+\delta k_{\mathrm{tr}},\quad \langle -k\rangle_{\mathrm{ref}}=k_0+\delta k_{\mathrm{ref}},$$

we write the relations (10) and (11) in the form

$$\overline{T}\delta k_{\mathrm{tr}}=-\overline{R}\delta k_{\mathrm{ref}}=\frac{\langle T'\rangle_{\mathrm{inc}}}{4l_0^2}. \tag{12}$$

Here we took into account that $R'=-T'$.

It is helpful to write the first equation in (12) in the form of a ''conservation law'' for the average wave number

$$\overline{T}\langle k\rangle_{\mathrm{tr}}+\overline{R}\langle -k\rangle_{\mathrm{ref}}=k_0. \tag{13}$$

The relation (13) reflects the fact that the average values of the particle momentum are identical in the initial and final states (the final state is a superposition of the transmitted and reflected packets). However, we can see that the average momenta of the transmitted and reflected packets are individually different from $\hbar k_0$. For example, for a GWP this difference vanishes only in the limit $l_0\to\infty$. This property of the wave packets also appears, at first glance, to be at variance with common sense: The momentum of a particle should not change as the particle passes through a stationary potential barrier. It can be shown, however, that a more accurate interpretation of the packet formalism can remove all contradictions which arise in it.

5. THE NEED TO SEPARATE THE TUNNELING AND REFLECTING COMPONENTS FROM THE INCIDENT WAVE PACKET

Indeed, according to the standard interpretation, in the problem at hand quantum mechanics describes statistically a series (strictly speaking, infinite) of experiments (measurements) in which the particle is incident on a potential barrier from the left side and then either passes through the barrier or is reflected from it. The incident wave packet describes the entire series of measurements, and the transmitted and reflected packets separately describe only part of them. This is why in determining the tunneling times of a particle it is pointless to compare separately the positions of the centers of the transmitted and reflected wave packets with the position of the center of the incident packet. The latter can be compared only with a superposition of the reflected and transmitted packets. In this case, the initial and final states describe the same collection of measurements. If, however, we wish to study the two scattering channels separately, then in determining the transmission (reflection) time of a particle only the collection of measurements in which the particle has successively tunneled (reflected) need be taken into account. In other words, we must separate (if it is possible to do so) from the incident wave packet two components that describe the incident particle in both series of measurements separately. We can then compare with these packets the positions of the centers of the transmitted and reflected packets.

One of the main conditions which the desired packets must satisfy is that the transition from the initial into a final state must occur for each channel separately (i.e. in each series of measurements) with the ''number of particles'' being conserved.

We write the incident wave packet in the form

$$\Psi_{\mathrm{inc}}(k,t)=\Psi_{\mathrm{inc}}^{\mathrm{tr}}+\Psi_{\mathrm{inc}}^{\mathrm{ref}}+\Psi_{\mathrm{inc}}^{\mathrm{int}},$$

where

$$\Psi_{\mathrm{inc}}^{\mathrm{tr}}=\sqrt{T(k)}\Psi_{\mathrm{inc}}(k,t), \tag{14}$$

$$\Psi_{\mathrm{inc}}^{\mathrm{ref}}=\sqrt{R(k)}\Psi_{\mathrm{inc}}(k,t), \tag{15}$$

$$\Psi_{\mathrm{inc}}^{\mathrm{int}}=[i-\sqrt{T(k)}-\sqrt{R(k)}]\Psi_{\mathrm{inc}}(k,t). \tag{16}$$

It is easy to verify that the functions $\Psi_{\mathrm{inc}}^{\mathrm{tr}}$ and $\Psi_{\mathrm{inc}}^{\mathrm{ref}}$ satisfy the relations

$$\langle\Psi_{\mathrm{inc}}^{\mathrm{tr}}|\Psi_{\mathrm{inc}}^{\mathrm{tr}}\rangle+\langle\Psi_{\mathrm{inc}}^{\mathrm{ref}}|\Psi_{\mathrm{inc}}^{\mathrm{ref}}\rangle=\langle\Psi_{\mathrm{inc}}|\Psi_{\mathrm{inc}}\rangle. \tag{17}$$

$$\langle\Psi_{\mathrm{inc}}^{\mathrm{tr}}|\Psi_{\mathrm{inc}}^{\mathrm{tr}}\rangle=\langle\Psi_{\mathrm{tr}}|\Psi_{\mathrm{tr}}\rangle,\langle\Psi_{\mathrm{inc}}^{\mathrm{ref}}|\psi_{\mathrm{inc}}^{\mathrm{ref}}\rangle=\langle\Psi_{\mathrm{ref}}|\Psi_{\mathrm{ref}}\rangle; \tag{18}$$

$$\langle k\rangle_{\mathrm{inc}}^{\mathrm{tr}}=\langle k\rangle_{\mathrm{tr}};\langle k\rangle_{\mathrm{inc}}^{\mathrm{ref}}=\langle -k\rangle_{\mathrm{ref}}. \tag{19}$$

$$\langle\hat{x}\rangle_{\mathrm{inc}}^{\mathrm{tr}}(t)=m^{-1}\hbar\langle k\rangle_{\mathrm{inc}}^{\mathrm{tr}}\cdot t; \tag{20}$$

$$\langle\hat{x}\rangle_{\mathrm{inc}}^{\mathrm{ref}}(t)=m^{-1}\hbar\langle -k\rangle_{\mathrm{inc}}^{\mathrm{ref}}\cdot t. \tag{21}$$

We see that the contributions containing $\Psi_{\mathrm{inc}}^{\mathrm{int}}$ and the interference terms do not enter into the relation (17). Therefore the entire set of measurements indeed separates into two sets. The relations (18) give a basis for the assumption that in the series of measurements in which the particle passes through the barrier the state of the particle before scattering is described by the function $\Psi_{\mathrm{inc}}^{\mathrm{tr}}(k,\ t)$. In the other series, in which a particle is reflected, the state of the particle is described by the function $\Psi_{\mathrm{inc}}^{\mathrm{ref}}(k,\ t)$. In both cases, the scattering of the particle occurs with conservation of the number of particles [see Eq. (18)] and momentum [see Eq. (19)]. According to the initial formulation of the problem, the two particles are initially located at the point $x=0$ [see Eqs. (20) and (21)].

Taking the relation (13) into account, it is easy to show that

$$\overline{T}\langle k\rangle_{\mathrm{inc}}^{\mathrm{tr}}+\overline{R}\langle k\rangle_{\mathrm{inc}}^{\mathrm{ref}}=k_0. \tag{22}$$

In other words, the average values of k in both scattering channels equal k_0.

We note that the above-described procedure for dividing the incident wave packet into two components, which describe an incident particle in both series of measurements separately, is not unique. It is easily verified that the relations (17)–(20) remain unchanged if $\Psi_{\mathrm{inc}}^{\mathrm{tr}}$ is multiplied by $\exp[iw_1(k-k_0)]$ and $\Psi_{\mathrm{inc}}^{\mathrm{ref}}$ by $\exp[iw_2(k-k_0)]$, where $w_1(k)$ and $w_2(k)$ are arbitrary odd real functions. In the case at hand, however, this nonuniqueness is not important, since in what follows we shall employ only the relations (17)–(22) to determine the particle-tunneling times.

6. PARTICLE TRANSMISSION AND REFLECTION TIMES

On this basis, we determine the transmission time τ_{tun} for the section Z_1Z_2 for the particle as the difference of the times t_2 and t_1 which satisfy the equations

$$\langle \hat{x}\rangle_{\mathrm{inc}}^{\mathrm{tr}}(t_1)=a-L_1;\ \ \langle \hat{x}\rangle_{\mathrm{tr}}(t_2)=b+L_2. \tag{23}$$

It thus follows that

$$\tau_{\mathrm{tun}}(L_1,L_2)=\frac{m}{\hbar\langle k\rangle_{\mathrm{tr}}}[\langle J'\rangle_{\mathrm{tr}}+L_1+L_2]. \tag{24}$$

We shall define the time required by the particle to leave the point Z_1, reflect from the barrier, and return to the initial point, i.e., the reflection time, τ_{ref}, as the difference t_3-t_1 in the times for which

$$\langle \hat{x}\rangle_{\mathrm{inc}}^{\mathrm{ref}}(t_1)=\langle \hat{x}\rangle_{\mathrm{ref}}(t_3)=a-L_1. \tag{25}$$

It is easy to show that

$$\tau_{\mathrm{ref}}(L_1)=\frac{m}{\hbar\langle -k\rangle_{\mathrm{ref}}}[\langle J'-F'\rangle_{\mathrm{ref}}+2L_1]. \tag{26}$$

We note that the expressions (24) and (26), which are asymptotically exact for large values of L_1 and L_2 contain contributions from the barrier region $\tau_{\mathrm{tun}}^{(0)}$ $[\tau_{\mathrm{tun}}^{(0)}=\tau_{\mathrm{tun}}(0,\,0)]$ and $\tau_{\mathrm{ref}}^{(-)}$ $[\tau_{\mathrm{ref}}^{(-)}=\tau_{\mathrm{ref}}(0)]$. These are the desired transmission and reflection times for a particle in the barrier region.

Here it is very important to keep in mind that expressions (24) and (26) become meaningless if $L_1=0$ or $L_2=0$. The temporal characteristics defined above therefore by no means describe measurements in which probes are located on the barrier boundaries. Incidentally, this arrangement of the probes in itself is meaningless, since it is impossible to identify, because of interference, the centers (or, in a special case, the maxima) of the packets. The conditions for interference not to influence the measurements and the conditions for our model to be applicable are virtually identical: The distances to the points Z_1 and Z_2, where the probes are located, must be sufficiently large. Therefore, to determine experimentally, for example, the barrier transmission time, it is first necessary to measure the time $\tau_{\mathrm{tun}}(L_1,\,L_2)$ and then subtract from it the transmission times of the sections $(a-L_1,a)$ and $(b,$ $b+L_2)$, where the particle is free. It is assumed that the average velocities of the particles on the free sections have already been determined separately for both scattering channels.

In the transfer matrix method[9,10] it is easy to establish that for waves incident on the barrier from the right, the phase F changes sign and the transmission coefficient and phase J do not depend on the direction of motion of the incident wave. Taking this circumstance into account, it is easy to show that the reflection time $\tau_{\mathrm{ref}}^{(+)}$ for a particle incident on the barrier from the right is determined by the expression

$$\tau_{\mathrm{ref}}^{(+)}=\frac{m\langle J'+F'\rangle_{\mathrm{ref}}}{\hbar\langle -k\rangle_{\mathrm{ref}}}. \tag{27}$$

It follows from the recurrence relations for the tunneling parameters of the wave (see Ref. 10) that for barriers which are symmetric with respect to the center of the barrier section $F'\equiv 0$. In this case the particle reflection time does not depend on the side from which the particle is incident on the barrier, $\tau_{\mathrm{ref}}^{(0)}=\tau_{\mathrm{ref}}^{(-)}=\tau_{\mathrm{ref}}^{(+)}$.

It is easy to verify that for wave packets which are sufficiently narrow in k space the tunneling times determined here are equal to the ''phase'' times (see Ref. 10).

7. OTHER APPROACHES IN PACKET ANALYSIS

Here it is very important to compare our results with those of Ref. 3. Both cases concern packet analysis for particles described by arbitrary wave packets. However, there is a fundamental difference between the two approaches. In our approach, the determination of the tunneling times is based on calculations of the average values of the operator $\hat{x}$. At the same time, in Ref. 3 averaging of the operator $\hat{\theta}$, where $\hat{\theta}=i\nu^{-1}\hat{x}$ and $\nu=\hbar k/m$, was actually used.

This procedure, in our view, has two serious drawbacks. The first one is connected with the fact that this operator is non-Hermitian. [Here it would be better to employ the corresponding Hermitianized operator (tempus), which was studied in Ref. 11.] The other, more serious, drawback is that the average value of the square of the operator $\hat{\theta}$ (and the Hermitianized operator) and therefore the variance of this quantity in a state which is described by a GWP are infinite. But if the variance of a physical quantity in some state is infinite, then its average value (though in itself finite) becomes physically meaningless in this state.

The unfounded use of the operator $\hat{\theta}$ (rather than $\hat{x}$) in determining the temporal parameters of tunneling leads, in our view, to another error as well.

The so-called dwell time τ_d of a particle in the barrier region is quite widely used in the literature (see Refs. 6 and 7) to describe both scattering channels together. Sometimes (see, for example, Ref. 4) this quantity is written as the average value of the tunneling times in both scattering channels

$$\tau_d=\overline{T}\theta_{\mathrm{tun}}+\overline{R}\theta_{\mathrm{ref}}, \tag{28}$$

where θ_{tun} is the tunneling time, and θ_{ref} is the reflection time.

Doubts as to the validity of such averaging with respect to both scattering channels were stated earlier in Refs. 6 and 7. However, a clear answer to the question of the validity of the expression (28) was not given. This equation, in our view, in general, is incorrect, because it implicitly assumes the validity of averaging the operator $\hat{\theta}$, which plays in this case the role of the ‘‘time operator.’’ We have shown that even for a GWP this operation becomes meaningless.

In our approach there arises a different relation for the temporal parameters which relates both scattering channels. Indeed, it is easy to verify that

$$\langle J' \rangle_{\mathrm{inc}} \equiv \overline{T} \langle J' \rangle_{\mathrm{tr}} + \overline{R} \langle J' \rangle_{\mathrm{ref}} . \tag{29}$$

Using expressions (24), (26), (27), and (29) and assuming that

$$\langle J' \rangle_{\mathrm{inc}} = v_0 \tau_d , \tag{30}$$

we write this identity in the form

$$\tau_d = \overline{T} \tau_{\mathrm{tun}}^{(0)} \frac{\langle v \rangle_{\mathrm{tr}}}{v_a} + \overline{R} \, \overline{\tau}_{\mathrm{ref}} \frac{\langle - v \rangle_{\mathrm{ref}}}{v_0} , \tag{31}$$

where $\overline{\tau} = \frac{1}{2} (\tau_{\mathrm{ref}}^{(-)} + \tau_{\mathrm{ref}}^{(+)})$, and $v_0 = m^{-1} \hbar k_0$.

As we can see, the relation (31) is different from Eq. (28). It assumes the same form as the relation (28)

$$\tau_d = T(k_0) \tau_{\mathrm{tun}}^{(0)} + R(k_0) \tau_{\mathrm{ref}}^{(0)} \tag{32}$$

only in a special case, for a GWP which is infinitely narrow in k space and for symmetric barriers. In this case the velocities of the particles are the same in both scattering channels. Therefore, if a tunneling time describing both scattering channels simultaneously is introduced, then this must be done, in our view, according to the expressions (30) and (31).

In conclusion, we note that in a future work we shall investigate numerically with the aid of the expressions obtained above the tunneling parameters of different packets. Specifically, the properties of Gaussian wave packets with different degrees of localization in x space will be studied.

[1] E. H. Hauge, J. P. Falck, and T. A. Fjeldly, Phys. Rev. B **36**, 4203 (1987).
[2] N. A. Teranishi, A. M. Kriman, and D. K. Ferry, Superlatt. Microstr. **3**, 509 (1987).
[3] W. Jaworski and D. M. Wardlew, Phys. Rev. A **37**, 2843 (1988).
[4] C. R. Leavens and G. C. Aers, Phys. Rev. B **39**, 1202 (1989).
[5] R. Landauer and Th. Martin, Solid State Commun. **84**, 115 (1992).
[6] E. H. Hauge and J. A. Stovneng, Rev. Mod. Phys. **61**, 917 (1989).
[7] R. Landauer and Th. Martin, Rev. Mod. Phys. **66**, 217 (1994).
[8] A. M. Steinberg, Phys. Rev. Lett. **74**, 2405 (1995).
[9] N. L. Chuprikov, Fiz. Tekh. Poluprovodn. **26**, 2040 (1992) [Sov. Phys. Semicond. **26**, 1147 (1992)].
[10] N. L. Chuprikov, Fiz. Tekh. Poluprovodn. **27**, 806 (1993) [Semiconductors **27**, 439 (1993)].
[11] D. H. Kobe and V. C. Aguillera-Navarro, Phys. Rev. A **50**, 933 (1994).

Translated by M. E. Alferieff

Transverse stability of an impact-ionization front in a Si $p^{+}-n-n^{+}$ structure

A. M. Minarskiĭ

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

P. B. Rodin

Institute of Theoretical Physics, Berlin Technical University, D-10623 Berlin, Germany

(Submitted May 31, 1996; accepted for publication July 5, 1996)
Fiz. Tekh. Poluprovodn. **31**, 432–436 (April 1997)

The transverse stability of an impact-ionization front in a large-area silicon $p^{+}-n-n^{+}$ structure is studied. An analytical model allowing for simultaneous motion of the ionization front and displacement of the majority carriers from the nondepleted part of the n base is proposed. The stability of a planar front is investigated, the growth increments are calculated, and the physical mechanisms of instability are indicated. A criterion is formulated for quasistable propagation of a wave. © *1997 American Institute of Physics.* [S1063-7826(97)01104-6]

1. INTRODUCTION

The excitation of an impact-ionization wave is a unique, with respect to its rate, and at the same time very universal ''nonoptical'' mechanism for modulating the conductivity of a semiconductor structure — diode, transistor, or thyristor.[1–4] In a recent report[5] we showed that the propagation of a wave in a large-area $p^{+}-n-n^{+}$ junction is accompanied by an increase in the transverse perturbations of the position of the front. Long-wavelength (with respect to the base length W) modes, whose buildup time is close to the time over which the ionization front (IF) traverses its own length, play the dominant role. The development of an instability can result in the formation of local switching channels with characteristic transverse size W. The nonuniform dynamics of the impact-ionization front, including a stage of formation of local switching channels, was described in Ref. 6. The transverse instability of the IF is a fundamental factor limiting the possibility of diode structures which are switched with the aid of an impact-ionization wave.

This paper continues the theoretical investigations of the stability of impact-ionization fronts.[5–8] It considers the problem of the stability of the front in the most important case in which an undepleted quasilinear region is present in the n base. The propagation of the IF in such a structure is accompanied by the displacement of carriers from the quasineutral region. As a result, the dynamics and stability of the IF should be studied together with the motion of the boundary between the space-charge region and the quasineutral region. An analytical model of the propagation of an impact-ionization wave in a Si-$p^{+}-n-n^{+}$ structure is proposed and its transverse stability is analyzed.

2. DESCRIPTION OF THE MODEL

We assume that initially the impact-ionization wave propagates uniformly over the entire area of the device. We studied a structure whose base contains, at least initially, an undepleted region, i.e. the condition $W>[2\varepsilon\varepsilon_0 u/qN_d]^{1/2}$ is satisfied (here u is the voltage applied to the structure, N_d is the density of donors in the n base, q is the electron charge, ε is the permittivity of the semiconductor, and ε is the permittivity of free space). Below we call such a structure a long-base structure, in contrast to the similar TRAPATT diode[4] structures, in which the depleted region occupies the entire n base.[5]

Three regions can be distinguished in the n base (Fig. 1); a neutral plasma region (NR); a space-charge region (SCR), including a completely depleted region, the IF itself, and the space-charge layer behind the front; and, a region of dense plasma behind the front. The characteristic carrier densities for the case of a Si diode are presented in the caption in Fig. 1. We assume below that the ionization coefficients are the same for electrons and holes and we use a step approximation for the impact-ionization coefficient as a function of the electric-field intensity E: $\alpha(E)=\alpha_0 H(E-E_a)$, where H is the unit step (Heaviside) function, E_a is the impact-ionization threshold, and α_0 is a model parameter, which represents the saturated value of the impact-ionization coefficient. We disregard the impact ionization in the NR. The field in the region of the dense plasma behind the front is assumed to be weak and the distribution of the field in the SCR is assumed to be trapezoidal.

The position of the IF $x_f(y,\ z)$, together with the position $d(y,x)$ of the boundary between the SCR and the NR, completely determine the distribution of the field $E(x,\ t)$ in a given section of the structure ($y=$ const, $z=$ const). The equations of motion for the variables $x_f(y,\ z)$ and $d(y,z)$ have the form

$$\frac{\partial x_f}{\partial t}=\frac{1}{\tau_f}l_f(u,x,d), \tag{1}$$

$$\frac{\partial d}{\partial t}=-v_n(E_0(u,x,d)), \tag{2}$$

where

$$\tau_f=\frac{1}{\alpha v_s}\ \ln\frac{N}{n_0},\quad N=\frac{\alpha_0\varepsilon\varepsilon_0 E_a}{q},$$

$$v_n(E)=\frac{\mu_n E}{1+E/E_s}.$$

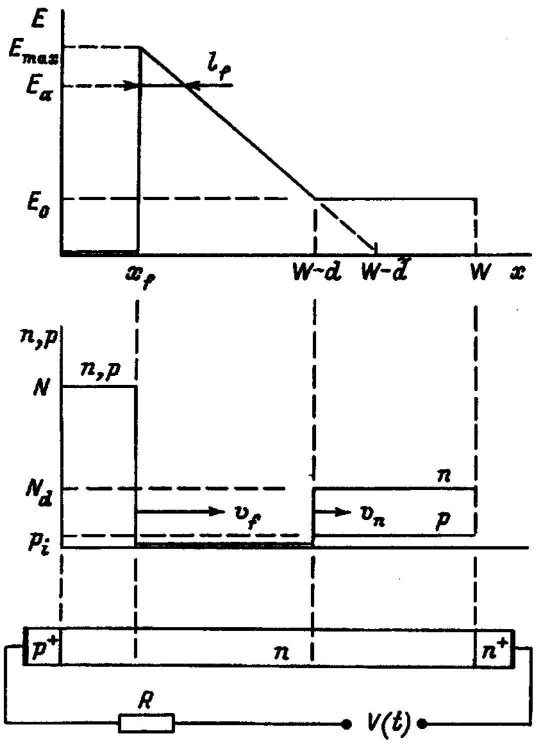

FIG. 1. Diagram of the electric field distribution $E(x,t)$ and carrier densities n and p in the base of a $p^+ - n - n^+$ structure during the propagation of an impact-ionization wave. The typical values of the densities are: $N=10^{17}-10^{18}$ cm^{-3}, $N_d=10^{14}$ cm^{-3}, $p_i=10^6$ cm^{-4}.

Here ν_s is the unsaturated carrier velocity, N is the density of the nonequilibrium electron-hole plasma behind the IF, n_0 is the carrier density in the SCR in front of the IF, $E_s=\nu_s/\mu_n \approx 3\times 10^3$ V/cm is the characteristic field in which the drift velocity saturates, and $\mu_n=1500$ V$\cdot$cm^2/s is the weak-field electron mobility. A direct calculation shows that the size l_f of the ionization region and the field E_0 in the quasineutral region are given by the expressions

$$l_f=\frac{\varepsilon\varepsilon_0}{qN_d}\left(\frac{u}{W-x_f}-E_a\right)+\frac{1}{2}\left(W-x_f-\frac{d^2}{W-x_f}\right). \tag{3}$$

$$E_0=\frac{u}{W-x-f}-\frac{qN_d}{2\varepsilon\varepsilon_0}\frac{(W-x_f-d)^2}{w-x_f}. \tag{4}$$

Equation (1) describes the propagation of the IF in a model proposed in Refs. 6 and 8. The derivation of this equation[6,8] depends on two universal laws of propagation of impact-ionization waves and streamers:[9] The front velocity v_f is proportional to the size l_f of the impact-ionization region; the displacement time τ_f of the electric field out of the ionization region is determined by the transport time of the space charge associated with the front through this region.

The velocity of the boundary of the NR equals the velocity of the electrons displaced to the n^+ emitter. The minority carriers (holes) are then extracted into the SCR. The density p of the minority carriers is low compared with N_d. As a result, their contribution to the space charge in the SCR is very small. In Eq. (2) an approximation of the velocity $v_n(E)$ is used — the carrier velocity as a function of the electric field strength in Si proposed in Ref. 10.

Equations (1)–(4) must be supplemented by the Kirchhoff equation

$$u=V(t)-qN_dRS\left\langle\frac{dx_f}{dt}\right\rangle, \tag{5}$$

which relates the average (over the area) velocity of the front with the voltage on the structure [$V(t)$ is the source voltage, R is the load resistance, and S is the area of the structure]. In Eq. (5) the relation of the current density j in a given section of the structure with the velocity of the front is taken into account:[6] $j=qN_d\nu_f$. This closed system of equations completely describes the dynamics of a planar IF and an IF that is perturbed by long-wavelength ($\lambda > W$) fluctuations. The condition that the total current is continuous along the x direction holds automatically for Eqs. (1)–(4).

3. DYNAMICS OF A PLANAR IONIZATION FRONT

In the present case of a long-base structure, the one-dimensional dynamics of the IF exhibits several characteristic features which have no analogs for structures such as TRAPATT diodes. Two new phenomena can be distinguished: retardation of the IF as a result of a redistribution of voltage between the SCR and the NR, whose most extreme manifestation is stopping of the wave, and avalanche breakdown of the NR. We shall give only a brief description of these effects, which is necessary for further understanding of the results of the stability analysis.

1. *Retardation of the* IF. Differentiating Eqs. (1) and (2) with respect to time shows that the sign of the acceleration of the IF is determined by the sign of the right side of the expression

$$\left(\tau_f+qN_dRS\frac{\partial l_f}{\partial u}\right)\frac{d\nu_f}{dt}=\frac{\partial l_f}{\partial u}\frac{dV(t)}{dt}+\frac{d\nu_n-\tilde{d}\nu_f}{W-x_f}, \tag{6}$$

where

$$\tilde{d}=-(W-x_f)\frac{\partial l_f}{\partial x_f}=(W-x_f)-\frac{\varepsilon\varepsilon_0 E_{\max}}{qN_d}, \quad \frac{\partial l_f}{\partial u}>0.$$

The meaning of the quantities $\tilde{d}$ and $E_{\max}$ is clear from Fig. 1. For sufficiently large W, the quantities d and $\tilde{d}$ are close to one another, while v_f can be several times greater than v_n. As a result, for $V(t)=$ const the velocity of the IF decreases as the front propagates, and it is possible for the wave to stop. The physical mechanism of the stopping of the wave is connected with the fact that the growth of the field occurring in the NR during the motion of the IF results in a voltage drop across the SCR.

2. *Breakdown of the quasineutral region.* Since $\nu_f > \nu_s$ and $\nu_n < \nu_s$, the IF always overtakes the receding boundary of the quasineutral region. The SCR decreases in size and the field in the NR increases. The impact-ionization threshold can be reached before the majority carriers are completely removed from the n base, as a result of which avalanche breakdown of the NR ceases. In the simplest case when the current in the external circuit is fixed, $\nu_f=$ const, and the carrier velocity in the NR equals the saturated value ν_s, the condition that the IF traverses the entire n base before E_0 reaches E_a can be represented in the form

$$W < \frac{v_f}{v_f - v_s}\,\frac{\varepsilon\varepsilon_0 E_a}{qN_d}. \tag{7}$$

On this basis it follows that wave breakdown in the initial form, similar to that occurring in a TRAPATT diode, does not occur in a diode structure with a sufficiently long base. For low velocities v_f the IF can stop and for high velocities breakdown of NR can occur.

4. TRANSVERSE STABILITY OF A PLANAR IONIZATION FRONT

1. *Basic equation.* Let us consider the stability of a planar IF which propagates with a prescribed law of variation of $V(t)$, with respect to transverse perturbations $\delta x_f(r,t)$ and $\delta d(r,\,t)$, whose wavelength is greater than the size of the ionization region and the thickness of the structure: $\lambda > W$, l_f. The fact that long-wavelength disturbances play a dominant role was substantiated in Refs. 5 and 8. We assume that the average values (over the area) $\langle \delta x_f(r,t)\rangle$ and $\langle \delta d(r,t)\rangle$ equal zero. As a result, the perturbations do not interact with the external circuit.

Linearizing the equations of motion of the perturbed system gives a system of equations which can be represented, using specific relations (3) and (4) for l_f and E_0, in compact form

$$\begin{aligned}
\tau_f \frac{\partial \delta x_f}{\partial t} &= \frac{\partial l_f}{\partial x_f}\delta x_f + \frac{\partial l_f}{\partial d}\delta d,\\
\tau_d \frac{\partial \delta d}{\partial t} &= -\left(\frac{\partial l_f}{\partial x_f}+1\right)\delta x_f - \left(\frac{\partial l_f}{\partial d}+1\right)\delta d,
\end{aligned} \tag{8}$$

where

$$\tau_d = \frac{\varepsilon\varepsilon_0}{qN_d}\left(\frac{\partial v_n}{\partial E_0}\right)^{-1},$$

$$\frac{\partial l_f}{\partial x_f} = -\frac{\tilde{d}}{W - x_f},\quad \frac{\partial l_f}{\partial d} = -\frac{d}{W - x_f}.$$

Equations (8) include the characteristic times τ_f and τ_d. The first time is the time over which the front traverses its own length $\tau_f = \tau_M(\mu E_a \ln(N/n_0)/v_s)$ and depends on the Maxwellian relaxation time τ_M in the dense plasma behind the ionization front $(\mu E_a \ln(N/n_0)/v_s) \approx 10^2$. The quantity $\tau_d = \tau_m(1 + E_0/E_s)^2$ is related to the Maxwellian relaxation time τ_m in the NR by a factor that determines how close E_0 is to E_s. In all cases $\tau_d > \tau_f$. In the limit $v_n \to v_s$ the stronger condition $\tau_d \gg \tau_f$ holds.

The characteristic equation has the form

$$\begin{aligned}
&\lambda^2 + b\lambda + c = 0,\quad b = \frac{1}{\tau_d}\left(\frac{\partial l_f}{\partial d}+1\right) - \frac{1}{\tau_f}\frac{\partial l_f}{\partial x_f},\\
&c = \frac{1}{\tau_f \tau_d}\left(\frac{\partial l_f}{\partial d} - \frac{\partial l_f}{\partial x_f}\right).
\end{aligned} \tag{9}$$

In the case where the carrier velocity is a monotonic function of the field, the free term is always negative:

$$c = -\frac{1}{\tau_f}\frac{\partial v_n}{\partial E_0}\frac{E_0}{W - x_f} < 0.$$

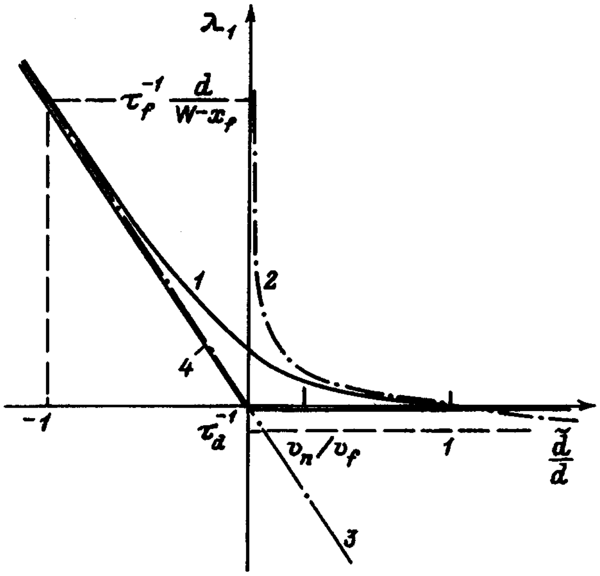

FIG. 2. Growth increment λ_1 of an inhomogeneous perturbation versus the dimensionless parameter $\tilde{d}/d$ (*1*). The dot-dashed lines 2 and 3 are determined by the relations (10) and (11), respectively. Curve *4* corresponds to the case of saturated carrier velocity in the neutral region $v_n = v_s$. In the parameter $v_n/v_s < \tilde{d}/d < 1$ the IF moves with retardation.

As a result, the roots of Eq. (9) are real and have different signs $\lambda_1 > 0$ and $\lambda_2 < 0$. Therefore, the instability occurs for all propagation regimes of the front. It follows from Eqs. (5) that the growing mode $(\lambda_1 > 0)$ corresponds to a displacement of the front as a whole: $\delta x_f > 0$, $\delta d < 0$. The decaying mode $(\lambda_2 < 0)$ corresponds to a change in the size of SCR: $(\delta x_f > 0,\ \delta d > 0)$.

Since the interval W traversed by the front is bounded, a conclusion about the instability of the front can be drawn in practice only if $\lambda_1 W/v_f$. The quantity λ_1 depends strongly on the profile of the field in the wave. Two cases corresponding to different instability mechanisms can be distinguished.

2. *Weak and strong instability of the front.* Assuming that the ratio τ_f/τ_d is small, we represent the positive root of the characteristic equation in the form

$$\lambda_1 = \frac{1}{\tau_d}\frac{d - \tilde{d}}{\tilde{d}},\quad \tilde{d} > 0, \tag{10}$$

$$\lambda_1 = -\frac{1}{\tau_f}\frac{\tilde{d}}{W - x_f},\quad \tilde{d} < 0. \tag{11}$$

Here the growth increments are expressed in terms of the characteristic times τ_f and τ_d and the geometric parameters of the profile of the field in the wave. The range of applicability of these expressions does not include the neighborhood of the value $\tilde{d} = 0$. The dependence of λ_1 on the parameter $\tilde{d}/d$ is shown in Fig. 2. It is clear from expressions (10) and (11) and Fig. 1 that regions of weak $(\lambda \sim \tau_d^{-1})$ and strong $(\lambda \sim \tau_f^{-1})$ instability can be distinguished in this problem.

Weak instability occurs for profiles possessing an extended neutral region $(\tilde{d} > 0)$. Its physical mechanism is associated with the positive relation between the electron departure velocity rate and the field E_0. An increase in the field

E_0 at a fixed voltage results in a more rapid motion of the NR boundary. As a result of the decrease in the length of the boundary, it also results in a redistribution of the voltage between the NR and the SCR. The relative increase in the voltage across the SCR also increases the velocity of the IF. Therefore, the development of instability in the present case is determined mainly by carrier drift in the NR. This determines the large (corresponding to τ_d) characteristic development times of the perturbation. In semiconductors with a nonmonotonic field dependence of the drift velocity this effect is absent in the range of fields where the differential mobility is negative.

To estimate correctly the role of the above-described mechanism, it should be noted that the weak instability is realized, as a rule, for fronts which with a constant source voltage $V(t)$ move with retardation. Indeed, according to Eq. (6), the front slows down if $\tilde{d}/d>\nu_n/\nu_f$ (see Fig. 2). Instability development in this case means not acceleration of the leading sections but rather only a decrease in the their retardation rate.

The condition of strong instability is $\tilde{d}<0$. Its mechanism is associated with acceleration of the IF itself and is completely analogous to the mechanism described for similar TRAPATT diode structures in Ref. 5. Indeed, for $\partial l_f/\partial x>0$ a local forward displacement of a section of the wave with a prescribed voltage on the device results in an increase in the size l_f of the ionization region. As a result, the velocity of the leading section increases and the nonuniformity intensifies. The characteristic times in this case are equal to the time over which the front traverses its own length.

As the flat front propagates, the parameter $\tilde{d}/d$ decreases, varying in the interval $(1,-\infty)$. The system then passes through a phase of weak instability and enters a region of strong instability. At the final stage of switching, when the majority carriers are completely displaced out of the base ($\tilde{d}<0$, $d=0$), the growth increment is described by the expression

$$\lambda_1=\frac{1}{\tau_f}\,\frac{E(xf)-E(W)}{2E(W)}, \tag{12}$$

which wass obtained previously in Ref. 5.

In the case where the carrier velocity in the NR saturates $\nu_n=\nu_s$, the characteristic time $\tau_d\to\infty$. Equation (11) then becomes exact, $\lambda_2=0$ for $\tilde{d}>0$. The condition $\tilde{d}>0$, or

$$\frac{\varepsilon\varepsilon_0 E_{\max}}{qN_d}>W-x_f, \tag{13}$$

then becomes the criterion for stability of a planar front. We note the clear geometric meaning of this criterion: The slope of the line connecting the points ($E_{\max}$, x_f) and (0,W) must be less than $qN_d/\varepsilon\varepsilon_0$.

5. DISCUSSION OF THE RESULTS AND COMPARISON WITH EXPERIMENT

The foregoing analysis shows that transverse instability is an internal property of the dynamics of an IF in a Si p^+-n-n^+ structure and occurs for all regimes and in all stages of front propagation. However, in practice, it is necessary to distinguish weak instability of the wave, which occurs in the presence of an extended neutral region in front of the SCR, and strong instability, which corresponds to the case of a short neutral region and the case of a completely depleted n base. As a result of the long characteristic growth times of the perturbation that correspond to a weak instability of the front, this mechanism ca be regarded as secondary. This circumstance makes the results of the present work applicable to GaAs diodes as well. The condition (13) is a criterion for quasistable propagation of the front. It is significant that this criterion necessarily breaks down as the IF approaches the n^+ emitter. As a result, the final stage of propagation of the IF always contains a transversely unstable phase.

The criterion (13) is closely related to the empirical criterion of malfunction-free operation of silicon diode peakers,[2,11] which require the presence of a neutral region in the n-base at the initial moment of the wave propagation:

$$W>\left(\frac{2\varepsilon\varepsilon_0 u}{qN_d}\right)^{1/2}. \tag{14}$$

Breakdown of the criterion (14) in the case of Si diodes results in local thermal damage.[2,11] Local switching of GaAs diodes, observed in Ref. 12, apparently is attributable to the same cause.

The condition (13) gives a theoretical justification for the criterion (14). This condition ensures quasistability of the initial phase of wave propagation. At the same time, because of the presence of a transversely unstable final phase, this empirical reliability criterion cannot be regarded as a sufficient condition for uniformity of the modulation of the volume of the structure after switching.

We now turn to the results of the one-dimensional numerical modeling of IF propagation, performed in Ref. 13 for a Si structure with the parameters $W=200$ μm and $N_d=10^{14}$ cm^{-3}. According to the computational results, the switching time is 600 ps.[13] Estimates of the growth increments of the transverse perturbation, made according to Eqs. (10) and (11) for the electric field profiles $E(x,\,t)$ obtained in Ref. 13 , show that the motion of a planar IF becomes unstable even at $t=200$ ps. A clear transition also occurs between the stable and unstable phases. It can be concluded from this analysis that one-dimensional numerical modeling is unsuitable for quantitative description of the switching process.

We thank A. V. Gorbatyuk, I. V. Grekhov, A. F. Kardo-Sysoev, and M. E. Levinshteĭn for stimulating discussions. This work was supported by the Institute of Theoretical Physics at Uppsala University.

[1] V. I. Grekhov and A. F. Kardo-Sysoev, Pis'ma Zh. Tekh. Fiz. **5**, 950 (1979) [Sov. Tech. Phys. Lett. **5**, 395 (1979)].

[2] V. M. Turkevich and I. V. Grekhov, *New Methods of Transmitting High Power with Semiconductor Devices* [in Russian], Nauka, Leningrad (1988).

[3] I. V. Grekhov, Solid–State Electron. **32**, 923 (1989).

[4] B. C. Deloach and D. L. Scharfetter, IEEE Trans. Electron. Dev. **20**, 9 (1970).

[5] A. M. Minarskiĭ and P. B. Rodin, Pis'ma Zh. Tekh. Fiz. **20**(12), 38 (1994) [Tech. Phys. Lett. **20**, 753 (1994)].

[6] A. M. Minarskiĭ and P. B. Rodin, Zh. Tekh. Fiz. **29**, 1506 (1995) [*sic*].
[7] A. M. Minarsky and P. B. Rodin, in *Proc. of 1995 International Semiconductor Device Research Symposium* (Charlottesville, USA. Dec., 1995).
[8] A. M. Minarskiĭ and P. B. Rodin, Preprint No. 1639, A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg (1995).
[9] M. I. D'yakonov and V. Yu. Kacharovskiĭ, Zh. Éksp. Teor. Fiz. **95**, 1850 (1989) [Sov. Phys. JETP **68**, 1070 (1989)].
[10] C. Jacobini, C. Canali, G. Ottaviani, and A. Alberigi, Solid–State Electron. **20**, 77 (1977).
[11] V. I. Grekhov, A. F. Kardo-Sysoev, L. S. Kostipa, and S. V. Shendereĭ, Zh. Tekh. Fiz. **51**, 1709 (1981) [Sov. Phys. Tech. Phys. **26**, 984 (1981)].
[12] S. N. Vaĭnshteĭn, Yu. V. Zhilyaev, and M. E. Levinshteĭn, Pis'ma Zh. Tekh. Fiz. **14**, 1526 (1989) [*sic*].
[13] Yu. D. Bilenko, M. E. Levinshteĭn, M. V. Popova, and V. S. Yuferev, Fiz. Tekh. Poluprovodn. **17**, 1812 (1983) [Sov. Phys. Semicond. **17**, 1156 (1983)].

Translated by M. E. Alferieff

Deep level transient spectroscopy under conditions of current-carrier exchange between two allowed bands

A. A. Lebedev

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

(Submitted June 28, 1996; accepted for publication July 8, 1996)
Fiz. Tekh. Poluprovodn. **31**, 437–440 (April 1997)

Some characteristic features of deep level transient spectroscopy (DLTS) spectra of deep levels which effectively exchange current carriers with both allowed bands are studied. It is shown that the maxima of the peaks shift very little on the temperature scale, but analysis of the spectra by the conventional methods leads to errors in the determination of the parameters of the deep levels (ionization energies, capture cross sections). Several methods for determining these parameters more accurately are proposed and a numerical example of such an analysis of DLTS spectra is presented.
[S1063-7826(97)01204-0]

The methods of capacitance spectroscopy — isothermal capacitance relaxation (ICR), thermally stimulated capacitance (TSC), nonstationary deep-level transient spectroscopy (NDLCS or DLTS in the foreign literature) — and other methods have been widely used in recent years to determine the density and parameters of deep levels (DLs) in semiconductors.[1–6] These methods are based on measurements of the charge-exchange kinetics of DLs in the space-charge layer (SCR) with a reverse bias on the diode. In studying charge relaxation on DLs, exchange of current carriers (CCs) between the DLs and only one allowed band is ordinarily taken into account, since it is assumed that the interaction with the other band is weak because of the large difference in the ionization energies E_n and E_p of the DLs or the electron and hole capture cross sections σ_n and σ_p, respectively.[3] This approach is justified in most cases, but for some DLs located near the center of the band gap it is not always correct. Specifically, in Si many DLs have been discovered near the center of the band gap.[7] In this paper we examine the characteristic features of the thermal relaxation of the capacitance in the case where the DLs interact with the two allowed bands and we present methods for determining the parameters of the DLs in this case.

1. KINETICS OF CHARGE EXCHANGE BETWEEN DEEP LEVELS

For definiteness, we shall study a diode with a sharp asymmetric p^+-n junction or Schottky barrier (SB) in n-type material. In the calculations below we assume that the DL possesses a fixed ionization energy and generation of CTs through excited states does not play a large role.[3,8] The kinetics of charge exchange on DLs in the dark in the part of the SCR, where under reverse bias the density of free CCs can be neglected $(n\approx p\approx 0)$, is described by the equation

$$\frac{dm}{dt}=e_p(M-m)-e_n m, \tag{1}$$

where

$$e_p=c_p N_v \exp\left(-\frac{E_p}{kT}\right)\equiv c_p p_1, \tag{2a}$$

$$e_n=c_n N_c \exp\left(-\frac{E_n}{kT}\right)\equiv c_n n_1 \tag{2b}$$

are the rates of thermal generation of electrons from the valence band to DLs and from the DLs into the conduction band, respectively; N_ν and N_c are the effective densities of states in the valence and conduction bands, respectively; $c_n=\nu_e\sigma_n$ and $c_p=\nu_n\sigma_p$ are the electron and hole trapping rates in DLs, respectively; $\nu_e$and ν_n are the thermal velocities of the electrons and holes, respectively; and, M and m are the total density of DLs and the electron density in the DLs, respectively.

The solution of Eq. (1) with $m_0=\delta M$ at $t=0$ has the form

$$\frac{m(t)}{M}\ \frac{e_p}{e_n+e_p}+\left(\delta-\frac{e_p}{e_n+e_p}\right)\exp\left(-\frac{t}{\theta_c}\right), \tag{3}$$

where $\theta_c^{-1}=e_c=e_n+e_p$. It follows from Eq. (3) that under conditions when CCs are exchanged with the two allowed bands, charge exchange occurs according to an exponential law with a time constant θ_c, but e_c is the sum of exponentials of the reciprocal of the temperature with different exponents and pre-exponential factors, and the Arrhenius law $\ln\theta=f(1/T)$ is linear only in the case where the DL lies exactly at the center of the band gap $(E_n=E_p)$. As calculations show, however, when θ_c is measured in a limited range of temperatures T, this dependence can be easily approximated by a straight line with activation energy E_a that falls between E_n and E_p. Figure 1 (curve *1*) is a computed example of the function $\log\theta_c=f(1/T)$ for a level with the parameters $E_n=0.5$ eV, $E_p=0.6$ eV, $c_nN_c=10^{12}\ \mathrm{s}^{-1}$, and $c_pN_\nu=10^{14}\mathrm{s}^{-1}$. We see that the points fall well on a straight line with activation energy $E_a=0.545$ eV. Methods for determining E_n and E_p more accurately will be examined below.

The limiting charge exchange on DLs $(t\gg\theta_c)$, as follows from Eqs. (2) and (3), is

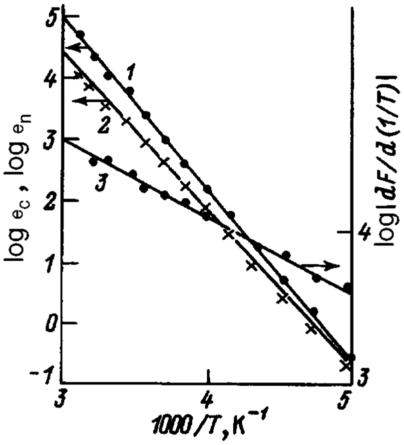

FIG. 1. Computed Arrhenius curves for e_c (*1*), e_n (*2*), and the function $dF/d(1/T)$ (*3*) for the example presented in the text.

$$\frac{m_{st}}{M}=\frac{e_p}{e_n+e_p}=\frac{1}{1+\dfrac{c_nN_c}{c_pN_\nu}\exp\left(\dfrac{E_g-2E_n}{kT}\right)}. \tag{4}$$

Here use was made of the fact that the band gap is $E_g=E_n+E_p$. It follows from Eq. (4) that under standard conditions the filling of the DLs in the SCR depends on the ratio c_n/c_p, E_n, and E_p. If σ_n and σ_p are of the same order of magnitude, then the DLs lying above the center of the band gap ($E_n<E_g/2$) will be empty and the levels located below the center of the band gap ($E_n>E_g/2$) will be completely filled with electrons, irrespective of the initial filling of the DLs, the type of base conductivity, and the type of level — donor or acceptor.[3–9]

2. INITIAL CHARGE EXCHANGE BETWEEN DEEP LEVELS

Charge exchange between DLs can be initiated by several methods: 1 — decreasing the reverse bias V_1 to 0 (dropping the voltage); 2 — partial dropping of the voltage from V_1 to $V_2<V_1$; 3 — injection from the same $p-n$ junction; 4 — base-side illumination of the diode with light that is absorbed ($h\nu>E_g$); 5 — injection from the emitter in a triode structure; and other methods. In the case of charge exchange by methods 4 and 5, complete removal of all electrons from the DLs is possible with a reverse bias. Without a reverse bias the two methods are similar to injection from a $p-n$ junction and are therefore not considered here.

When the voltage is dropped, an equilibrium filling of the DLs $m_0/M=n_0/(n_0+n_1)$, where n_0 is the equilibrium electron density in the conduction band, is established in the quasineutral region of the diode. If the DL lies near the Fermi level and in some temperature range $e_p>e_n$, then it is possible that in the SCR the DL is not emptied of electrons but rather is filled from the valence band, i.e., $m_0<m_{st}$. This will result in a change in sign of the capacitance relaxation. This situation is not studied in the present paper, and it is assumed below that $n_0\gg n_1$. If $n_0\gg M$, then the DL is filled with electrons according to an exponential law with time constant $\theta_p=(c_nn_0)^{-1}$everywhere except in the transitional layer near the boundary of the $p-n$ junction with the quasineutral region.[1–4,10] In most cases this layer is small, and we shall assume in the calculations below that the region of incomplete filling of the DL is taken into account and the duration of the filling pulse is $t_f\gg\theta_p$. Then $\delta=1$. The kinetics of filling of the DL with a drop of the voltage is an additional method for measuring σ in the absence of an electric field.

In the case of injection from a $p-n$ junction, the distribution of the injected CCs is $\Delta p=\Delta n\sim\exp(-X/L_D)$, where L_D is the diffusion length. In the calculations below we shall assume that the injection level is high ($\Delta n\gg n_0$) and the width of the SCR $W\ll L_D$ for the reverse bias employed. Then the initial filling of the DL will be

$$\delta=\frac{c_n}{c_n+c_p}\equiv\frac{1}{1+\lambda},$$

where $\lambda=c_p/c_n$. This filling can change when the direct bias is changed to a reverse bias. First of all, when the voltage is switched, at the SCR boundary, which moves into the volume of the base, $\Delta p\approx0$ and electrons are trapped. If the boundary of the SCR moves relatively slowly (compared with θ_p), then an appreciable change in the degree of filling of DL is possible. Second, holes which accumulated in the base under direct bias pass through the SCR and can be trapped in the DL.[11] Such charge transfer at the DL is substantial at high direct-current densities and long lifetime of nonequilibrium CCs in the diode base. These factors all produce some uncertainty in the initial filling of the DL in the case of injection from a $p-n$ junction.

3. DETERMINATION OF THE PARAMETERS OF DEEP LEVELS

Charge exchange on DLs results in relaxation of the capacitance of the structure. In the calculations below we shall assume that the shallow and deep impurities are distributed uniformly in the crystal and that the base resistance R is small ($\omega^2R^2C^2\ll1$, where ω is the angular frequency of the measuring signal) and the structure can be assumed to be one-dimensional.

Let us examine the DLTS spectra in the case of a low density of DLs ($M\ll N$, where N is the density of shallow donors) or when measurements are performed in the constant-capacitance regime. Then $\Delta C(t)$ or $\Delta V(t)\sim m(t)$. When CCs are exchanged with the two allowed bands, an amplitude factor A_1 which depends on T will appear in the DLTS spectra. If the initial charge transfer is produced by dropping the voltage, then $\delta=1$ and

$$A_1=\frac{e_n}{e_n+e_p}=\frac{1}{1+\dfrac{c_pN_\nu}{c_nN_c}\exp\left(-\dfrac{E_p-E_n}{kT}\right)}$$

$$\equiv\frac{1}{1+\lambda b\,\exp\left(-\dfrac{\Delta E}{kT}\right)}, \tag{5}$$

where $\Delta E=E_p-E_n$, and $b=N_\nu/N_c$.

The temperature dependence of A_1 is shown in Fig. 2 (curves *1*–*7*) in dimensionless coordinates for several values

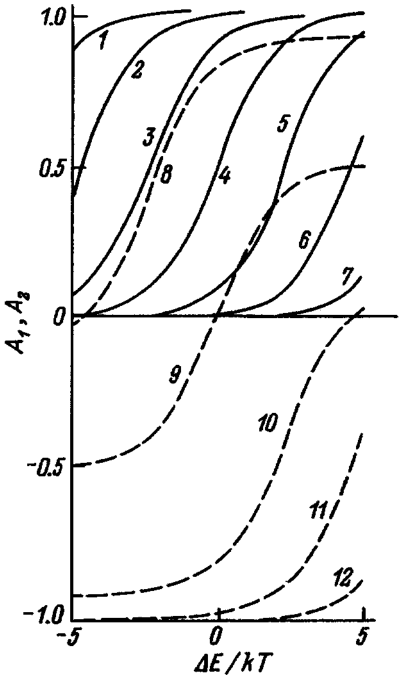

FIG. 2. Amplitude factors A_1 (*1–7*) and A_2 (*8–12*) versus $\Delta E/kT$.

of λb. From the character of the dependence $A_1(T)$ a conclusion can be drawn about the energy position of the DL: A_1 increases with T, if the DL lies below the center of the band gap ($E_n > E_p$), and it drops for the opposite relation $E_n < E_p$, since $|\Delta E/kT| \to 0$ as T increases. The values of A_1 for the example considered above fall on the curve *6* in Fig. 2. It is also obvious from the figure that appreciable capacitance relaxation occurs only for a certain ratio of ΔE and λ. Estimates show that even for a small displacement of the DL from the center of the band gap, the trapping cross sections for CCs on the DL must differ by several orders of magnitude in order for an appreciable exchange of CCs with both allowed bands to occur.

As numerical calculations showed, the amplitude factor has virtually no effect on the form of the DLTS peak and the position of its maximum. Typical displacements of the maximum are in the range 1–5 K. This makes it possible to determine θ_c by the standard method.[3–6] It can be determined more accurately from the intersection of the DLTS curves.[12]

To determine E_n and E_p separately, we shall make use of the circumstance that $A_1\theta_c = e_n$. We can then calculate E_n from the Arrhenius relation for e_n, and $E_p = E_g - E_n$. The Arrhenius curve for the values of $e_n(T)$ obtained in this manner is shown in Fig. 1 (curve *2*). It follows that the activation energy in this case is 0.505 eV and corresponds, to a high degree of accuracy, to the value of E_n which is incorporated in the calculation. The cross sections for trapping of CCs can be calculated if N and M are known. As shown above, M can be estimated for several methods of initial filling of the DLs.

Now we know $\theta_c(T)$, E_n, and E_p, but we do not know the pre-exponential factors c_nN_c and c_pN_ν. They can be calculated if the known functions $\exp(E_n/kT)$ and $\exp(E_p/kT)$ at some temperature are substituted into $e_c(T)$. If the operation is repeated at some other temperature, then we obtain a system of equations where c_nN_c and c_pN_ν are the unknowns and can be easily calculated.

It is also possible to use the known quantity $\exp(E_n/kT)$. Then

$$F=\frac{\exp\left(\frac{E_n}{kT}\right)}{\theta_c}=c_nN_c+c_pN_\nu \exp\left(-\frac{\Delta E}{kT}\right).$$

The derivative of this function is

$$\frac{dF}{d(1/T)}=-\frac{c_pN_\nu\Delta E}{k}\exp\left(-\frac{\Delta E}{kT}\right),$$

and ΔE and c_pN_ν and then c_nN_c can be calculated from it. The Arrhenius curve for the values of $-dF/d(1/T)$ obtained in analyzing the numerical example presented above is shown in Fig. 1 (curve *3*). The activation energy in this case is 0.105 eV instead of 0.1 eV, the value incorporated in the calculation. Calculations in other numerical models showed that the indicated quantities are determined with an accuracy of no worse than 10%.

Some estimates can also be made in the case if the amplitude of the DLTS peaks increases or decreases rapidly as T varies. Such a sharp temperature dependence means that in this case $\lambda\ \exp|\Delta E/kT| \gg 1$. Then ΔE and λ can be estimated from the Arrhenius relation.

Accordingly, all basic parameters of a DL can be determined, to a high degree of accuracy, from the DLTS spectra by several methods: the energy position of the DL in the band gap and the cross section for trapping of electrons and holes.

Initial charge exchange on a DL is also possible by injection from the same p^+-n junction. In this case with a high injection level $\delta = c_n/(c_n+c_p)$ and the amplitude factor is

$$A_2=\frac{c_n}{c_n+c_p}-\frac{e_p}{e_n+e_p}=\frac{1}{1+\lambda}-\frac{\lambda b\exp\left(-\frac{\Delta E}{kT}\right)}{1+\lambda b\exp\left(-\frac{\Delta E}{kT}\right)}$$

$$=\frac{1-\lambda^2 b\exp\left(-\frac{\Delta E}{kT}\right)}{(1+\lambda)\left[1+\lambda b\exp\left(-\frac{\Delta E}{kT}\right)\right]}. \tag{6}$$

It should be noted that in the case of injection-initiated charge exchange on a DL the position of the corresponding DLTS peak on the temperature scale does not change, since θ_c does not depend on the initial filling of the DL.

The temperature dependence $A_2 = f(T)$ in dimensionless coordinates is shown in Fig. 2 (curves *8–12*) for several values of λ and $b=1$. It follows from Eq. (5) that for $\lambda^2 b = \exp(\Delta E/kT)$ the amplitude factor $A_2 = 0$ and the capacitance relaxation changes sign (Fig. 2, curves *8–10*). The equality $A_2 = 0$ makes it possible to check whether or not the parameters of the DL were determined correctly.

The difference of the amplitude factors is

$$\Delta A = A_1 - A_2 - \frac{\lambda}{1+\lambda}$$

and for $\Delta A \approx 1$ for large values of λ, i.e., the DL is virtually completely freed of electrons and hence the total density M of DLs can be estimated. For $\lambda \lesssim 0.1$ the curves $A_1(T)$ and $A_2(T)$ virtually merge.

For a high injection level electrons can also fill the DLs which in the neutral region of the diode lie above the Fermi level. This leads to the appearance of additional DLTS peaks after injection. The amplitude of such peaks increases with the injection level and decreases rapidly with increasing temperature. This case will be examined in detail in a separate paper.

After some obvious changes, the formulas presented above also hold for sharp $n-p$ junctions or Schottky barriers on p-type semiconductors.

4. CONCLUSIONS

The temperature dependence of the height of the DLTS peaks is a characteristic indication that the DL interacts efficiently with the two allowed bands. Analysis of such spectra by the standard method can give incorrect results. More reliable methods for determining the parameters of DLs (E_n, E_p, σ_n, σ_p) for this case were proposed.

The above-examined methods for analyzing DLTS spectra are valid only for ''simple'' DLs. If a center with a DL has a complicated structure and the center undergoes restructuring in the process of charge exchange on the DL, then the amplitude of the DLTS signal can depend on entirely different factors. Then different methods based on other physical models must be developed to analyze the results of the measurements.

The exponential capacitance (or voltage) relaxation due to the dependence of the parameters of DLs on the electric field in the SCR or charge exchange on several DLs with different θ complicates the analysis of the spectra. Complications can also arise in the case where the cross sections for trapping of CCs and generation of CCs via excited states of a center with a DL are strongly temperature dependent.

[1] C. T. Sah and V. G. K. Reddi, IEEE Trans, Electron. Dev. **ED-11**, 345 (1964).

[2] C. T. Sah, L. Forbes, L. L. Rosier, and A. F. Tasch, Solid−State Electron. **13**, 759 (1970).

[3] L. S. Berman and A. A. Lebedev, *Capacitance Spectroscopy of Deep Levels in Semiconductors* [in Russian], Nauka, Leningrad (1981).

[4] A. A. Lebedev in *Methods of Diagnostics of Point Defects in Semiconductors and Device Structures Based on Them* [in Russian], Namangan (1970), p. 17.

[5] D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).

[6] S. M. Sze, *Physics of Semiconductor Devices*, Wiley, N. Y., 1969 [Russian trans., Énergiya, Moscow, 1973].

[7] E. V. Astrova and A. A. Lebedev, Fiz. Tekh. Poluprovodn. **20**, 683 (1986) [Sov. Phys. Semicond. **20**, 432 (1986)].

[8] L. M. Kapitanova, L. S. Kostina, A. A. Lebedev, A. T. Mamdalimov, and Sh. Makhkamov, Fiz. Tekh. Poluprovodn. **8**, 694 (1974) [Sov. Phys. Semicond. **8**, 444 (1974)].

[9] A. A. Lebedev and V. Ékke, Fiz. Tekh. Poluprovodn. **20**, 1806 (1986) [Sov. Phys. Semiconductors **20**, 1134 (1986)].

[10] A. A. Lebedev, Fiz. Tekh. Poluprovodn. **28**, 1980 (1994) [Semiconductors **28**, 1092 (1994)].

[11] E. V. Astrova and A. A. Lebedev, Preprint No. 1333, A. F. Ioffe Physicotechnical Institute, Leningrad (1989).

[12] E. V. Astrova and A. A. Lebedev, Fiz. Tekh. Poluprovodn. **24**, 549 (1990) [Sov. Phys. Semicond. **24**, 344 (1990)].

Translated by M. E. Alferieff

Effect of inhomogeneities of Bi_2T_3 crystals on the transverse Nernst–Ettingshausen effect

M. K. Zhitinskaya and S. A. Nemov

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

T. E. Svechnikova

A. A. Boĭkov Institute of Metallurgy, Russian Academy of Sciences, 119877 Moscow, Russia

(Submitted July 4, 1996; accepted for publication July 8, 1996)
Fiz. Tekh. Poluprovodn. **31**, 441–443 (April 1997)

The effect of inhomogeneities in p-Bi_2Te_3 crystals, grown by different methods, on the coefficients of the transverse Nernst–Ettingshausen and Hall effects has been studied. The degree of homogeneity of the crystals was estimated on the basis of thermo-emf (Seebeck coefficient) data. It was established that the concentration dependences of the Nernst–Ettingshausen coefficient at $T=300$ K in homogeneous and inhomogeneous samples are different. This makes it possible to use the Nernst–Ettingshausen effect data for quality assessment of Bi_2Te_3 crystals. It is shown that the anomalous drop of the Hall coefficient with increasing temperature in a number of samples is due to inhomogeneities, and in Sn-doped Bi_2Te_3 crystals it is due to the existence of an impurity level against the background of the valence-band states. © *1997 American Institute of Physics.* [S1063-7826(97)01304-5]

Bismuth telluride Bi_2Te_3 is a layered crystal with $R3mm$ structure. The anisotropy of the physical properties and the characteristic features of its phase diagram are the reasons why there are at least two types of inhomogeneities in these crystals. One type is due to the nonuniform distribution of atoms in layers, which results in the appearance of diverse point defects. The other type of possible inhomogeneities are layered inhomogeneities. It is obvious that defects, inhomogeneities, and imperfections can influence the electrical properties of crystals. The parameters of the energy spectrum which are determined from measurements of these properties can, naturally, differ for samples prepared by different methods, since the number of defects and the relation between them changes. In this connection, the effect of the degree of perfection of crystals on their electrical properties and the parameters of the energy spectrum which are obtained from them have been discussed repeatedly in the literature.[1–4] It is obvious that the most reliable information about the energy spectrum of current carriers can be obtained by investigating perfect single crystals.

Additional interest in this problem has now appeared in connection with active investigations of the effect of different impurities on the electrical and thermoelectric properties of Bi_2Te_3-based materials.[5–9] When this compound is doped, especially with metals (for example, Pb, Sn, In, and Tl), the number of defects and the degree of their nonuniformity increase. In inhomogeneous semiconductors, a situation in which the temperature dependences of the transport coefficients (specifically, the Hall coefficient) can simulate the behavior of a homogeneous semiconductor with impurity levels is possible. This danger must be kept in mind when analyzing and interpreting experimental results. For this reason, it is important to estimate the effect of inhomogeneities on the electrical properties.

In the present work we investigated the effect of different degrees of inhomogeneity (crystal quality) on the transverse Nernst–Ettingshausen effect (TNNE) and the Hall effect.

We have accordingly prepared samples by methods which are ordinarily employed for preparing Bi_2Te_3 crystals: Perfect single crystals were grown by Czochralski's method.[4] Single and block crystals were obtained by the method of directed crystallization. Polycrystals with grain size $d=100-200\ \mu$m were obtained by the metalloceramic method. The ceramic samples were annealed at $T=350$ K for 120 h.

To obtain a range of hole densities, the samples were additionally doped with different impurities. The composition of the samples corresponded to the formula $(Bi_{1-x}Me_x)_2$, where (Me = Sn, In, Tl, and Pb) and $x=0.01-6\times10^{19}\ \mathrm{cm}^{-3}$. Tin-doped single crystals were specially grown for this investigation, since the possible existence of the Sn impurity states in p-Bi_2Te_3 against the background of the valence band has been reported in the literature. The homogeneity of the samples was estimated by measuring the Seebeck coefficient with the aid of a thermal probe at room temperature.

The following independent tensor components were measured on single-crystalline samples: the Hall coefficient R_{321}, the thermoelectric power S_{11} and S_{33}, and the Nernst–Ettingshausen coefficient Q_{123}. (The axis 3 corresponds to the direction of the trigonal axis C_3 of the crystal.) The indices in the coefficients correspond to the order of their arrangement: The first index corresponds to the direction of the measured electric field, the second correspond to the direction of the electric current or temperature gradient, and the third corresponds to the direction of the magnetic field. For block crystals, which consist of several single crystals slightly disoriented with respect to one another, and also ceramic samples, which are polycrystals, the results obtained

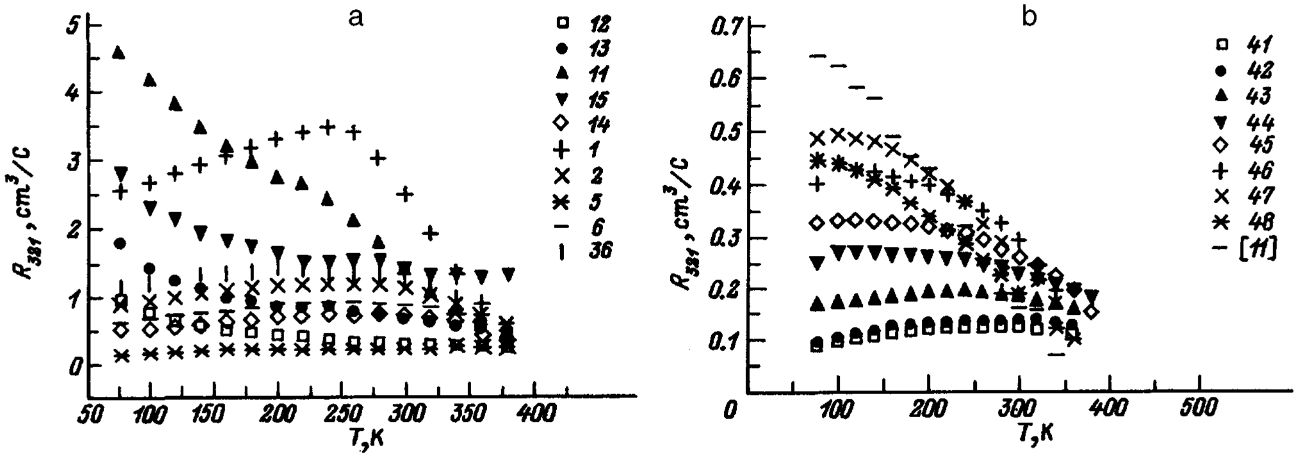

FIG. 1. a–b — Temperature dependences of the Hall coefficient (R_{321}) in p-Bi_2Te_3 samples with different degrees of inhomogeneity. Sample preparation methods: Nos. 1–6, 21–26, 31–36 — Czochralski's method, single crystals. Doping: Nos. 1–46 — deviation from the stoichiometric composition: 11–15 — Sn, 7 — Tl, 21–26 — Pb, 31–36 — In, 41–48 — In, Pb and deviation from quasistoichiometry.

were corrected by the method described in Ref. 10 . Measurements were performed mainly in the temperature range 77–420 K. The coefficient Q was measured under isothermal conditions in a classically weak magnetic field. The hole density p was determined from Hall effect data on the large component R_{321} at liquid-nitrogen temperature. According to published and our data,[5,6,8] an impurity resonance level against the background of the valence band can exist in tin-doped samples. In this case the variation of the Hall coefficient with temperature reflects the change in the current-carrier density. For this reason, the hole density in Sn-doped samples at $T = 300$ K was determined from the corresponding Hall coefficient $R(T)$.

The main experimental data are shown in Figs. 1 and 2. Let us examine the results obtained. As one can see from Fig. 1, the samples can be divided into two groups according to the character of the temperature dependences of the Hall coefficients: 1) samples with normal (ordinary) temperature dependence $R(T)$ for p-Bi_2Te_3, characterized by the presence of a maximum of R in the temperature range 250–400 K and a monotonically decreasing value of R with further increase of T (samples 1, 2, 5, 6, 14, 36, 41–43), and 2) samples in which R decreases starting at 77 K (samples 11–13, 15, 44–48). We note that this unusual, for Bi_2Te_3, character of $R(T)$ was also observed in Refs. 6 , 8 , and 11 . The decrease in the Hall coefficient with increasing temperature could be due to the onset of the intrinsic conduction, the presence of impurity levels, as well as the presence of inhomogeneities.[12]

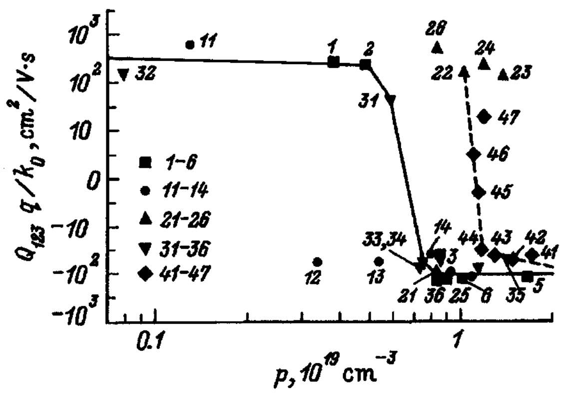

FIG. 2. Density dependence of the transverse Nernst–Ettingshausen coefficient — Q_{123q}/k_0 in p-Bi_2Te_3 samples at $T = 300$ K (k_0 — Boltzmann's constant, q — electron charge). Dots — experiment; the sample numbers are the same as in Fig. 1; solid line — homogeneous samples; dashed line — inhomogeneous samples.

The onset of intrinsic conduction can be judged according to the TNEE data. The point is that the TNEE coefficient is extremely sensitive to the appearance of minority carriers. The onset of mixed electron-hole conductivity in Bi_2Te_3 results in a change in sign of Q (from negative to positive)[13] and is characterized by fast exponential growth of Q. It follows from the TNEE data that the observed decrease of the Hall coefficient at temperatures $T > 300$ K is due to the onset of intrinsic conduction, and in the most inhomogeneous samples the intrinsic conduction starts at lower temperatures.

As one can see from Fig. 1, for a number of samples with high hole densities ($R < 0.5$) a decrease of the Hall coefficient with increasing temperature is observed near 77 K, i.e. at temperatures substantially lower than the onset temperature of intrinsic conduction. This behavior of the Hall coefficient as a function of temperature could be due to the existence of impurity levels[6] or the manifestation of inhomogeneities.[12] In our case, the decreasing Hall coefficients as a function of temperature are observed in inhomogeneous samples (samples 44–48, Fig. 1b). High-quality Sn-doped Bi_2Te_3 single crystals grown by Czochralski's method (samples 11–15, Fig. 1a) are exceptions. The strong T dependence of R is characteristic of narrow-gap IV–VI semiconductors with resonance impurity states.[14] For this reason, it is natural to attribute the observed temperature dependence $R(T)$ in Bi_2Te_3Sn to the existence of a resonance level of tin. Estimates made by analogy with Ref. 14 show that the Sn level lies between the subbands of the valence band, and extrapolation of the temperature dependence of the chemical potential to zero temperature gives an energy of the impurity states $\varepsilon = 0.02 - 0.03$ eV.

Let us now discuss the TNEE data. It is well known that the TNEE coefficient in p-Bi_2Te_3 at $77 < T < 300$ K is characterized by negative values, which decrease in absolute value according to the law of variation of the Hall mobility as a function of temperature. Near room temperature, in samples with typical current carrier densities (p, $n \sim 1 \times 10^{19}$ cm^{-3}), the coefficient changes sign and the character of the temperature dependence changes from power-law to exponential because of the onset of intrinsic conduction. The smallness of the coefficient Q at $T \sim 300$ K makes the Nernst–Ettingshausen effect very sensitive to different types of inhomogeneities. This makes the isotherm of the density dependence of the TNEE coefficient at $T = 300$ K worth examining. We see in Fig. 2 where it is shown that a qualitative difference is observed in the properties of homogeneous and inhomogeneous samples. They possess different density dependences $Q(p)$, and for inhomogeneous samples, characteristically, the TNEE changes sign sooner (i.e., at $T < 300$ K) and Q takes on large positive values. It should be noted that the TNNE data, as one can see from Fig. 2, confirm the high quality of the Sn-doped Bi_2Te_3 crystals. Therefore, the experimentally observed characteristic features in the transport coefficients of these crystals attest to differences in the energy spectrum of samples with and without tin.

In summary, the data on the transverse Nernst–Ettingshausen effect near room temperature can be used together with Hall and Seebeck effect data to estimate the degree of perfection of Bi_2Te_3 crystals.

[1] L. P. Caywood and G. R. Miller, Phys. Rev. B **2**, 3209 (1970).

[2] V. A. Kudinov, Fiz. Tekh. Poluprovodn. **8**, 2057 (1974) [Sov. Phys. Semicond. **8**, 1340 (1974)].

[3] J.-P. Fleurial, L. Gailliard, R. Triboulet, H. Scherrer, and S. Schererer, J. Phys. Chem. Sol. **49**, 1237 (1988).

[4] T. E. Svechnikova, M. A. Korzhuev, N. L. Korenovskiĭ, N. M. Maksimova, N. V. Polikarpova, S. N. Chizhevksaya, P. P. Konstantinov, and G. T. Alekseeva, Neorg. Mater. **30**, 1123 (1994).

[5] M. K. Zhitinskaya, S. A. Nemov, T. G. Abaidulina, and T. E. Svechnikova in *Proc. XIV Int. Conf. on Thermoelectrics* (St. Peterburg, Russia, 1995), p. 56.

[6] V. A. Kulbachinskii, M. Inoue, M. Sasaki, H. Negishi, W. X. Gao, K. Takase, Y. Giman, J. Horak, and P. Lostak, Phys. Rev. B **50**, 16 921 (1994).

[7] J. Pancir, J. Horak, and Z. Stary, Phys. St. Sol. A **103**, 517 (1987).

[8] G. T. Alekseeva, P. P. Konstantinov, V. A. Kugasov, L. N. Luk'yanova, T. E. Svechnikov, and S. N. Chizhevskaya, Fiz. Tverd. Tela (Leningrad) **33**, 3539 (1991) [Sov. Phys. Solid State **33**, 1988 (1991)].

[9] M. K. Zhitinskaya, S. A. Nemov, Yu. I. Ravich, T. G. Abaĭdulina, V. V. Kompaneets, G. S. Bushmarina, and I. A. Drabkin, Fiz. Tekh. Poluprovodn. **27**, 1724 (1993) [Semicond. **27**, 952 (1993)].

[10] D. J. Ryden, J. Phys. C **4**, 1193 (1971).

[11] G. Bergmann, Z. Naturforsch. A **19**, 800 (1964).

[12] A. V. Karpov, A. Ya. Shik, and B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. **16**, 1406 (1982) [Sov. Phys. Semicond. **16**, 901 (1982)].

[13] M. K. Zhitinskaya, V. I. Kaĭdanov, and S. A. Nemov, Dep. No. 3628-76.

[14] V. I. Kaĭdanov and Yu. I. Ravich, Usp. Fiz. Nauk **145**, 51 (1985) [Sov. Phys. Usp. **28**, 31 (1985)].

Translated by M. E. Alferieff

Investigation of carrier transport in a system of undoped quantum wells under pulsed excitation

A. M. Georgievskiĭ, V. A. Solov'ev, B. S. Ryvkin, N. A. Strugov, E. Yu. Kotel'nikov, and V. E. Tokranov

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

A. Ya. Shik

Center for Chemical Physics, University of Western Ontario, London, Canada N6A 3K7

(Submitted July 18, 1996; accepted for publication September 10, 1996)
Fiz. Tekh. Poluprovodn. **31**, 444–450 (April 1997)

A phenomenological theory describing nonstationary carrier transport processes in a system of undoped quantum wells with short carrier generation pulses is developed. An experimental method for finding the characteristic carrier trapping and carrier emission times from the wells is proposed on the basis of the theoretical expressions. The first experimental results making it possible in principle to determine these parameters are reported. © *1997 American Institute of Physics.* [S1063-7826(97)01404-X]

1. INTRODUCTION

Semiconductor $p-i-n$ heterostructures with an i layer containing a system of noninteracting identical quantum wells (KWs) are now the subject of active investigations. Carrier transport through a QW layer in a reverse-biased $p-i-n$ heterostructure have a determining effect on the characteristics of optoelectronic devices, such as optical modulators based on electroabsorption of light[1] and devices with a self-electrooptic effect (SEED),[2] which are used in optical communications and information processing systems.

In a phenomenological theory carrier transport through a system of undoped QWs is described by the characteristic carrier trapping and emission times of a well and the characteristic carrier drift time between neighboring wells.[3,4] A knowledge of these times as a function of the well depth and width, temperature, and electric field[5–7] is necessary in order to produce the above-mentioned optoelectronic devices with prescribed characteristics, and it is also of physical interest in itself.

As a rule, in experiments investigating carrier transport through a layer of QWs electron-hole pairs are excited in the layer optically. As a result, all characteristics are determined by simultaneous transport of electrons and holes. This circumstance makes it much more difficult to interpret the results obtained.[8]

In Ref. 4 theoretical expressions describing the separate transport of electrons and holes in a system of undoped QWs were obtained for the case of stationary excitation outside the system of QWs and monopolar injection within the system. Furthermore, an experimental method, making it possible to estimate the characteristic carrier trapping, emission, and drift times from the form of the response of the current through the structure to the switching on and off of generation was proposed.

In this study, which is a continuation of an earlier study,[4] we develop a phenomenological theory describing nonstationary carrier transport processes in the presence of short generation pulses. On the basis of this regime, an alternative method is proposed for determining the characteristic times from the parameters of the response of the current through the layer of QWs. This method includes a procedure for checking of a specific experimental system of QWs the applicability of the theory developed for describing carrier transport. The first experimental results showing that the characteristic carrier transport times can in principle be determined by comparing the experimental curves with the theoretical calculations are also presented.

2. BASIC EQUATIONS

To describe carrier transport in the structure of interest with an applied electric field, we shall characterize the structure, just as in Ref. 4 , by the number n_k^b of carriers localized in the kth well and the number of carriers n_k in delocalized states above the well. A number of simplifications are made in so doing. It is assumed that the currents are generated outside the layer of QWs and carriers of only one type participate in the transport processes. We neglected the possibility of carrier tunneling between wells (this is an acceptable simplification for barrier thickness greater than 100 Å). The carrier injection level is assumed to be low enough so that the redistribution of the electric field between QWs can be neglected. The field within the QW layer is assumed to be constant and the characteristic carrier drift time τ_d and carrier capture and emission times τ_c and τ_e, respectively, of a well do not depend on the number of wells.

In this case the behavior of a system of M QWs is described by the system of equations

$$dn_k/dt=n_k^b/\tau_e-n_k/\tau_c+(n_{k-1}-n_k)/\tau_d, \tag{1}$$

$$dn_k^b/dt=n_k/\tau_c-n_k^b/\tau_e, \quad k=1,2,\ldots,M. \tag{2}$$

The quantity $n_0(t)$ in Eq. (1) with $k=1$ is viewed as a prescribed density in the injecting electrode and is determined by the form of the carrier generation pulse outside the QW layer. The initial conditions for Eqs. (1) and (2) are

$$n_k^b(t=0)=n_k(t=0)=0,$$

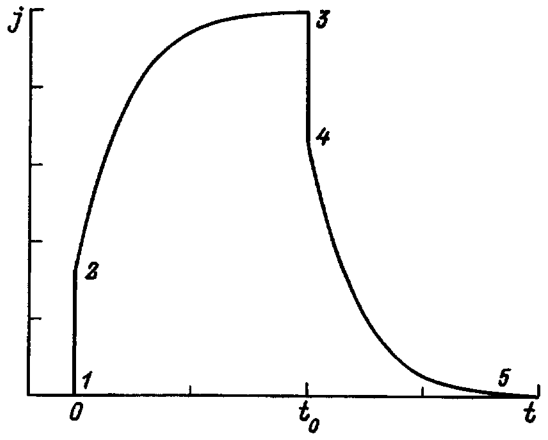

FIG. 1. Schematic representation of the behavior of the current through a structure with multiple QWs versus time with pulsed excitation: *1–2* — Current jump occurring when generation is switched on; *2–3* — slow relaxation with constant excitation; *3–4* — downward current jump occurring when carrier generation is switched off; *4–5* — slow relaxation associated with carrier ejection from the QW layer.

i.e., initially there are no carriers in the region of the QWs. Next, when an excitation pulse is switched on, carriers of one type are injected through the boundary of the QW layer.

In all calculations we assumed that the carrier emission rate from a well is much smaller than the carrier trapping and drift rates ($\tau_e \gg \tau_c,\ \tau_d$). This simplification is acceptable, since in the absence of tunneling between wells the characteristic emission time τ_e from a well contains an exponentially large factor $\exp(\Delta E/kT)$, where ΔE is the ionization energy of the ground state in a QW. Moreover, the carrier generation pulse duration t_0 is required to satisfy the condition $t_0 \gg \tau_1$, where $\tau_1 = \tau_c\tau_d/(\tau_c+\tau_d)$ is the characteristic scale of the fast relaxation time[4] associated with establishment of the carrier distribution above the wells.

3. COMPUTATIONAL RESULTS

Let us consider the case of a square generation pulse with duration t_0

$$n_0(t)=\begin{cases} n_0=\mathrm{const}(t), & 0<t<t_0 \\ 0, & t>t_0. \end{cases}$$

Figure 1 shows schematically the behavior of the current through a structure with many QWs with such an excitation pulse. For $t<t_0$ the densities in the QWs vary in time just as in the stationary excitation case studied in Ref. 4. At times of the order of τ_1 a jump *1–2* occurs in the current through the structure. This jump is associated with establishment of a carrier distribution above the wells from $n_k=0$ to

$$n_k=n_0\left(\frac{1}{1+\tau_d/\tau_c}\right)^k. \tag{3}$$

Next, carrier redistribution occurs as a result of emission from the wells on a time scale $\tau_2=\tau_e(1+\tau_d/\tau_c)$ (section *2–3*). For $\tau_2 \gg \tau_1$ this process is described by the expression

$$n_k(t)=n_0\left(\frac{1}{1+\tau_d/\tau_c}\right)^k\left\{1+\sum_{j=0}^{k-1}\frac{k!}{j!(k-j)!}\left(\frac{\tau_d}{\tau_c}\right)^{k-j}\times\left[1-\exp\left(-\frac{t}{\tau_2}\right)\sum_{p=0}^{k-j-1}\frac{1}{p!}\left(\frac{t}{\tau_2}\right)^p\right]\right\}, \tag{4}$$

a derivation of which is presented in Ref. 4.

The processes associated with the switching off of the excitation are described in a manner that is different from Ref. 4 since the system has not reached saturation. This is expressed in the nonequilibrium carrier distribution over the wells at the moment excitation is switched off. At the time $t=t_0$, when the carrier generation pulse ceases, $n_k^b(t_0)$ is determined from the relation

$$n_k^b(t_0)=n_k(t_0)\frac{\tau_e}{\tau_c}-[n_{k-1}(t_0)-n_k(t_0)]\frac{\tau_e}{\tau_d}, \tag{5}$$

where $n_k(t_0)$ is given by Eq. (4) and $n_0(t_0)=n_0$. As a result of the rapid relaxation, over times of the order of τ_1 there is not enough time for the quantity n_k^b to change (the jump *3–4* in Fig. 1) and n_k drops to values determined from the recurrence relation

$$n_k=n_{k-1}\frac{\tau_1}{\tau_d}+n_k^b\frac{\tau_1}{\tau_e}, \tag{6}$$

where $n_0=0$ and $n_k^b=n_k^b(t_0)$ is given by Eq. (5). Here we have assumed that the number of carriers above the wells is negligible compared with the number of carriers in the wells ($n_k \ll n_k^b$). The relations (5) and (6) are obtained from Eq. (1), where we set $dn_k/dt=0$, since outside the fast-relaxation region this quantity is small, $\sim n_k/\tau_2$.

We shall now describe the slow relaxation after excitation is switched off, when $n_0=0$ (the section *4–5*). We can then write the recurrence relation (6) for $n_k(t)$ in the form of a sum

$$n_k(t)=\frac{\tau_1}{\tau_e}\sum_{p=1}^{k}n_p^b(t)\left(\frac{1}{1+\tau_d/\tau_c}\right)^{k-p}. \tag{7}$$

Substituting expression (7) into Eq. (2), we obtain a system of equations describing carrier redistribution processes over wells after the passage of a carrier generation pulse ($t>t_0$),

$$\frac{dn_k^b}{dt}=-\frac{n_k^b}{\tau_e}+\frac{\tau_d}{\tau_c\tau_2}\sum_{p=1}^{k}n_p^b\left(\frac{1}{1+\tau_d/\tau_c}\right)^{k-p}. \tag{8}$$

The initial condition for the system of equations (8) $n_k^b(t_0)$ is obtained from relation (5).

4. NUMERICAL MODELING

The solution of Eq. (8) can be written in the general form as

$$n_k^b(t)=\exp\left(-\frac{t-t_0}{\tau_2}\right)\sum_{p=0}^{k-1}A_{k,p}\left(\frac{t-t_0}{\tau_2}\right)^p, \tag{9}$$

where $A_{k,p}$ are time-independent coefficients, and $A_{k,0}=n_k^b(t_0)$ is the initial condition. Substituting expression

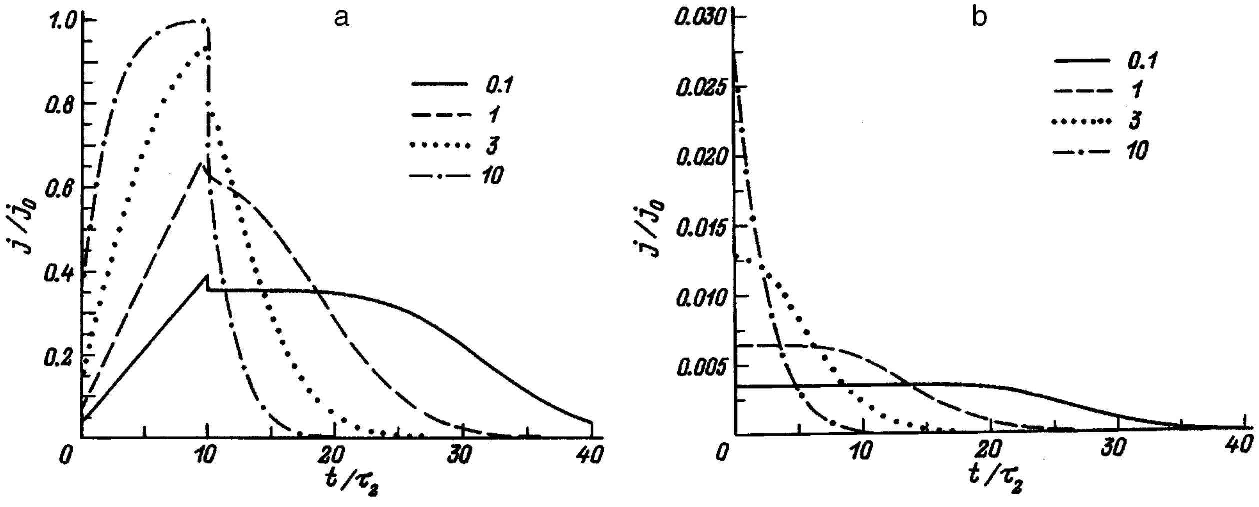

FIG. 2. Response of the current through a layer with $M = 30$ quantum wells to an excitation pulse as a function of the parameter τ_c/τ_d; a — $t_0=10\tau_2$, b — $t_0=0.1\tau_2$; the ratios τ_c/τ_d, corresponding to different curves, are indicated in the figure.

(9) into Eq. (8) and equating to zero the coefficients in front of the powers of $(t-t_0)/\tau_2$, we obtain recurrence relations for the coefficients $A_{k,p}$:

$$A_{k,0}=n_k^b(t_0)\quad \text{for } k=1,2,\dots,M;$$

$$A_{k,p}=\frac{\tau_d}{\tau_c}\frac{1}{1+\tau_d/\tau_e}\frac{1}{p}A_{k-1,p-1}+\frac{1}{1+\tau_d/\tau_c}A_{k-1,p}$$

$$\text{for } k=2,3,\dots,M \quad \text{and}$$

$$p=1,2,\dots,k-2;$$

$$A_{k,p}=\frac{\tau_d}{\tau_c}\frac{1}{1+\tau_d/\tau_c}\frac{1}{p}A_{k-1,p-1}\quad \text{for}$$

$$k=2,3,\dots,M \quad \text{and} \quad p=k-1. \tag{10}$$

To calculate the slow relaxation after the generation pulse is switched off (the section *4–5* in Fig. 1), first we find the carrier redistribution over the wells from Eqs. (9) and (10). Next, the redistribution of the carriers in the delocalized states above the wells is obtained from Eq. (7). We thus find the current flowing through the QW layer from the formula[4]

$$j(t)=\frac{e\nu}{M+1}\sum_{k=0}^{M} n_k(t). \tag{11}$$

Here $\nu=d/\tau_d$, and d is the period of the structure. On the slow-relaxation section, during the carrier generation pulse (the section *2–3* in Fig. 1), the value of the current density is obtained by substituting the carrier distribution above the wells (4) into expression (11).

The behavior of the current in time is shown in Figs. 2a and 2b for carrier generation pulse durations $\tau_2<t_0<M\tau_2$ and $t_0\ll\tau_2$, respectively. The formation of a plateau on the current curve for small ratios τ_c/τ_d after the generation pulse is switched off is attributable to the fact that in accordance with expressions (3) and (4), for short generation pulses the distribution over the wells is of an exponential character and over some period of time the total number of carriers contributing to the current changes very little. The characteristic change in the carrier distribution over the wells after the generation pulse ends is shown in Fig. 3. One can see from the figure that after generation is switched off $(t=t_0)$ the distribution of carriers in the wells is carried out of the layer of QWs in the form of a wave and the current through the structure starts to drop only when the distribution function reaches the boundary of the layer. The presence of a plateau in $j(t)$ is important for separating experimentally fast and slow relaxation processes in the method described below for determining the well parameters.

5. SHORT GENERATION PULSES

The expressions which are obtained and which describe the response of the system to an excitation pulse can be

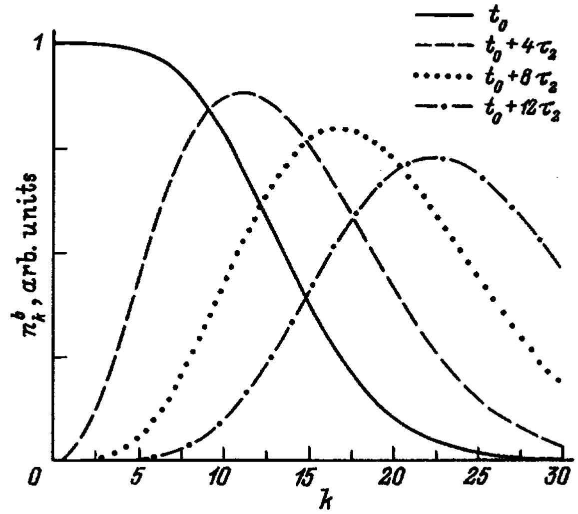

FIG. 3. Character of the variation of the carrier distribution over the wells at different times after the generation pulse stops for $M=30$, $t_0=10\tau_2$, and $\tau_c/\tau_d = 0.3$.

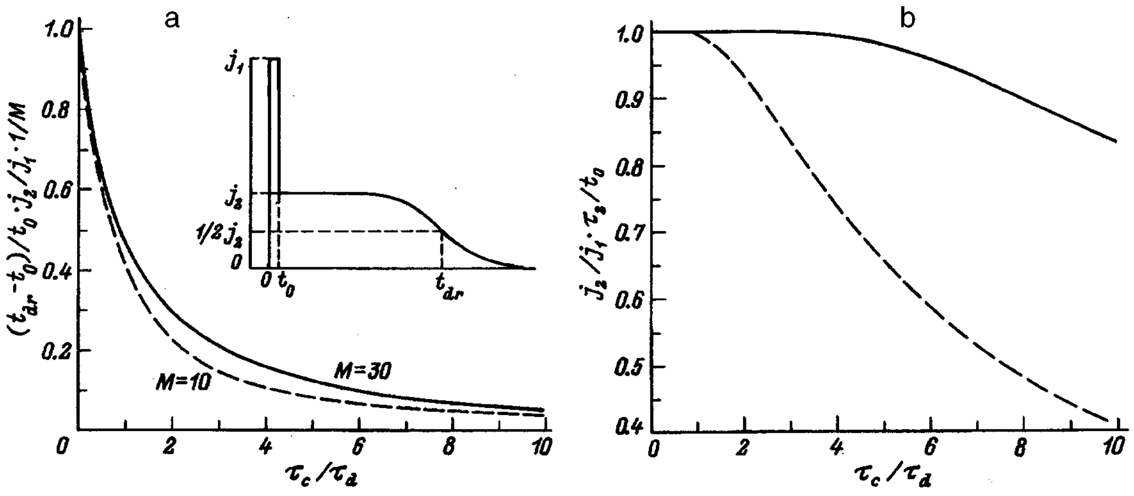

FIG. 4. Curves relating the characteristic emission time from a well, capture time, and drift time with the parameters of the current response to a short excitation pulse ($t_0 \ll \tau_2$) for $M=30$ (solid line) and $M=10$ (dashed line). Inset: Schematic diagram of the current response to an excitation pulse.

simplified substantially and, as already noted, become more convenient for comparison with experimental data in the case of short carrier generation pulses with duration t_0, which satisfy the inequalities $\tau_1 \ll t_0 \ll \tau_2$. For such pulses we can disregard the slow relaxation processes occurring upon removal of the excitation.

The current response to an excitation pulse is shown schematically for this case in the inset in Fig. 4a. When the generation pulse is switched on, the current jumps by an amount j_1 determined by Eqs. (3) and (11):

$$j_1=\frac{e\nu n_0}{M+1}\sum_k\left(\frac{1}{1+\tau_d/\tau_c}\right)^k=\frac{e\nu n_0}{M+1}\,\frac{1+\tau_d/\tau_c}{\tau_d/\tau_c}\times\left[1-\left(\frac{1}{1+\tau_d/\tau_c}\right)^{M+1}\right]. \tag{12}$$

The number of carriers in the wells at the moment the generation pulse is switched off ($t=t_0$) is determined by the expression

$$n_k^b(t_0)=n_k\frac{t_0}{\tau_c}. \tag{13}$$

Expression (13) was obtained from the starting equations (2) under the condition that no carriers are emitted from the wells over the pulse passage time; n_k is given by Eq. (3). We obtain the following expression for the quasisaturation current j_2, after the generation pulse ends and after fast relaxation has occurred, and there is not enough time for the number of carriers in the wells to change:

$$j_2=\frac{e\nu n_0}{M+1}\,\frac{t_0}{\tau_2}\,\frac{\tau_d}{\tau_c}\sum_k k\left(\frac{1}{1+\tau_d/\tau_c}\right)^k=\frac{e\nu n_0}{M+1}\,\frac{1+\tau_d/\tau_c}{\tau_d/\tau_c}\,\frac{t_0}{\tau_2}\times\left\{1-\left[(M+1)\frac{\tau_d}{\tau_c}+1\right]\left(\frac{1}{1+\tau_d/\tau_c}\right)^{M+1}\right\}. \tag{14}$$

6. METHOD FOR DETERMINING THE QW PARAMETERS

It is obvious from relations (12) and (14) that in the case of excitation by short pulses the ratio $(j_2/j_1)(\tau_2/\tau_1)$ depends on the parameter τ_c/τ_d and the number M of wells. This dependence, shown in Fig. 4b, gives a unique relation between the emission time τ_e from a well and the parameter τ_c/τ_d.

The slow relaxation processes after the generation pulse ends ($t>t_0$) are described by expression (9) with the initial condition (13). The current through the layer of quantum wells can be written in the general form as

$$\frac{j(t)}{j_2}=\exp\left(-\frac{t-t_0}{\tau_2}\right)\sum_k B_k\left(\frac{t-t_0}{\tau_2}\right)^k, \tag{15}$$

where B_k are the expansion coefficients, which depend only on the parameter τ_c/τ_d. The duration of the response of the system to a short carrier generation pulse is found as the solution of the equation $j((t_{dr}-t_0)/\tau_2)/j_2=1/2$, depending on the parameter τ_c/τ_d and the number of wells. The quantity

$$\frac{t_{dr}-t_0}{\tau_2}\,\frac{j_2}{j_1}\,\frac{\tau_2}{t_0}=\frac{t_{dr}-t_0}{t_0}\,\frac{j_2}{j_1}=\mathrm{const}\,(t_0) \tag{16}$$

is shown in Fig. 4a as a function of the parameter τ_c/τ_d. This dependence makes it possible to obtain the ratio τ_c/τ_d from the experimentally determined current response parameters.

We note that as the parameter τ_c/τ_d approaches zero, i.e., in the case where the carriers are rapidly captured by the wells, the drop-off time $t_{dr}-t_0$, as one can see from Fig. 4a, approaches the value $M\tau_2$. The same result can be obtained from the assumption that each carrier visited all wells in succession.

The following procedure for analyzing the experimental data follows from what has been said above. Having determined the parameter (16) from experiment, we find with the

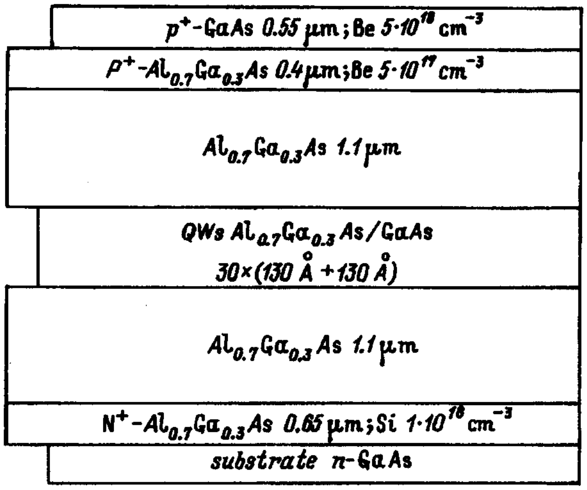

FIG. 5. Schematic representation of the experimental structure.

aid of Fig. 4a the ratio τ_c/τ_d, after which τ_2 is determined on the basis of Fig. 4a, i.e., ultimately the emission time τ_e from a QW.

The quantity (16) is found from the experimental data and should not depend on the generation pulse duration τ_0. This can serve as a criterion of applicability of the theoretical expressions given above for describing relaxation processes in the experimental structure.

We assumed above that the generation pulse is square and that it has ideal edges, i.e., it is switched on and off over times much shorter than τ_1. However, in the case where the duration t_f of the leading edge satisfies the condition $t_1 \ll \tau_f \ll \tau_2$, the same expressions can be used, assuming the leading edge of the pulse is a sequence of infinitely small jumps with characteristic times τ_1. In this case, the fast relaxation processes will occur over times of the order of τ_f. Therefore, in the case of pulses with duration $\tau_1 \ll t_0 \ll \tau_2$, the duration τ_f of the leading edge can be comparable to the pulse duration. Then the same expressions can be used to describe the slow relaxation processes, i.e., the slow component of the response of the current flowing through the QW layer for this case does not depend on the shape of the excitation pulse.

7. EXPERIMENTAL RESULTS

The experimental structure, grown by molecular-beam epitaxy in a TsNA-4 apparatus, consisted of $M=30$ QWs in a GaAs/Al$_x$Ga$_{1-x}$As system. The parameters of the layers in the grown structure are shown in Fig. 5. A (100) n-GaAs substrate was used for growth. All layers between the N^+ and P^+ emitters were not specially doped and actually possessed a weak n-type conductivity. To optimize and check the conditions of growth during the epitaxy process, the intensity of the directly reflected, high-energy, electron diffraction (RHEED) reflection was recorded.[9] This made it possible to select a temperature, growth rate, and excess As pressure above the group-III elements. As a result, the nonradiative recombination rate in the QWs could be reduced to a minimum and sharp heterojunctions could be obtained.[9]

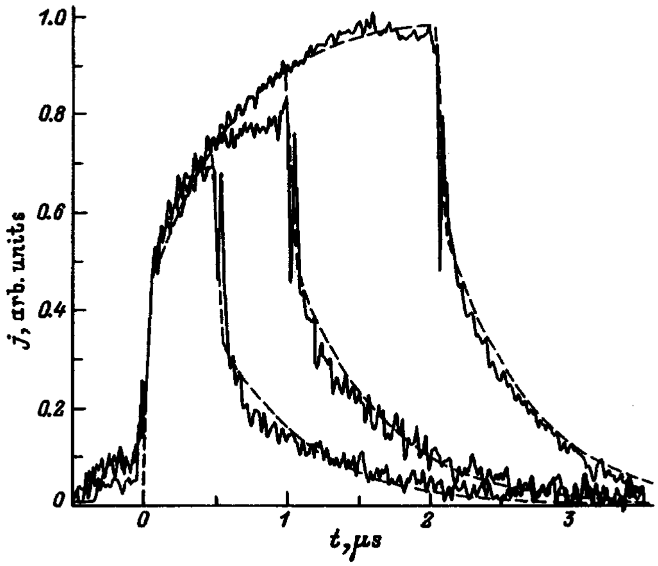

FIG. 6. Kinetics of the electron current through a system of QWs with a voltage of 20 V for generation pulses with duration $t_0=0.51$, 1.03, and 2.08 μs. Solid curves — experimental results, dashed curves — theoretical calculation with $\tau_2=340$ ns and $\tau_c/\tau_d=15$.

Carrier transport in the system of QWs was investigated by making time-resolved measurements of the current induced by an electron probe.[10] The application of this method, instead of the conventional excitation of carriers by light, made it possible to avoid producing additional narrow-gap layers (carrier generators) and additional barriers separating them from the system of QWs. Electron-hole pairs were excited in the undoped region of the top wide-gap emitter by a beam of electrons with energy of the order of 10 keV, which is greater than the excitation energy of an electron-hole pair. This makes it possible to disregard the contribution of the primary electrons to the current response. The current of the electronic probe was chose to be at the level 5×10^{-9} A. Under such conditions the current response to the excitation pulse ceased to depend on the probe current at voltages of 16–18 V on the structure. It can be assumed, therefore, that for voltages on the structure exceeding 16–18 V the processes in the QW region do not depend on the carrier injection level.

Measurements of the current through the structure under the conditions of excitation with short pulses were performed with a voltage of 20 V on the structure (which corresponded to electric fields of $\sim 7\cdot10^4$ V/cm in the QW region) and generation pulse durations $t_0=0.51$, 1.03, and 2.08 μs. Since the electron-hole pairs were excited in the wide-gap emitter near the P^+ layer, it can be assumed that hole emission occurs over times much shorter than the electron emission time, and the hole current contributes only to the fast component of the current. Therefore, the electron current can be obtained from the curve of the total current through the structure by subtracting the hole current — a square pulse with duration t_0. The amplitude of the subtracted pulse is obtained from the condition that the number of electrons equals the number of holes. The electron current obtained in this manner through the system of QWs is shown in Fig. 6 by the solid line.

The theoretical curves (Fig. 6, dashed line) were obtained in accordance with the theory described above by optimizing the parameters τ_2 and τ_0/τ_d for all three curves simultaneously. The best agreement with theory was obtained for $\tau_2=340$ ns and $\tau_c/\tau_d=15$, which gives $\tau_e \sim 300$ ns. Such a large value of τ_e for electrons in the structure under study apparently stems from the fact that when an electron is emitted from a QW, it must overcome not only the potential barrier, but also pass into a different valley.[11]

In summary, the first preliminary experimental results have shown that it is in principle possible to determine the characteristic carrier transport times in a system of QWs by comparing the measured current kinetics with the kinetics calculated in accordance with the theory presented in Secs. 2–3.

The ultrashort excitation pulse method described in Secs. 5–6 cannot be used for analysis in the case of the present structure because the signal-to-noise ratio is too low. Furthermore, as one can see from Fig. 2b, for large values of τ_c/τ_d a plateau is not formed on the current curve. This makes it much more difficult to perform measurements of the parameters of the response to an ultrashort excitation pulse.

8. CONCLUSIONS

A phenomenological theory describing nonstationary current transport processes in a system of QWs with short generation pulses was developed in this paper. The proposed theory makes it possible to determine the characteristic carrier transport times by comparing the current kinetics with the theoretical calculations. The special case of ultrashort generation pulses was examined separately. The present method includes a procedure for checking the applicability of the theory developed for describing carrier transport for a specific experimental QW system. The first experimental results showing that the characteristic carrier transport times can in principle be determined by comparing the experimental curves with the theoretical calculations were presented.

The theoretical part of this work was supported by the Volkswagen–Stiftung Foundation and the experimental part was supported by INTAS (Grant INTAS-93-2687).

We thank J. Gutowski and A. C. Walker for a helpful discussion.

[1] G. D. Boyd, D. A. B. Miller, D. S. Chemla, S. L. McCall, A. C. Gossard, and J. H. English, Appl. Phys. Lett. **50**, 1119 (1987).

[2] D. A. B. Miller, Opt. Quant. Electron. **22**, S61 (1990).

[3] D. C. Hatchings, C. B. Park, and A. Miller, Appl. Phys. Lett. **59**, 3009 (1991).

[4] S. S. Kucherenko, B. S. Ryvkin, and A. Ya. Shik, Fiz. Tekh. Poluprovodn. **29**, 487 (1995) [Semiconductors **29**, 252 (1995)].

[5] S. V. Kozyrev and A. Ya. Shik, Fiz. Tekh. Poluprovodn. **19**, 1667 (1985) [Sov. Phys. Semicond. **19**, 1024 (1985)].

[6] J. A. Brum and G. Bastard, Phys. Rev. B **33**, 1420 (1986).

[7] P. W. M. Blom, C. Smit, J. E. M. Haverkort, and J. H. Wolter, Phys. Rev. B **47**, 2072 (1993).

[8] A. M. Fox, D. A. B. Miller, G. Livescu, J. E. Cunningham, and W. Y. Jan, IEEE J. Quant. Electron. **27**, 2281 (1991).

[9] D. Z. Garbuzov, V. P. Evtiihiev, N. I. Katsavets, A. B. Komissarov, T. E. Kudrik, I. V. Kudryashov, V. B. Khalfin, R. K. Bauer, Zh. I. Alferov, and D. Bimberg, J. Appl. Phys. **75**, 4152 (1994).

[10] S. G. Konnikov, V. E. Umanskiĭ, V. M. Chistyakov, and I. I. Lodyzhenskiĭ, Fiz. Tekh. Poluprovodn. **22**, 1803 (1988) [Sov. Phys. Semicond. **22**, 1140 (1988)].

[11] F. Beltram, F. Capasso, J. F. Walker, and R. J. Malik, Appl. Phys. Lett. **53**, 376 (1988).

Translated by M. E. Alferieff

Electron localization in sound absorption oscillations in the quantum Hall effect regime

I. L. Drichko, A. M. D'yakonov, A. M. Kreshchuk, T. A. Polyanskaya,[1)] I. G. Savel'ev,
I. Yu. Smirnov, and A. V. Suslov[2)]

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

(Submitted July 22, 1996; accepted for publication September 10, 1996)
Fiz. Tekh. Poluprovodn. **31**, 451–458 (April 1997)

The absorption coefficient for surface acoustic waves in a piezoelectric insulator in contact with a GaAs/$Al_{0.25}Ga_{0.75}As$ heterostructure (with two-dimensional electron mobility $\mu = 1.3\times10^5$ cm^2/(V·s) at $T = 4.2$ K) via a small gap has been investigated experimentally as a function of the frequency of the wave, the width of the vacuum gap, the magnetic field, and the temperature. The magnetic field and frequency dependences of the high-frequency conductivity (in the region 30–210 MHz) are calculated and analyzed. The experimental results can be explained if it assumed that there exists a fluctuation potential in which current carrier localization occurs. The absorption of the surface acoustic waves in an interaction with two-dimensional electrons localized in the energy ''tails'' of Landau levels is discussed. © *1997 American Institute of Physics.* [S1063-7826(97)01504-4]

1. INTRODUCTION

The problems of quantum interference of electrons, i.e., weak localization in a two-dimensional electron gas (2DEG), just as problems of strong charge-carrier localization (associated with defects of the crystal lattice), occupy a central place in two-dimensional nanoelectronics, but they have still not been adequately investigated. The characteristic features of localization in a 2DEG can be studied by investigating rf conductivity of a 2DEG in the quantum Hall effect regime.[1]

One proven method for investigating rf conductivity is the acoustic method, which makes it possible to measure the rf conductivity of semiconductors without any electrical contacts on the sample. The crux of this method is that a piezoelectrically active sound wave accompanied by an electric field undergoes absorption as a result of interaction with electrons as it propagates in the semiconductor, and this absorption is directly related to the electrical conductivity of the sample.

Acoustic methods have been successfully used to study the rf conductivity of a three-dimensional electron gas in doped compensated semiconductors at low temperatures.[2,3] It has been shown that if the electrons are in a free (delocalized) state, then the ultrasonic absorption coefficient of the electrons in the semiconductor, which is a piezoelectric material, in a magnetic field H is determined completely by its dc conductivity $\sigma_{xx}^{dc}(H)$. If, however, electrons are localized at separate impurity centers or in a random fluctuation potential produced by the impurities, as in the case of a strongly doped and strongly compensated semiconductor, then the conduction mechanisms in a constant electric field and in the rf field accompanying the acoustic wave are different. In the static case the conductivity σ_{xx}^{dc} is of a hopping character and the rf conductivity σ_{xx}^{ac} can remain ''metallic,'' and in addition $\sigma_{xx}^{dc} \neq \sigma_{xx}^{ac}$. By studying the ultrasonic absorption coefficient of the electrons as a function of the magnetic field intensity, temperature, ultrasonic frequency, and ultrasonic wave intensity in the three-dimensional case it was possible to determine not only the character of the carrier localization but also the corresponding numerical parameters.

On this basis, it is helpful to extend the acoustic method of investigating rf conductivity to semiconductor structures with a 2DEG. The interaction of surface acoustic waves (SAWs) with two-dimensional electrons was first observed in GaAs/AlGaAs type heterostructures.[4,5] It was shown that the SAW energy is absorbed as a result of this interaction, and the magnitude of the absorption is determined by the conductivity of the 2DEG. For this reason, in a quantizing magnetic field, when Shubnikov–de Haas (SdH) oscillations appear, the absorption coefficient for the surface acoustic wave also oscillates. The oscillations of this coefficient are sharpest in the region of magnetic fields where the quantum Hall effect regime obtains. In this connection, by analogy with the three-dimensional semiconductors, the characteristic features of localization in a 2DEG in a quantum Hall effect regime can be studied by studying the temperature, frequency and magnetic field dependences of the sound absorption and, which is especially important, this can be done by studying the conductivity measured by a contact-free method.

Two experimental configurations are now employed. In the first one SAWs are excited in the standard manner on the surface of a piezoelectric crystal (for example, $LiNbO_3$), and the experimental structure with the 2DEG is placed above the surface of a piezodielectric crystal with a gap, whose width is less than the wavelength. In the second configuration the experimental structure itself, which must be a piezoelectric, serves as the sound duct. Surface acoustic waves are excited in it and at the same time the interaction of sound with the 2DEG is investigated. The first configuration has the advantage that the gap eliminates the deformation interaction of the two-dimensional (2D) electrons with SAWs, and only the interaction with the electric fields produced by the SAWs in the piezoelectric crystal remains. However, this method is inconvenient because it is almost impossible to determine the width of the gap by direct measurements. In experiments employing the second configuration there is no gap at all, but

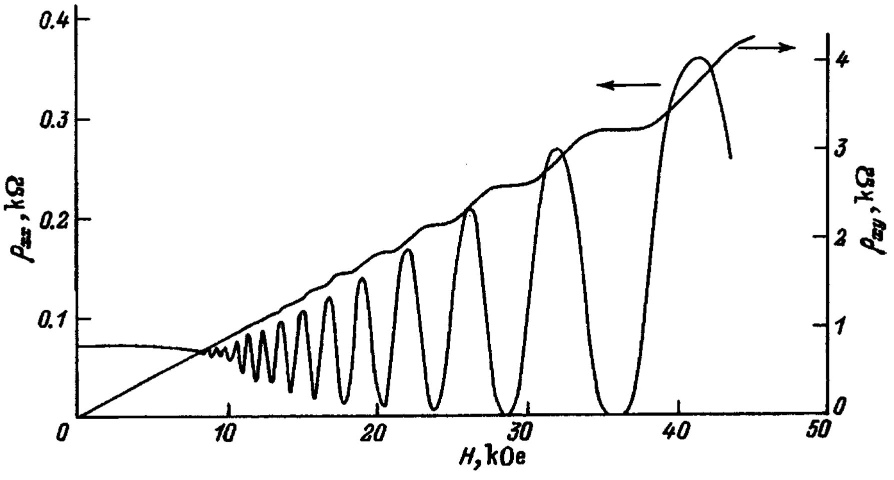

FIG. 1. Resistivity tensor components ρ_{xx} and ρ_{xy} versus the magnetic field H in a GaAs/$Al_{0.25}Ga_{0.75}As$ heterostructure at $T=4.2$ K.

in this case the deformation interaction of the 2D electrons with the SAWs cannot be ignored.[6]

The first configuration was used in Ref. 4 and the second was used in Refs. 5–8. In all of these investigations it was concluded that the character of the dependence of the SAW absorption coefficient of the two-dimensional electrons on the magnetic field H is determined by the dc conductivity σ_{xx}^{dc} and its dependence on H. However, in Ref. 9, where the SAW absorption of a 2DEG in a $In_{0.47}Ga_{0.53}As$/InP heterostructure was measured in the first experimental configuration with $T=4.2$ K, it was shown that this assertion is valid only for high Landau-level numbers and in stronger magnetic fields (i.e., small filling numbers $\nu = nhc/2eH$, where n is the Hall electron density) the variation of the SAW absorption coefficient as a function of H is not determined by $\sigma_{xx}^{dc}(H)$.

In the present work the dependences of the SAW absorption coefficient Γ of 2D electrons on the magnitude and direction (longitudinal and transverse with respect to the normal to the plane of the 2DEG) of the magnetic field up to $H \simeq 60$ kOe, the SAW frequency in the range $f=30-210$ MHz, and the temperature in the range $T=1.4-4.2$ K are investigated using the first experimental configuration mentioned above for the purpose of studying the rf conductivity and the character of electron localization in structures of this type in the integral quantum Hall effect regime. To get a better physical picture of the phenomena, galvanomagnetic measurements of the components of the dc resistivity tensor were performed as a function of the magnetic field up to $H \simeq 50$ kOe.

2. EXPERIMENTAL PROCEDURE AND RESULTS

The GaAs/$Al_{0.25}Ga_{0.75}As$ structures investigated were prepared by molecular-beam epitaxy. A layer of gallium arsenide with residual impurity density 10^{15} cm^{-3} and thickness of the order of 1 μm (the so-called buffer layer) was grown on a semi-insulating GaAs (001) substrate of thickness $d \simeq 300$ μm. Next, a 50-Å-thick undoped layer of the solid solution $Al_{0.25}Ga_{0.75}As$ (so-called spacer) and an approximately 500-Å-thick layer of a solid solution of the same composition but doped with silicon with an electron density $n \simeq 10^{18}$ cm^{-3} were grown. The top (contact) layer of the structure consisted of a 200-Å gallium arsenide film doped with silicon $[n \simeq (3-4) \times 10^{18}$ cm$^{-3}]$.

The samples for the dc galvanomagnetic measurements and for measurements by acoustic methods were prepared from neighboring regions of the heterostructure. The measurements were performed with dc current $I \simeq 10-20$ μA (which ensured that there would be no heating effects) in the temperature range 1.8–4.2 K on mesa structures etched out in the form of a double cross with contact pads. We present below the results of our investigations for a structure with Hall density $n=6.9\times10^{11}$ cm^{-2} and Hall mobility $\mu=1.3\times10^{5}$ cm^2/(V·s) at $T=4.2$ K. The density calculated according to the SdH oscillations was $n_{SH}=6.8\times10^{11}$ cm^{-2}. The results of the measurements of the components ρ_{xx} and ρ_{xy} of the resistivity tensor as a function of the magnetic field H (in the quantum Hall effect region) at $T=2.1$ K are presented in Fig. 1.

A piezoelectrically active surface wave, excited on the Y-cut surface of a $LiNbO_3$ crystal (see inset in Fig. 2) and propagating in the Z direction, was used for the acoustic

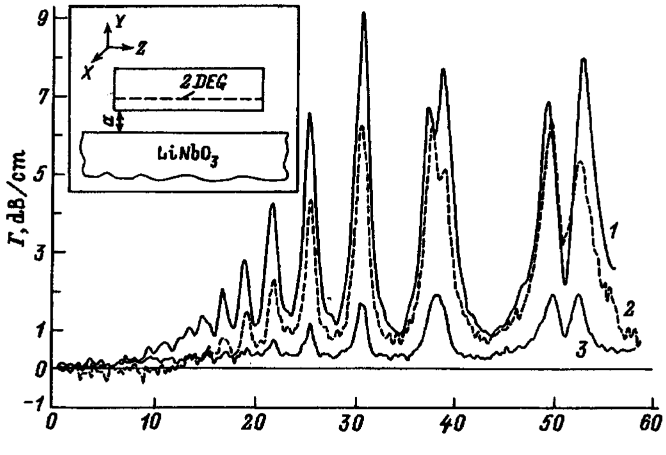

FIG. 2. Surface acoustic wave absorption coefficient Γ versus the magnetic field intensity H at $f=30$ MHz with vacuum gap widths a (μm): *1*—0.3, *2*—0.5, and *3*—1.0. Inset: Arrangement of the sample with respect to the crystal axes of lithium niobate.

measurements. The SAW was produced and detected with interdigital transducers with fundamental frequency 30, 90, and 150 MHz. In addition, higher-order harmonics were also used in the experiment. The sample was placed directly on the lithium niobate surface, along which a SAW propagated, and detected with the aid of a spring. The sample was located in a cryoduct in a vacuum chamber. The vacuum gap between the piezodielectric and the sample, determined by the roughness of their surfaces, is designated by the letter a in the diagram shown in the inset in Fig. 2. A longitudinal electric field (along the direction of propagation of the acoustic wave), which arises with the motion of the SAW in the piezodielectric crystal, penetrates into the experimental object, and the 2DEG is then located in an alternating electric field whose frequency equals the frequency of the surface acoustic wave. The interaction of the 2D electrons in the experimental heterostructure with the SAW electric field causes the wave to decay, and this decay is recorded by the transducer employed. The independence of the absorption coefficient Γ from the intensity of the sound wave was checked in the experiment, i.e., the measurements were performed in a regime linear in the electric field.

Figure 2 shows the dependence of the absorption coefficient for a SAW with frequency $f=30$ MHz on the intensity of a magnetic field, oriented in a direction normal to the surface of the heterostructure, at $T=4.2$ K and for different vacuum gap widths a. Similar curves were also obtained for other frequencies and temperatures. An interesting feature is observed on the curves in Fig. 2: In strong magnetic fields the maxima of the absorption Γ split and the lower the temperature, the lower the magnetic field intensity at which splitting is observed. The maxima of Γ are equally spaced as function of $1/H$ and for large Landau numbers they correspond, with respect to the magnetic field, to the minima of $\sigma_{xx}^{dc}(H)$; for higher fields H, when σ_{xx}^{dc} becomes small at the minima of the oscillations ($<10^{-7}\ \Omega^{-1}$) the absorption maxima split in two and the conductivity minimum (and also the center of the Hall plateau) corresponds to an absorption minimum.

We also measured the absorption of SAWs in a magnetic field oriented parallel to the surface of the sample. In this geometry the SAW absorption oscillations were not observed. This confirms the two-dimensional character of the absorption oscillations which we observed in a transverse magnetic field.

An interesting feature of the absorption of SAWs in strong magnetic fields is the inequality, observed for small filling numbers, in the amplitudes of the SAW absorption maxima in the presence of splitting. In Ref. 5 this fact was attributed to the inhomogeneity of the experimental samples. We performed a very large number of measurements on several GaAs/AlGaAs differently positioned (with different vacuum gaps) samples, and we observed both a different asymmetry of the peaks and the same height of the peaks on the same sample; this is illustrated in Fig. 2. It is not yet clear what causes this effect.

3. ANALYSIS OF THE EXPERIMENTAL DATA

The equal spacing of the peaks in the dependence of the absorption of surface acoustic waves (SAWs) on the quantity $1/H$ makes it possible to determine the density of 2DEG from the period of the oscillations by the standard method. The density obtained in this manner equals, within the limits of accuracy of the experiment, the density n_{SH} calculated according to the SdH oscillations.

The absorption coefficient Γ, which is associated with the interaction of a SAW with two-dimensional electrons, in the experimental configuration employed in our work was calculated in Ref. 10, taking into account the diffusion of current carriers. If the absorption coefficient Γ is introduced as $A=A_0\exp(-\Gamma x)$, where A_0 and A are the signal amplitudes at the entrance and exit, respectively, and x is the length of the sample, then Γ is (in cm^{-1})

$$\Gamma=\chi^{(1)}k\frac{\pi\sigma f_2(k)/\varepsilon_s\nu}{1+[Dk/\nu+2\pi\sigma f_1(k)/\varepsilon_s\nu]^2}. \tag{1}$$

In Ref. 10 $\chi^{(1)}\simeq K^2$ is the electromechanical coupling constant of LiNbO$_3$; k and ν are, respectively, the SAW wave vector and velocity; D is the diffusion coefficient; ε_s is the permittivity of the semiconductor; and σ is the conductivity of the 2DEG. We shall estimate the contribution of diffusion to the absorption coefficient without expanding the functions f_1 and f_2 (expressions for which are presented in Ref. 10). For a degenerate 2DEG the diffusion coefficient is

$$D=\sigma\pi\hbar^2/e^2m^*=\pi\sigma a_{\mathrm{B}}/\varepsilon_s, \tag{2}$$

where m^* is the electron effective mass, and a_B is the effective Bohr radius (for electrons in GaAs $a_B\simeq 100$ Å). We note that for all frequencies and vacuum gap widths employed in the experiment, the value of f_1 calculated according to Eq. (10) from Ref. 10 ranged from 0.6 to 1.4, so that the two terms in the denominator of Eq. (1) satisfy the relation

$$D\frac{k}{\nu}\ll\frac{2\pi\sigma}{\nu\varepsilon_s}f_1 \tag{3}$$

(which reduces to the inequality $k\ll 2f_1/a_{\mathrm{B}}$), and it is natural to use below the formula for the absorption coefficient neglecting the diffusion term. If expression (1) is rewritten in the ''symmetric'' form, setting σ equal to σ_{xx}—the conductivity of a 2DEG in a magnetic field and assuming the distance from the channel with the 2DEG to the surface of the sample equals 0 (which corresponds to the experimental structure), then Γ in dB/cm is given by

$$\Gamma=34.72(K^2)k\exp(-2ka)(\varepsilon_1+\varepsilon_0)\varepsilon_s\varepsilon_0^2b(k,a)$$
$$\times\frac{(4\pi\sigma_{xx}/\varepsilon_s\nu)r(k,a)}{1+[(4\pi\sigma_{xx}/\varepsilon_s\nu)r(k,a)]^2}, \tag{4}$$

where $\varepsilon_1=51$ (Ref. 10), $\varepsilon_s=12$, and $\varepsilon_0=1$ are the permittivities of LiNbO$_3$, GaAs, and of the vacuum gap between the sample and LiNbO$_3$ respectively;

$$b(k,a)=\{[c-t\exp(-2ka)][c+m-t\exp(-2ka)$$
$$-p\exp(-2ka)]\}^{-1},$$

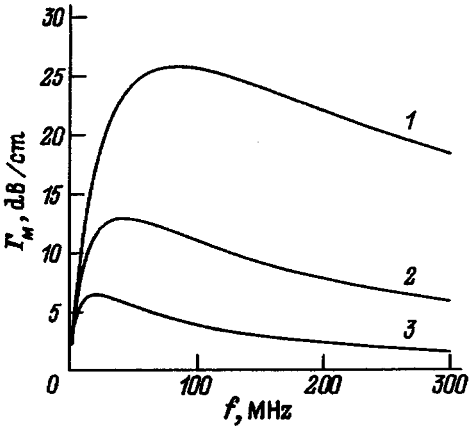

FIG. 3. Theoretical value of the maximum absorption Γ_M versus acoustic wave frequency f for different vacuum gap widths a (μm): *1*—0.3, *2*—0.6, *3*—1.2.

$$r(k,a)=\frac{c+m-t\exp(-2ka)-p\exp(-2ka)}{2[c-t\exp(-2ka)]},$$

where $c=(\varepsilon_1+\varepsilon_0)(\varepsilon_s+\varepsilon_0)$, $t=(\varepsilon_1-\varepsilon_0)(\varepsilon_s-\varepsilon_0)$, $m=(\varepsilon_1+\varepsilon_0)(\varepsilon_s-\varepsilon_0)$, and $p=(\varepsilon_1-\varepsilon_0)(\varepsilon_s+\varepsilon_0)$.

Let us now analyze Γ as a function of σ_{xx} in accordance with Eq. (4). If the conductivity is high, then

$$(4\pi\sigma_{xx}/\varepsilon_s v)r(k,a)\gg 1 \tag{5}$$

and $\Gamma\sim 1/\sigma_{xx}$; if (for low conductivity)

$$(4\pi\sigma_{xx}/\varepsilon_s v)r(k,a)\ll 1, \tag{6}$$

then the absorption coefficient $\Gamma\sim\sigma_{xx}$. Therefore, it is obvious that for

$$(4\pi\sigma_{xx}/\varepsilon_s v)r(k,a)=1 \tag{7}$$

the function $\Gamma(H)$ possesses a maximu Γ_M, and from Eqs. (4) and (7) it is obvious that the value of Γ_M does not depend on the value of σ_{xx}. The equality (7) in our case corresponds to

$$\sigma_{xx}r(k,a)=v,$$

since $4\pi/\varepsilon_s\simeq 1$ for GaAs. Let us now consider the dependence of the maximum absorption coefficient Γ_M on the wave frequency f, which follows from Eq.(4) and which is shown in Fig. 3 for different values of a. It is obvious from Fig. 3 that the dependence $\Gamma_M(f)$ has a maximum, and the width of the vacuum gap increases with decreasing value of Γ_M, so that the maximum occurs at a lower frequency.

If we now refer to Fig. 2, then on the basis of the foregoing analysis of Eq. (4) it is possible to explain the splitting of the maxima of Γ at low filling numbers. Indeed, near the conductivity maxima, where $\sigma_{xx}\simeq 10^{-5}\ \Omega^{-1}$, the condition (5) is satisfied and the absorption coefficient Γ assumes its minimum values. As the magnetic field increases, within the same Landau level the quantity σ_{xx} starts to decrease rapidly, and the coefficient Γ in accordance with Eq. (5) increases, until σ_{xx} decreases to a value corresponding to condition (7). In this case $\Gamma(H)$ reaches its maximum value Γ_M. As σ_{xx} decreases further, the condition (6) is satisfied and the coefficient Γ, correspondingly, also decreases. This occurs right up to values of H for which $\sigma_{xx}(H)$ reaches its minimum value. As the magnetic field increases further, σ_{xx} again increases, the changes in Γ occur in the reverse order. As a result, a second maximum appears in the dependence $\Gamma(H)$. In weaker magnetic fields the quantity $\sigma_{xx}(H)$ does not reach at the minima the values for which the equality (7) holds. For this reason, splitting of the maxima of $\Gamma(H)$ does not occur. As the temperature decreases, the oscillations of $\sigma_{xx}(H)$ become sharper and deeper, which is reflected in the corresponding curves for $\Gamma(H)$ as a splitting of the maxima with the larger numbers.

To calculate from Γ the dissipative conductivity $d\sigma_{xx}$, it is necessary to know the gap width a between the insulator and the experimental object. In our case the heterostructure is pressed directly to the lithium niobate surface and a is a poorly controllable parameter, since it is determined by irregularities of unknown amplitude on both surfaces. The effective value of a can be determined from acoustic measurements by investigating the frequency dependence of the maximum value of the absorption coefficient Γ_M. For the same placement of the sample, i.e., for the same value of a but different SAW frequencies, we obtain [see Eqs. (4) and (7)]

$$\frac{\Gamma_M(k_1)}{\Gamma_M(k_2)}=\frac{k_1}{k_2}e^{-2a(k_1-k_2)}\frac{b(a,k_1)}{b(a,k_2)}. \tag{8}$$

The value of a determined from this equation was found to be 0.25–1 μm for different placements of the sample. Knowing a, it is possible to determine from relation (4) and the experimentally measured values of Γ the conductivity σ_{xx}^{ac} and its dependence on the magnetic field. It was found that this quantity, which was determined from the acoustic measurements for different values of a in the range of magnetic fields corresponding to absorption coefficient minima (conductivity maxima), corresponds within 20% to the values of σ_{xx}^{dc} (for the same values of H) which were calculated from the galvanomagnetic measurements of the dependences $\rho_{xx}(H)$ and $\rho_{xy}(H)$ in a direct current:

$$\sigma_{xx}^{dc}(H)=\frac{\rho_{xx}(H)}{[\rho_{xy}(H)]^2+[\rho_{xx}(H)]^2}. \tag{9}$$

This equality apparently means (by analogy with the three-dimensional case) that in the indicated range of magnetic fields the static conductivity σ_{xx}^{dc} is determined by carriers in delocalized states.

To analyze further the experimental SAW absorption curves we employed the vacuum gap a as a parameter, choosing it so that near the conductivity maxima σ_{xx}^{ac} would equal σ_{xx}^{dc}. In this method of determination, the vacuum gap a differs by no more than 20% from the gap calculated from Eq. (8).

4. DISCUSSION

Figure 4 shows the magnetic field dependences of the dissipative conductivity $\sigma_{xx}^{ac}(H)$, which were obtained from the SAW absorption coefficient for different frequencies and gap widths a, as well as the dependence $\sigma_{xx}^{dc}(H)$ near the maximum with filling number $\nu=nhc/2eH=3.5$, where n is

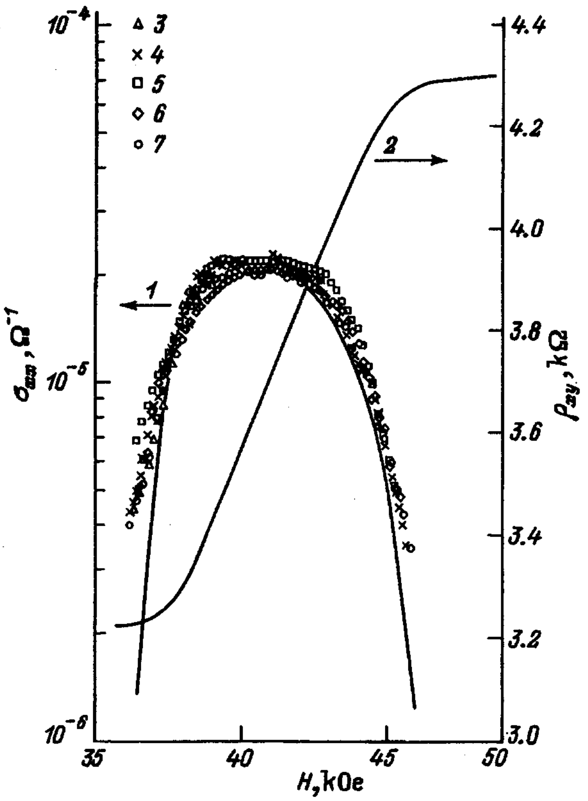

FIG. 4. σ_{xx}^{dc} (solid line) and σ_{xx}^{ac} (symbols) (*1*) versus the magnetic field H near the region of delocalized states; *2*—$\rho_{xy}(H)$. The symbols correspond to frequencies f (MHz) and vacuum gap widths a (in μm): *3*—213 and 0.3, *4*—30 and 0.5, *5*—150 and 0.3, *6*—30 and 0.4, *7*—90 and 1.2.

the Hall electron density. It should be noted that if the value of a determined by the method described in the preceding section is substituted into Eq. (4) and the value of σ_{xx}^{dc} measured in a direct current is used for σ_{xx}, then it is not possible to describe the entire experimental dependence $\Gamma(H)$.

It is obvious from Fig. 4 that the conductivities are equal to each other in magnetic fields in the range 37–45 kOe, but as the magnetic field moves away from the maximum in different directions, they separate, and σ_{xx}^{ac} is always greater than σ_{xx}^{dc}. In our view, the difference of the rf and static conductivities describes the transition through the mobility threshold from free to localized electronic states and correspondingly to different mechanisms of conduction in static and rf electric fields.

To explain the experimental data on SAW absorption, we assume that the Landau levels are smeared in space by the fluctuation potential which may not be small. Indeed, in the best GaAs/AlGaAs-type structures the mobility of the 2DEG at $T=4.2$ K reaches values of $\mu \simeq 10^7$ cm^2/(V·s), i.e., it is two orders of magnitude greater than the mobility in the structure investigated by us. The simplest explanation for the difference in σ_{xx}^{ac} and σ_{xx}^{dc} could be as follows: Far from the maxima the conductivity σ_{xx} is of the character of hops along localized states and should therefore increase with frequency f as $\sigma \sim f^s$ $(s=1)$.[1] However, as we shall show below, this is at variance with the experimental conclusion that in measurements of γ at different frequencies and calculations of σ_{xx}^{ac} from them we did not observe a frequency dependence of σ_{xx}^{ac} within the limits of the experimental error, i.e., according to our data $s \simeq 0$. Therefore, a more adequate model is the model of a large-scale fluctuation potential and the assumption that for 2DEG a ‘‘mobility threshold’’ exists in our samples. As a result, in the range of magnetic fields where the Fermi level lies above the mobility threshold (region of maximum σ_{xx}), the electrons are delocalized and $\sigma_{xx}^{dc}(H) = \sigma_{xx}^{ac}(H)$. As the magnetic field varies, the Fermi level crosses the mobility threshold, causing a percolation-type transition, in which the electrons are localized in the random potential, forming so-called ‘‘lakes’’ with metallic-type conductivity. In the latter case the static conductivity mechanism becomes different from the rf conductivity mechanism: The static conductivity is determined by the activation transfer of electrons to the percolation level and the rf conductivity, whose magnitude is larger than that of the static conductivity, is determined by the conductivity in the lakes (along closed orbits for free 2D electrons, determined by the relief formed by fluctuations of the potential energy). This model is supported by the fact that in the range of magnetic fields where $\sigma_{xx}^{dc}(H) = \sigma_{xx}^{ac}(H)$ the quantity ρ_{xy} increases abruptly at a transition between two quantum plateaus (Fig. 4), which corresponds to the position of the Fermi level in the region of delocalized states.[11]

It is very difficult to calculate the SAW absorption coefficient of 2D electrons localized in a random potential. For this reason, we assume, by analogy with the three-dimensional case,[3] that for the two-dimensional conductivity the formula for Γ, which describes SAW absorption in this case, will have the same structure as Eq. (4), but a factor Σ, which depends on the magnetic field and temperature, appears in this case in front of the fraction. For a three-dimensional electron gas, the physical meaning of the coefficient $\Sigma(H)$ is that of a relative volume occupied by the conducting ‘‘drops,’’ whereas in our case (2DEG) $\Sigma(H)$ is approximately proportional to the density of electrons which are localized on the minima of the fluctuation potential (lakes) and which realize in the lakes a conduction in the electric field of the SAW. As a result,

$$\Gamma = \Sigma(H)\Gamma(\sigma_{xx}^{l}), \tag{10}$$

where $\Gamma(\sigma_{xx}^{l})$ corresponds to expression (4) with σ_{xx} replaced by σ_{xx}^{l}—the conductivity in the lakes. In the range of fields H where the electrons are delocalized, the function $\Sigma(H)=1$ and

$$\sigma_{xx}^{dc}(H) = \sigma_{xx}^{ac}(H),$$

and then the value of Σ becomes less than 1. Therefore, to calculate $\sigma_{xx}^{ac}(H)$ in the entire range of magnetic fields, it is necessary to know the value of Σ and its magnetic-field dependence.

The following operation was performed to determine $\Sigma(H)$: The ratio of the experimentally measured values of $\Gamma(H)$ and $\Gamma(H_{\max})$ was determined near each maximum of σ_{xx} with respect to the magnetic field in the range of magnetic fields where $\Gamma \sim 1/\sigma_{xx}$, assuming that near the percolation level $\sigma_{xx}^{ac} = \sigma_{xx}^{dc}$. Then

$$\Gamma(H)/\Gamma(H_{\max}) = \Sigma(H)\sigma_{xx}^{dc}(H_{\max})/\sigma_{xx}^{dc}(H), \tag{11}$$

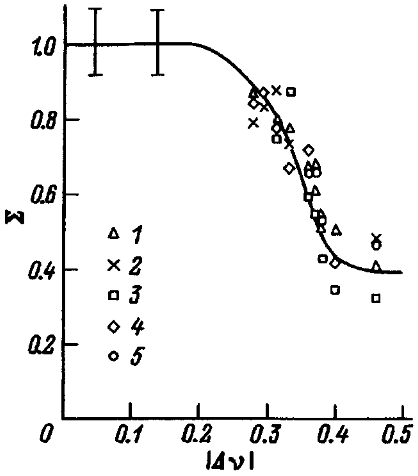

FIG. 5. Σ versus $|\Delta\nu|$— absolute value of the change in the filling factor ν, measured from $\nu=3.5$. The symbols correspond to frequencies f (MHz) and vacuum gap widths a (in μm): *1*—213 and 0.3, *2*—30 and 0.5, *3*—150 and 0.3, *4*—30 and 0.4, *5*—90 and 1.2.

since $\Gamma(H_{\max})$ is determined according to Eq. (4). As indicated above, the ratio $\sigma_{xx}^{dc}(H_{\max})/\sigma_{xx}^{dc}(H)$ is calculated from the components $\rho_{xx}(H)$ and $\rho_{xy}(H)$ of the resistivity tensor which are measured for this sample.

Another method to obtain Σ was to determine it from the quantity Γ_M (see Fig. 2). This method did not require a knowledge of σ_{xx}, since we obtain from Eqs. (4), (7), and (10) the expression

$$\Gamma=\Sigma(H)17.36K^2k\ \exp(-2ka)(\varepsilon_1+\varepsilon_0)\varepsilon_s\varepsilon_0^2b(k,a). \tag{12}$$

The dependence of the function Σ, which is determined by the methods described above, on $|\Delta\nu|$ at $T=4.2$ K in the magnetic field range corresponding to filling numbers ν from 3 to 4, where $\Delta\nu$ is measured from $\nu=3.5$, is shown in Fig. 5. It should be noted that the points on the plot were obtained from the values of the absorption coefficient for different SAW frequencies and different values of a. The value of Σ depends on the temperature and is a subject of our further investigations. The solid line in Fig. 5 was estimated by eye and extrapolated to the region $\Delta\nu=0.5$, so that it was impossible to calculate the value of Σ near the conductivity minima because of the lack of accurate data on the value of the static conductivity in the limit $\sigma_{xx}^{dc}\rightarrow 0$.

Knowing the vacuum gap and the function $\Sigma(H)$, the value of $\Sigma(H)$ could be determined according to Eqs. (4) and (10) from the experimental values of Γ. The results of such calculations are presented in Fig. 6 for $T=4.2$ K. The different points correspond to curves obtained at different frequencies and with different vacuum gaps. The solid line represents the function $\sigma_{xx}^{dc}(H)$. As one can see from Fig. 6, for large Landau numbers $\sigma_{xx}^{ac}(H)=\sigma_{xx}^{dc}(H)$, just as in the weaker magnetic fields, but only near the conductivity maxima. In the regions of the magnetic field where $\sigma_{xx}^{dc}\rightarrow 0$, however, the values of σ_{xx}^{dc} and σ_{xx}^{ac} diverge, and σ_{xx}^{ac} is always greater than σ_{xx}^{dc}, as should be the case on the basis of the model proposed above. The equality $\sigma_{xx}^{ac}(H)=\sigma_{xx}^{dc}(H)$ for high Landau-level numbers is attributable to the fact that the mobility thresholds at $T\simeq 4.2$ K for neighboring levels are smeared in energy and overlap, so that the condition for carrier localization is not achieved in weak

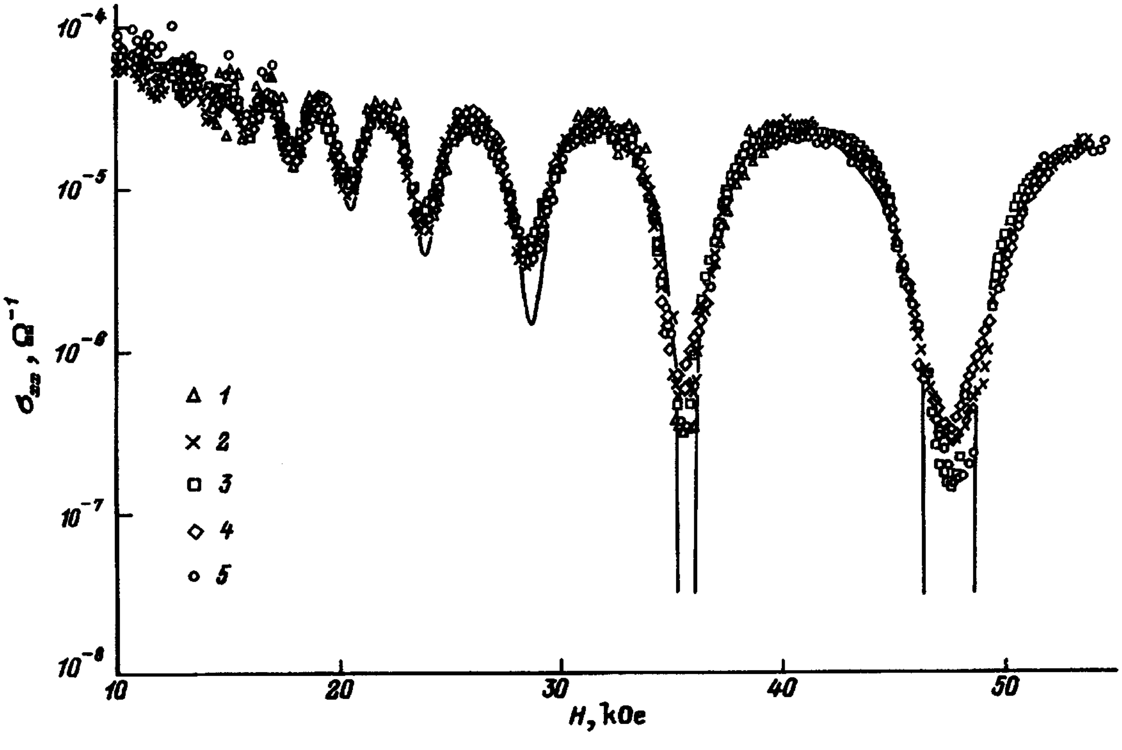

FIG. 6. σ_{xx}^{dc} (solid line) and σ_{xx}^{ac} (symbols) versus the magnetic field H. The different symbols have the same meaning as in Fig. 5.

magnetic fields, even when the chemical potential falls between the Landau levels.

Since σ_{xx}^{ac} is calculated from Γ after a rather complicated procedure of analysis of the experimental data, we estimate the error in σ_{xx}^{ac} to be of the order of 40%. To this accuracy, the quantity σ_{xx}^{ac} at the minima with respect to the magnetic field for small filling numbers, i.e., in localized states, does not depend on the sound frequency (and in our experiments the frequency varied by a factor of 7).

It should be noted that our results are at variance with the data of Ref. 12, where σ_{xx}^{ac} was studied in the Corbino disk geometry in the frequency range 50–600 MHz. The authors found that σ_{xx}^{ac} does not equal to σ_{xx}^{dc} either at the maxima or at the minima of the oscillations of $\sigma_{xx}(H)$ when the filling numbers ν are small. They explained the observed frequency dependence of σ_{xx} in terms of the theory given in Ref. 13.

5. CONCLUSIONS

The magnetic field and frequency dependences of the high-frequency conductivity of the 2DEG in the heterostructure GaAs/$Al_{0.25}Ga_{0.75}As$ (in the region 30–210 MHz) were calculated and analyzed. It was shown that the experimental results can be explained if it is assumed that there exists a fluctuation potential in which carrier localization occurs. The character of the SAW absorption in the presence of an interaction with localized carriers is discussed.

We thank Yu. M. Gal'perin, V. D. Kagan, and A. Ya. Shik for helpful discussions and G. O. Andrianov for assisting in the work.

This work was supported by the Russian Fund for Fundamental Research Grants Nos. 95-02-0466a and 95-02-04042a as well as INTAS Grants Nos. 93-1403, 93-1403-EXT, and 95-IN/RU-553.

---

1) e-mail: pta@nano.ioffe.rssi.ru

2) e-mail: sus@frost.shuv.ioffe.rssi.ru

[1] D. G. Polyakov and B. I. Shklovskii, Phys. Rev. B **48**, 11 167 (1993).

[2] Yu. M. Gal'perin, E. M. Gershenzon, I. L. Drichko, and L. B. Litvak-Gorskaya, Fiz. Tekh. Poluprovodn. **24**, 3 (1990) [Sov. Phys. Semicond. **24**, 1 (1990)].

[3] Yu. M. Gal'perin, I. L. Drichko, and L. B. Litvak-Gorskaya, Fiz. Tverd. Tela (Leningrad) **30**, 3118 (1988) [Sov. Phys. Solid State **30**, 1792 (1988)]; B. A. Aronzon and I. L. Drichko, Fiz. Tekh. Poluprovodn. **26**, 1446 (1992) [Sov. Phys. Semicond. **26**, 811 (1992)].

[4] A. Schenstrom, Y. J. Quan, M. F. Xu, H. P. Baum, M. Levy, and B. K. Sarma, Solid State Commun. **65**, 739 (1988).

[5] A. Wixforth, J. P. Kotthaus, and G. Weimann, Phys. Rev. Lett. **56**, 2104 (1986); A. Wixforth, J. Scriba, M. Wassermeier, J. P. Kotthaus, G. Weiman, and W. Schlapp, Phys. Rev. **40**, 7874 (1989).

[6] V. W. Rampton, K. McEnaney, A. G. Kozorezov, P. J. A. Carter, C. D. Wilkinson, M. Henin, and O. H. Hughes, Semicond, Sci. Techn. **7**, 647 (1992).

[7] R. Boulet, P. Coleridje, F. Guillon, M. D'Iorio, and A. Sachrajda, Can. J. Phys. **69**, 461 (1991).

[8] R. L. Willett, M. A. Paalanen, R. R. Ruel, K. W. West, L. N. Pfeiffer, and D. J. Bishop, Phys. Rev. Lett. **65**, 112 (1990).

[9] I. L. Drichko, A. M. D'yakonov, V. D. Kagan, A. M. Kreshchuk, G. D. Kipshidze, T. A. Polyanskaya, I. G. Savel'ev, I. Yu. Smirnov, A. V. Suslov, and A. Ya. Shik, Fiz. Tekh. Poluprovodn. **29**, 1306 (1995) [Semiconductors **29**, 677 (1995)].

[10] V. D. Kagan, Fiz. Tekh. Poluprovodn. **31**, 478 (1997) [Semiconductors **31**, 407 (1997)].

[11] E. I. Rashba and V. B. Timofeev, Fiz. Tekh. Poluprovodn. **20**, 977 (1986) [Sov. Phys. Semicond. **20**, 617 (1986)].

[12] I. E. Batov, A. V. Polisski, M. L. Reznikov, and V. I. Tal'yanskii, Solid State Commun. **76**, 25 (1990).

[13] S. Yu. Apenko and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **89**, 573 (1985) [Sov. Phys. JETP **62**, 328 (1985)].

Translated by M. E. Alferieff

Spin relaxation and weak localization of two-dimensional electrons in asymmetric quantum wells

A. M. Kreshchuk, S. V. Novikov, T. A. Polyanskaya, and I. G. Savel'ev[1)]

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

(Submitted August 13, 1996; accepted for publication September 10, 1996)
Fiz. Tekh. Poluprovodn. **31**, 459–467 (April 1997)

The anomalous alternating-sign magnetoresistance in a two-dimensional electron gas on an $In_{0.53}Ga_{0.47}As/InP$ heterostructure was investigated experimentally at liquid-helium temperatures in a wide range of electron densities, including the case of two filled quantum-well subbands. The data obtained are analyzed in terms of a theory that takes into account terms in the spin splitting of the electron spectrum which are cubic and linear in the wave vector. The linear term is related to the asymmetry of the quantum well, i.e., the presence of an electric field at the heterojunction. It is shown that the new theoretical model describes the experiment better. © *1997 American Institute of Physics.* [S1063-7826(97)01604-9]

1. INTRODUCTION

The theory of weak electron localization due to the interference of the electron wave functions as the electrons diffuse along a closed trajectory was developed in the beginning of the 1980s (Refs. 1–3). This theory explained the anomalous behavior of the negative magnetoresistance. This effect has long been observed in substances with metal-type conductivity in weak magnetic fields at low temperatures, when the classical (Lorentzian) magnetoresistance equals zero (when the condition $kT \ll \varepsilon_F$, is satisfied; here ε_F is the Fermi energy, $\omega_{c\tau} \ll 1$, ω_c is the cyclotron frequency, and τ is the momentum relaxation time). It was also shown then[4,5] that because of the different behavior of triplet and singlet terms of the interference wave function of the electrons, the specific experimental manifestation of this effect depends on the ratio of the characteristic times: the relaxation time τ_ϕ of the phase of the electron wave function as a result of inelastic or quasielastic collisions and the spin relaxation time due to the spin-orbit interaction. If the spin relaxation is very rapid $\tau_s \ll \tau_\phi$, then the anomalous magnetoresistance becomes positive, as is observed in experiments on degenerate *p*-type semiconductors, where the spin relaxation rate is nearly equal to the elastic scattering rate and is much greater than $1/\tau_\phi$. In the case $\tau_s \ll \tau_\phi$ the theory predicts an alternating-sign magnetoresistance (MR), i.e., in a weak magnetic field B the singlet term predominates and a positive MR is observed; as B increases, the MR becomes negative, since the contribution of the singlet term to the MR saturates and the triplet term comes into play.

The dimension of the conductor from the standpoint of the weak localization effect is determined by the ratio of the dimensions of the conductor and the electron diffusion length $l_\phi = \sqrt{D\tau_\phi}$ during the loss of phase coherence of the electron wave function (D is the diffusion coefficient). If the film thickness is less than l_ϕ but the carrier density of states is three-dimensional, then the conductivity of the film is said to be quasi-two-dimensional. In contrast to quasi-two-dimensional structures, we shall call structures two-dimensional if the thickness of the conducting layer is comparable to the electron wavelength and in which size-quantization of the energy spectrum is observed.

The appearance of spin-orbit scattering in weak localization effects in quasi-two-dimensional conductors has been studied in detail experimentally for the example of metal films[6] with thickness less than l_ϕ. It was shown that the main spin scattering mechanism in these films is scattering by impurity nuclei (Elliot–Yafet) (EY) effect.[7] However, it follows from the theory of Ref. 4 that for this mechanism scattering in a direction normal to the conducting layer is important. Therefore, the EY mechanism should not appear in two-dimensional structures with one filled quantum-well subband. Size quantization is never observed in metal films because of the very short electron wavelength and it is quite easily achieved in semiconductor heterostructures with quantum wells. Indeed, the alternating-sign magnetoresistance in heterostructures with a two-dimensional electron gas (2DEG) was observed for a long time, only in structures with two[8] or a large number of filled quantum-well subbands,[9] where scattering in a direction normal to the surface is possible. However, experimental studies where the magnetoresistance in very weak magnetic fields was investigated in GaAs-based heterostructures with 2DEG and an alternating-sign MR was observed in structures with one filled quantum-well subband have appeared in recent years.[10,11] The authors interpreted this behavior as a manifestation of a spin scattering mechanism associated with spin splitting of the electronic spectrum because of the absence of an inversion center in the crystals—the D'yakonov–Perel' mechanism (DP mechanism). In Ref. 10 a volume Hamiltonian was used for the DP mechanism that took into account only the term proportional to the cube of the electron wave vector. The theory of Ref. 12, however, attests to the fact that in the two-dimensional case, in addition to the cubic term, it is also necessary to include terms which are linear in the wave vector and are related to either the absence of an inversion center of the crystal (Dresselhaus mechanism)[13] or the asymmetry of the quantum well, i.e., the presence of an electric field at the heteroboundary (Rashba mechanism[14]). The contribution of these terms to the spin relaxation time of two-dimensional electrons was taken into account in Ref. 12 in an analysis of spin-orbit scattering. However, a further theoretical analysis[15] showed that the contribution of these three terms is not additive and that the expression obtained in Ref. 4 for the magnetoresistance must be modified in order to make the linear terms comparable to the cubic terms. At present, there exists only one experimental study[16] where the experiment is analyzed in terms of this theory and it is shown that the linear terms play a large role in AlGaAs/$In_{0.15}Ga_{0.85}As$/GaAs heterostructures in the density range $(1-1.8)\times 10^{12}$ cm^{-2}.

In the present paper we investigate the alternating-sign magnetoresistance in a 2DEG in a $In_{0.53}Ga_{0.47}As$ layer at the heteroboundary in selectively doped $In_{0.53}Ga_{0.47}As$/InP and InAlAs/$In_{0.53}Ga_{0.47}As$/InP heterostructures. Preliminary investigations of these structures in strong magnetic fields[17] have shown that different spin effects are more pronounced in them than in GaAs-based heterostructures. This made it possible to investigate the effect of an alternating-sign magnetoresistance and spin-orbit scattering in a wider range of two-dimensional carrier densities n_s from 1.8×10^{11} to 2×10^{12} cm^{-2} and to determine the effect of the filling of a second quantum-well subband on the spin-orbit scattering of two-dimensional carriers.

2. THEORY OF ANOMALOUS MAGNETORESISTANCE IN THE PRESENCE OF SPIN-ORBIT SCATTERING

Negative magnetoresistance in weak magnetic fields is due to the suppression of the interference of the wave functions of electrons moving on closed trajectories in opposite directions or, in other words, the suppression of weak carrier localization, i.e., the action of a magnetic field reduces to destroying the coherence of the electron wave functions and is observed when the magnetic length $l_H=(hc/eB)^{1/2}$ becomes less than the spatial coherence length l_ϕ and, in the case of scattering with spin flip, less than the spin coherence length $l_s=\sqrt{D\tau_s}$. An expression for the quantum corrections to the conductivity in the two-dimensional case was obtained in Ref. 4 and can be represented in the form

$$\frac{\Delta\sigma(B)}{G_0}=-\Psi\left(\frac{1}{2}+\frac{H_{tr}}{B}\right)+\Psi-\left[\left(\frac{1}{2}+\frac{H_\phi}{B}+\frac{H_s}{B}\right)\frac{1}{2}\right.$$

$$\left.-\Psi\left(\frac{1}{2}+\frac{H_\phi}{B}\right)-\Psi\left(\frac{1}{2}+\frac{H_\phi}{B}+\frac{2H_s}{B}\right)\right], \tag{1}$$

$$\frac{\Delta\sigma(0)}{G_0}=\ln\left(\frac{H_\phi+H_s}{H_{tr}}\right)+\frac{1}{2}\ln\left(\frac{H_\phi+2H_s}{H_\phi}\right), \tag{1a}$$

where $\Psi(x)$ is the digamma function, B is the magnetic field, and

$$G_0=\frac{e^2}{2\pi^2\hbar},\quad \frac{H_{tr}}{B}=\frac{l_H^2}{4D\tau},\quad \frac{H_\phi}{B}=\frac{l_H^2}{4l_\phi^2},\quad \frac{H_s}{B}=\frac{l_H^2}{4l_s^2}.$$

As shown in Ref. 15, this expression is valid if the Hamiltonian for the spin splitting of the conduction band contains only a term that is cubic in the component of the wave vector k in the plane of the heterojunction

$$\Omega_3=\gamma\frac{k^3}{4}. \tag{2}$$

However, in a III–V based quantum well, size quantization leads to the appearance of a term that is linear in the wave vector (Dresselhaus mechanism)[13]

$$\Omega_1=\gamma k\left(\langle k_z^2\rangle-\frac{1}{4}k^2\right), \tag{3}$$

where $\langle k_z^2\rangle$ is the mean-square value of the component of the wave vector in a direction perpendicular to the plane. Furthermore, in an asymmetric quantum well an additional term, which was proposed by Rashba[14] and which is linear in the electron wave vector

$$\Omega_{1R}=\alpha Fk \tag{4}$$

(F is the electric field at the heterojunction) appears in the Hamiltonian.[14] The coefficients α and γ are constants characterizing the energy spectrum of a specific semiconductor. These three terms appear additively in the expression for calculating the spin relaxation rate

$$\frac{1}{\tau_s}=2(\Omega_1^2\tau_1+\Omega_{1R}^2\tau_1+\Omega_3^2\tau_3), \tag{5}$$

where

$$\frac{1}{\tau_n}=\int W(\theta)(1-\cos n\theta)d\theta,\quad n=1,3.$$

Here $\tau_1=\tau$ is the transport electron momentum relaxation time, and $W(\theta)$ is the probability of scattering by an angle θ.

However, if the spin-splitting Hamiltonian contains terms which are linear in k, it is necessary to take into account the correlation between the motions of electrons in the coordinate and spin spaces. In this case, as shown in Ref. 15, the contribution of the terms which are linear and cubic in the wave vector k is not additive. Additionally, it was demonstrated in Ref. 12 that the contributions of the Rashba and Dresselhaus terms are also not additive, and the magnetoconductivity is determined not by their sum but rather by their difference. An analytic expression for the magnetoconductivity in the case where the cubic term is comparable to one of the linear terms and the linear terms differ markedly in magnitude was obtained in Ref. 15 in the form

$$\frac{\Delta\sigma(B)}{G_0}=-\frac{1}{a_0}-\frac{2a_0+1+H_s/B}{a_1(a_0+H_s/B)-2H_{s1}/B}+\sum_{n=0}^{\infty}$$

$$\times\left(\frac{3}{n}-\frac{3a_n^2+2a_nH_s/B-1-2(2n+1)H_{s1}/B}{(a_n+H_s/B)a_{n-1}a_{n+1}-2[(2n+1)a_n-1]H_{s1}/B}\right)$$

$$-2\ln\frac{H_{tr}}{B}-\Psi\left(\frac{1}{2}+\frac{H_\phi}{B}\right)-3C, \tag{6}$$

where $a_n=n+1/2+\ H_\phi/B+H_s/B$, and C is Euler's constant. Expression (6), in contrast to expression (1), contains two characteristic magnetic fields, which are important for describing spin-orbit scattering. In addition to H_s, which is determined by the total spin relaxation time τ_s [Eq. (5)]

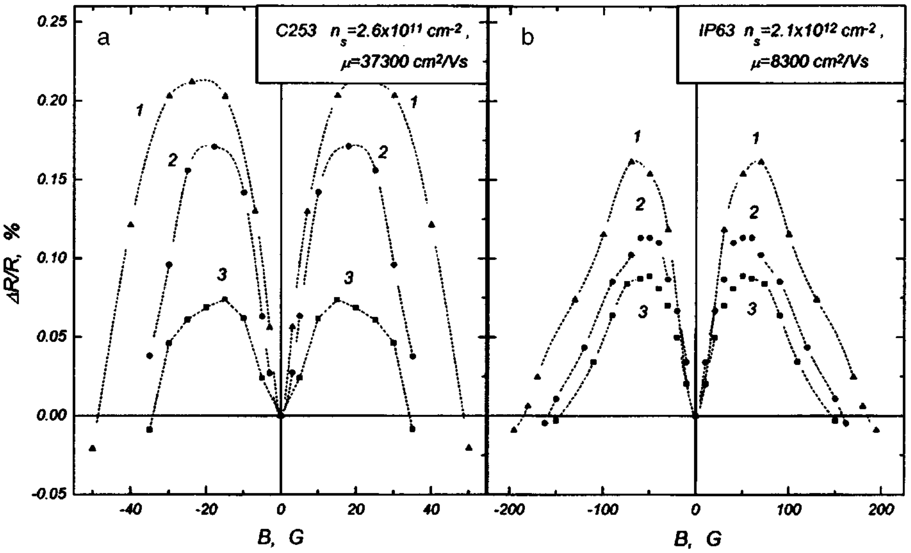

FIG. 1. Magnetoresistance $\Delta R/R$ versus magnetic field B for InP/InGaAs structures—sample S253 (a) and InAlAs/InGaAs structures—sample IP63 (b), measured at temperature T, K: *1*—1.86, *2*—2.5, and *3*—4.2.

$$H_s=\frac{2}{4\hbar eD}(\Omega_1^2\tau_1+\Omega_{1R}^2\tau_1+\Omega_3^2\tau_3), \qquad (7)$$

we have an additional parameter H_{s1}, which is determined by the largest term linear in the wave vector

$$H_{s1}=\frac{2\tau_1\max\{\Omega_1^2,\Omega_{1R}^2\}}{4\hbar eD}. \qquad (8)$$

3. SAMPLES AND EXPERIMENTAL PROCEDURE

Selectively doped heterostructures, in which the 2DEG was located in a square potential well in a layer of the solid solution $In_{0.53}Ga_{0.47}As$ isoperiodic to InP, were used as the object for investigating the spin-orbit scattering in effects due to weak localization of two-dimensional carriers. Heterostructures in which the layers of InP and $In_{0.52}Al_{0.48}As$ were the electron source and at the same time a barrier layer were used to increase the range of the parameter of the 2DEG.

Selectively doped $InP/In_{0.53}Ga_{0.47}As$ heterostructures were grown by liquid-phase epitaxy on semi-insulating InP (100) substrates. The structures consisted of successive layers of a p-type InP buffer layer with hole density $p<10^{15}$ cm^{-3}and thickness $d=1$ μm, a InP layer as a source of electrons with donor density from 10^{16} to 2×10^{17} cm^{-3}, and an $In_{0.53}Ga_{0.47}As$ layer containing the 2DEG and having parameters $p<10^{15}$ cm^{-3} and $d=0.3$ μm.[18] In the process of establishing thermal equilibrium in such a structure, electrons from donors in the InP layer are transferred into the narrow-gap InGaAs layer. Here they are partially captured by acceptors in the layer, recharge the defects on the surface of the InGaAs layer, and the remaining electrons end up in the potential well in the InGaAs layer at the heteroboundary and form a 2DEG. Heterostructures with a wide range of two-dimensional carrier densities from 1.8×10^{11} to 7×10^{11} cm^{-2} and similiar mobilities in the range from 2.5×8^4 to 3.5×10^4 cm^2/(V·s), which attests to the very small contribution of scattering by the long-range potential of ionized impurities to the electron momentum relaxation time in the structures, were selected for the investigations. As preliminary investigations showed,[19] the main mechanisms of elastic scattering in these structures are scattering by the alloy potential and by the irregularities of the heterostructure.

$InP/In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As$ heterostructures were grown by molecular-beam epitaxy. Here the electron source consisted of a doped $In_{0.52}Al_{0.48}As$ layer. The parameters of the structures were changed by illuminating the sample gradually by pulses from a GaAs light-emitting diode. The measurements were performed with direct current on samples in the form of a double Hall cross. The contacts to the samples were established by brazing indium droplets in vacuum at a temperature of 400 °C, which ensured that the contacts would be ohmic.

4. ANALYSIS OF THE MAGNETIC FIELD DEPENDENCE OF THE MAGNETORESISTANCE

The experimental magnetic field dependences of the magnetic resistance MR $\Delta R/R=[R(B)-R]/R$ (where $R(B)$ and R are the resistance, normalized to a square, of a sample in the magnetic field and without a field, respectively) in weak magnetic fields are shown in Fig. 1 for samples with 2DEG density $n_s=2.6\times10^{11}$ cm^{-2} (a) and 2.1×10^{12} cm^{-2} (b) and in Fig. 2 for a sample with $n_s=1.8\times10^{11}$ cm^{-2}.

We see that the alternating-sign character of the magnetoresistance is observed in a wide range of 2DEG densities. For a structure with a low-density 2DEG, negative magnetoresistance is observed in the entire experimental range of magnetic fields (Fig. 2, curve *1*). However, even for this structure, the magnetic field dependence of the derivative $\partial(\Delta R/R)/\partial(B^2)$ of the magnetoresistance with respect to the square magnetic field is a curve with a minimum (Fig. 2, curve *2*), which also attests to an appreciable contribution of spin scattering for this sample. The magnetoconductivity was determined from experiment on the basis of an expression that is valid in weak magnetic fields, where the components of the resistivity tensor satisfy the relation $\rho_{xy}\ll\rho_{xx}$,

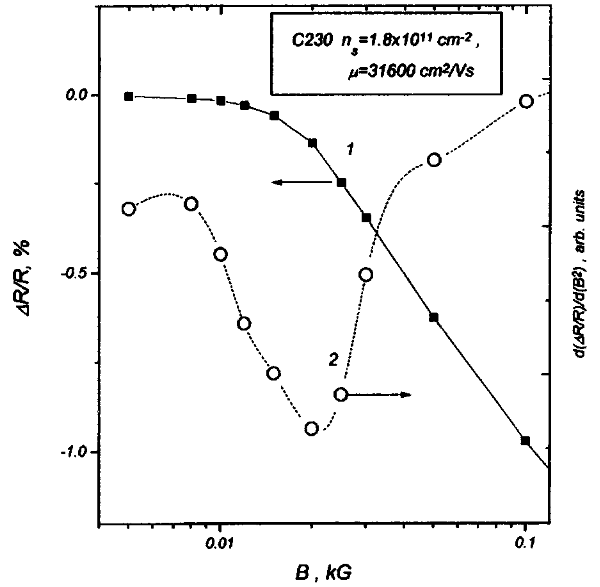

FIG. 2. Magnetoresistance $\Delta R/R$ versus magnetic field B for a InP/InGaAs structure—sample C230 (*1*) together with the derivative of this function with respect to the squared magnetic field (*2*) at $T=1.86$ K.

$$\frac{\Delta\sigma(B)}{G_0}=-\frac{[R(B)-R]}{R(B)RG_0}.$$

It is presented in the form of magnetic field dependences in Fig. 3. An initial analysis of the experimental dependences $\Delta\sigma(B)/G_0$ was made in terms of the theory of Ref. 4, i.e., expression (1). It follows from this analysis that the dimensionless magnetoconductivity at the minimum is determined by the ratio of the characteristic magnetic fields H_ϕ and H_s. Therefore, describing the initial section of the experimental magnetic field dependence of the magnetoconductivity with the help of expression (1) (Fig. 3, dashed curves) and making the experimental and computed values of $\Delta\sigma/G_0$ coincide at the minimum point, we determined the adjustable parameters H_ϕ and H_s which characterize the phase interruption time τ_ϕ of the wave function and the spin-orbit scattering time τ_s. In the entire experimental range of the parameters of the 2DEG, it was found that τ_s is temperature-independent, and τ_ϕ is inversely proportional to the temperature (see Fig. 4, where the typical temperatures dependences of the characteristic times are presented). Such a temperature dependence of τ_ϕ corresponds to the theoretical dependence predicted for electron-electron scattering with small momentum transfer, characteristic of slightly disordered 2DEG in a so-called ''dirty'' conductor.^{20} The experimental values of the coefficient $K_\phi=h/\tau_\phi T$ for structures with one filled subband are plotted in the inset in Fig. 4 as a function of the sample resistance R. The theoretical curve calculated on the basis of the expression

$$K_\phi=\frac{\hbar}{\tau_\phi T}=2\pi G_0 R\ln\left(\frac{1}{2\pi G_0 R}\right) \tag{9}$$

for electron-electron scattering in a dirty two-dimensional semiconductor^{20} is also presented there.$^{2)}$ We see that expression (9) describes well, both qualitatively and quantitatively, the dependence of τ_ϕ on the temperature and the parameters of the 2DEG when only one quantum-well subband is filled, i.e., it can be asserted that at low temperature the main mechanism for the relaxation of the phase of the wave function of two-dimensional electrons at the InP/InGaAs heterojunction is electron-electron scattering of the type (9), just as in most two-dimensional conductors. The phenomena associated with filling of the second subband will be discussed in Sec. 6.

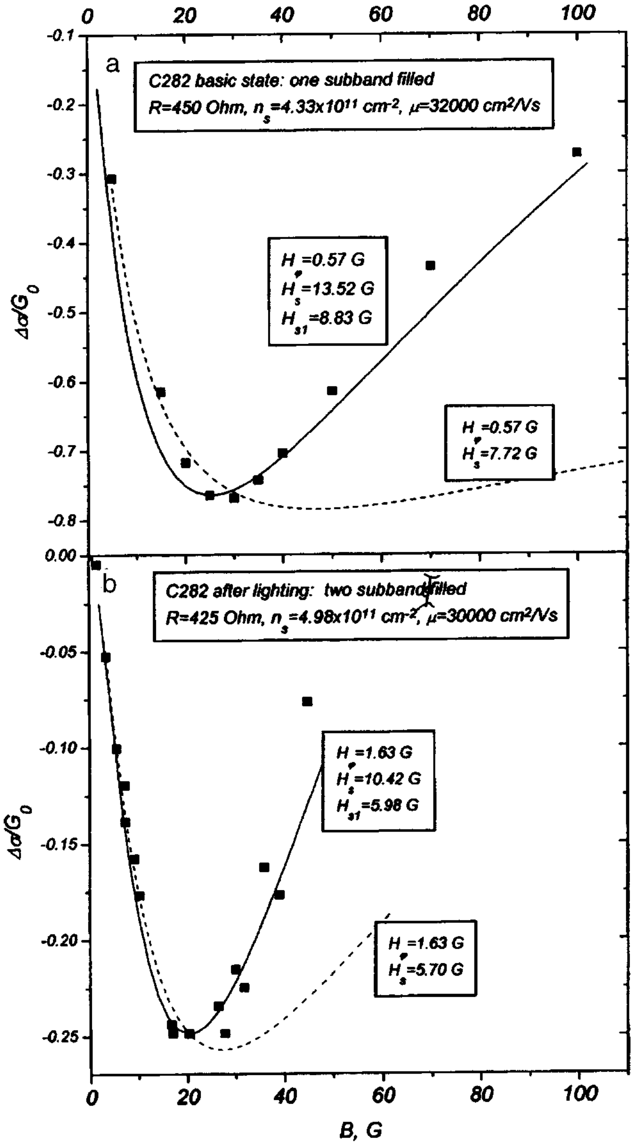

FIG. 3. Comparison of the experimental dependences of the normalized magnetoconductivity $\Delta\sigma/G_0$ versus magnetic field B (symbols) with the computed curves obtained on the basis of expression (1) (dashed curves) and expression (6) (solid curves). The data are presented for two states of sample C282: a—Initial state (one quantum-well subband is filled), b—in the frozen photoconductivity regime (two subbands are filled with carrier density 4.45×10^{11} cm^{-2} and 4.6×10^{11} cm^{-2} in the first and second subband, respectively).

To analyze the parameters characterizing the spin-orbit scattering, we take into account the fact that expression (1) was obtained for the case of a simplified Hamiltonian describing spin-orbit (SO) scattering. In addition, it is evident from Fig. 3 that the curve calculated on the basis of this

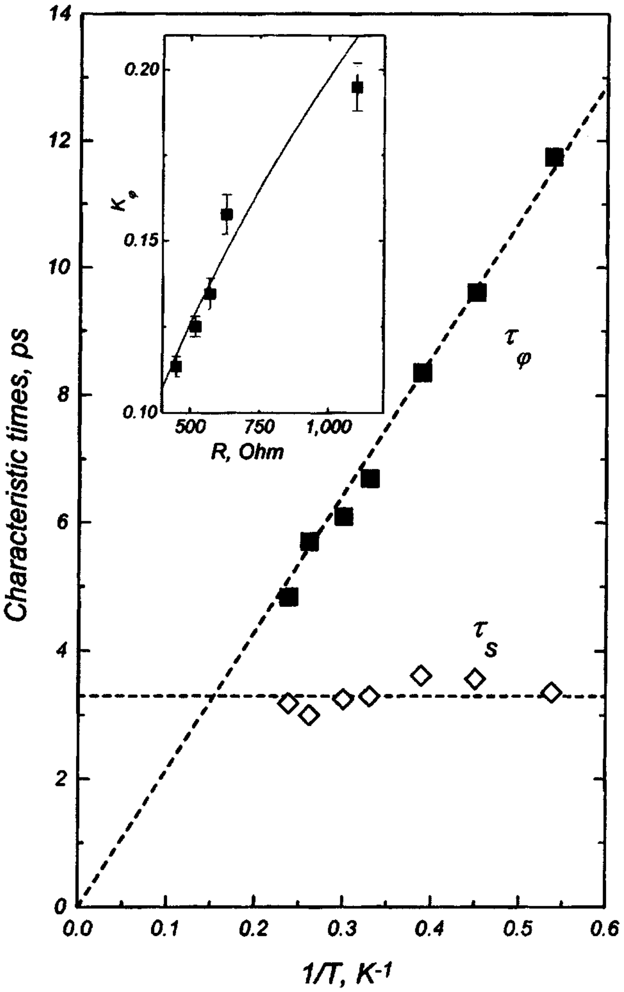

FIG. 4. Times τ_ϕ and τ_s, characterizing the relaxation of phase and spin, respectively, determined from analysis of the experimental magnetic-field dependences of the magnetoconductance on the basis of the expression (1) for sample C282 versus the reciprocal of the temperature. Inset: $K_\phi \equiv \hbar/\tau_\phi T$ versus the sample resistance R with one filled subband; experimental data—symbols, calculation based on the expression (9)—solid line.

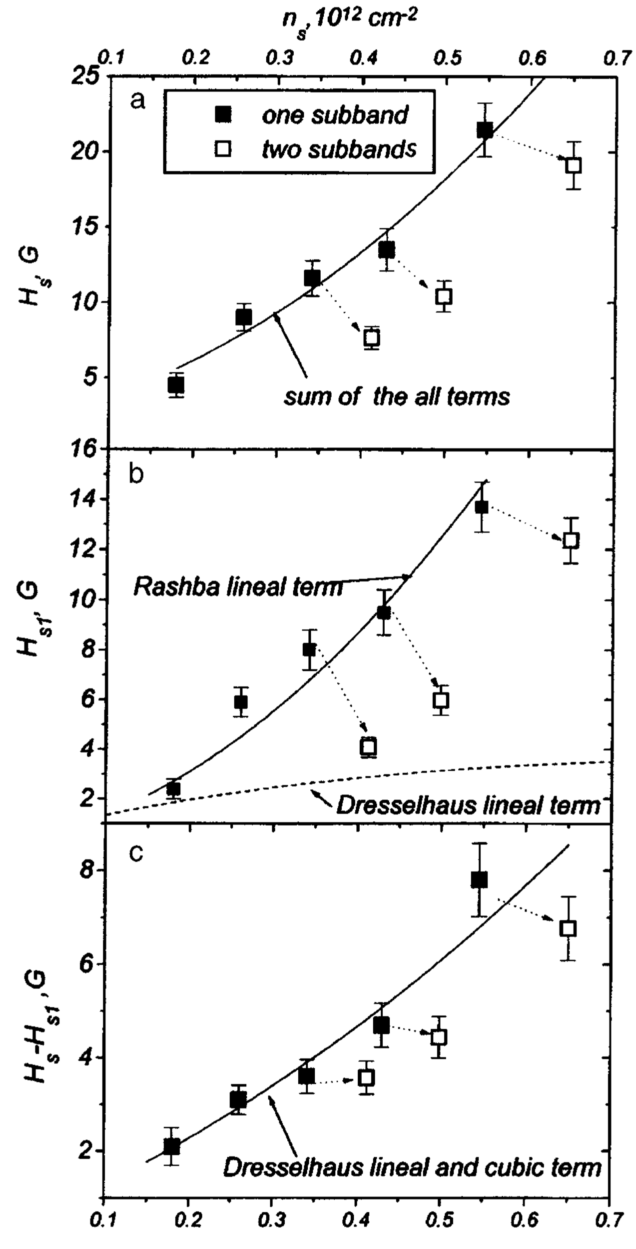

FIG. 5. Comparison of the experimental (symbols) and computed (curves) density dependences of the magnetic fields characterizing spin relaxation. The filled symbols correspond to structures with one filled subband and the open symbols correspond to structures with two filled subbands.

expression (Fig. 3, dashed curve) describes quite well the experimental magnetic field dependence of the magnetoconductivity (the section after the minimum). All this indicates that a more complicated expression (6), which takes into account the Hamiltonian for SO scattering terms which are cubic and linear in the wave vector, must be used. We accordingly introduce the additional adjustable parameter H_{s1}, which is associated with the largest term which is linear in the wave vector (8). There is only one published study[16] where the magnetoresistance of a 2DEG is analyzed on the basis of the simplified (1) and complete (6) theories of the suppression of weak localization by a magnetic field in the presence of SO scattering. It was found that the values of H_ϕ obtained in describing the experimental dependences with the help of expressions (1) and (6) are virtually identical. The good agreement of the quantities τ_ϕ (see Fig. 4), which we found as adjustable parameters in comparing the experimental curves with the theoretical relation (1), attests to the validity of the values of H_ϕ obtained by the method described above. This makes it possible to use as adjustable parameters in describing our data with the help of expression (6) only the quantities H_s and H_{s1} which characterize SO scattering and to take for the quantities H_ϕ the values found previously from expression (1). Therefore, as one can see from the data in Fig. 3, virtually the entire experimental curve $\Delta\sigma(B)/G_0$ with a maximum can be described by continuous curves.

5. ANALYSIS OF THE PARAMETERS CHARACTERIZING SPIN-ORBIT SCATTERING OF TWO-DIMENSIONAL ELECTRONS

Figure 5 shows the concentration dependences of the characteristic magnetic fields H_s (Fig. 5a) and H_{s1} (Fig. 5b), which describe the SO scattering and are found as adjustable parameters in describing the experimental curves $\Delta\sigma(B)/G_0$ with the aid of the help of theoretical relation (6) for all experimental samples. The filled symbols in Fig. 5 correspond to samples with one filled subband and the open symbols correspond to samples with two filled subbands. In this section we analyze again only the data for samples with one filled subband.

Comparing the results presented in Figs. 5a and 5b, it is easy to see that the parameter H_s, which characterizes the total spin relaxation time, and the parameter H_{s1}, which is associated only with the term linear in the wave vector in the SO scattering Hamiltonian, grow rapidly with increasing 2DEG density. Furthermore, the values of H_{s1} are close in order of magnitude to H_s. This shows that the linear term must be taken into account in describing the magnetoconductivity. In accordance with expression (8), to analyze the concentration dependence of H_{s1}, it is necessary to compare the magnitudes of the two SO scattering terms which are linear in k and which are predicted by the Dresselhaus [expression (3)] and Rashba [expression (4)] theories. We accordingly take into account that in a degenerate electron gas electrons with the Fermi energy and wave vector $k_F=\sqrt{2\pi n_s}$ participate in conduction. Furthermore, we note that in the experimental structures the two-dimensional electrons occupy an asymmetric quantum well at one heterojunction. One wall of this potential well is formed by the offset of the conduction band at the heterojuntion and the second is formed by the built-in electric field, which is determined by the charge distribution in the structure, specifically, the two-dimensional-carrier density and the intensity N_0 of the charged residual impurities in the $\mathrm{In}_{0.53}\mathrm{Ga}_{0.47}\mathrm{As}$ layer:

$$F=4\pi e(n_s+N_0)/\chi, \tag{10}$$

where χ is the permittivity of the semiconductor. This makes it possible to estimate the average value $\langle k_z^2\rangle$ of the wave vector of the electrons moving in the direction perpendicular to the heterojunction by using the wave function proposed by Fang and Howard for a square potential well,[22]

$$\langle k_z^2\rangle=\frac{1}{4}\left[48\pi e^2 m\left(N_0+\frac{11}{32}n_s\right)\Big/\hbar^2\chi\right]^{2/3}, \tag{11}$$

where m is the effective mass of the two-dimensional electrons. As a result, we find for the linear term due to the absence of an inversion center of the crystal (the Dresselhaus mechanism)

$$H_{s1D}=1.132\times10^7\gamma^2\left(\frac{m}{m_0}\right)^2\left\{0.0233\times\left[\left(\frac{m}{m_0}\right)\frac{1}{\chi}\left(\frac{N_0+\frac{11}{32}n_s}{10^{12}}\right)\right]^{2/3}-1.57\times10^{-4}\frac{n_s}{10^{12}}\right\}^2. \tag{12}$$

Here and below the magnetic field H is measured in G, and the density is measured in cm^{-2}. For the Rashba mechanism, which is associated with the asymmetry of the quantum well, using Eq. (10), we obtain

$$H_{s1R}=3710\,\alpha^2\left(\frac{m}{m_0}\right)^2\left(\frac{1}{\chi}\right)^2\left(\frac{n_s+N_0}{10^{12}}\right)^2. \tag{13}$$

Analysis of these expressions shows that in the investigated range of 2DEG densities for typical values $n_0=5\times10^{10}$ cm^{-2} H_{s1D} varies much less than H_{s1R}. When n_s varies by an order of magnitude, H_{s1D} varies only by a factor of 2 and H_{s1R} increases almost by two orders of magnitude. One can see from Fig. 5b that the experimental values of H_{s1} increase rapidly with increasing 2DEG density. Therefore, it can be stated that the parameter H_{s1} in the experimental structures is determined by the Rashba mechanism, i.e., it is due to the asymmetry of the quantum well and, correspondingly, the condition $H_{s1D}<H_{s1R}$ holds. Describing the experimental dependence of H_{s1D} on n_s (the filled squares in Fig. 5b) with the help of expression[3)] (13) and using the quantity α (Fig. 5b, solid curve) as an adjustable parameter, we find $\alpha=(33\pm3)$ $(\text{Å})^{-2}$.

Let us now analyze the second parameter H_s, which is proportional to the total spin-orbit scattering rate. It follows from the theory that $H_s=H_{s1R}+H_{s1D}+H_{s3D}$, where the parameter H_{s3D} is determined by the term cubic in the wave vector in the Hamiltonian (2), which characterizes the SO scattering in a crystal without an inversion center. The expression for this parameter can be written in the form

$$H_{s3D}=0.283\gamma^2\frac{\tau_3}{\tau_1}\left(\frac{m}{m_0}\right)^2\left(\frac{n_s}{10^{12}}\right)^2. \tag{14}$$

Since in our case $H_{s1D}\ll H_{s1R}$, the parameters in H_s, which characterize the SO scattering mechanism associated with the incomplete symmetry of the crystal (the adjustable parameter γ) and with the asymmetry of the quantum well (adjustable parameter α), can be distinguished from one another. Indeed, the quantity H_s-H_{s1} in this case is determined only by the parameters associated with the absence of an inversion center in the crystal: $H_s-H_{s1}=H_{s3D}+H_{s1D}$. In Fig. 5c the symbols represent the values of H_s-H_{s1} obtained by subtracting the experimental data for H_{s1} (Fig. 5b) from the data for H_s (Fig. 5a). The computed curve in Fig. 5c was obtained by describing the experimental data with the aid of a sum of the theoretical expressions (12) and (14) using the coefficient γ as an adjustable parameter. In so doing, allowance was made for the fact that $\tau_3/\tau_1\approx1$ for the mechanisms of momentum dissipation, which predominate in the structures investigated, both by a short-range potential of the heterojunction irregularities and by the composition nonuniformities of the solid solution. The agreement in Fig. 5c between theory and experiment is obtained with the adjustable parameter $\gamma=160\pm10$ $\mathrm{eV}\cdot(\text{Å}^3)$. Knowing this parameter, it is possible to check additionally the assertion made above that $H_{s1D}<H_{s1R}$. We shall calculate the density dependence of H_{s1D} on the basis of expression (12) (Fig. 5b, dashed curve) and compare with the previously calculated dependence for H_{s1R} (Fig. 5b, solid curve). One can see that the condition $H_{s1D}<H_{s1R}$ is indeed satisfied in the entire

investigated range of 2DEG densities; this confirms our conclusion that the Rashba mechanism, which is associated with the asymmetry of the quantum well, plays a large role. The validity of our analysis is also confirmed by the good agreement between the experimental values of H_s (filled symbols in Fig. 5a) and the theoretical curve calculated as the sum of expressions (12), (13), and (14) with the values obtained above for the coefficients α and γ (solid curve).

The values of the coefficients α and γ are determined by the ratio of the different parameters of the band structure of the semiconductor. We were not able to find published estimates of these parameters for a $In_{0.53}Ga_{0.47}As$ layer isoperiodic to InP. However, theoretical estimates of these parameters for GaAs and InAs have been published.[23] Comparing these data with our results, it can be concluded that the coefficient characterizing the Rashba effect, $\alpha = 33$ $(\text{Å})^{-2}$, is in good agreement with the theoretical estimates; i.e., it is greater than the theory for GaAs predicts and less than for InAs. However, the value $\gamma = 160$ $\text{eV} \cdot (\text{Å})^3$ which characterizes the effects associated with the symmetry of the crystal, is much greater than the values predicted for a three-dimensional electron gas in GaAs and InAs. The reason for this large discrepancy remains unknown at present.

6. SPIN-ORBIT SCATTERING IN THE WEAK LOCALIZATION OF TWO-DIMENSIONAL ELECTRONS FILLING TWO QUANTUM-WELL SUBBANDS

To clarify the role of the second quantum-well subband, we investigated the alternating-sign magnetoresistance in samples in which two quantum-well subbands are filled in the initial state, specifically, in a heterostructure grown by molecular-beam epitaxy with 2DEG density $n_s = 2 \times 10^{12}$ cm^{-2}. Furthermore, Fig. 3 shows the curves of the magnetoconductivity for sample S282 in the initial state (Fig. 3a) and in the frozen photoconductivity state (Fig. 3b). Analysis of the Shubnikov–de Haas oscillations shows that in the initial state of the sample the two-dimensional carriers fill only the first quantum-well subband and illumination with interband light changes the state of the sample and results in the appearance of carriers in the second subband. Because of the frozen photoconductivity, at low temperature this state of the sample remains unchanged after the light is switched off during the experiment. Comparing the data in Figs. 3a and 3b, it can be concluded that the appearance of carriers in the second subband decreases the so-called ''autolocalization'' minimum in the magnetoconductivity. This attests to the fact that, in contrast to the results obtained in Ref. 8, the filling of the second subband decreases the influence of the spin-orbit interaction on the localization effect.

The experimental curves of the magnetoconductivity for structures with two filled subbands were analyzed in exactly the same manner using the same expressions as described previously for the case where only one subband is filled.[4] The result turned out to be the values of the characteristic magnetic fields H_ϕ, H_s, and H_{s1}, shown in Figs. 5 and 6 for three samples in which filling of the second subband is observed in the frozen photoconductivity regime. The arrows mark the light-induced transition from the initial state (filled symbols) to a frozen photoconductivity state (open symbols).

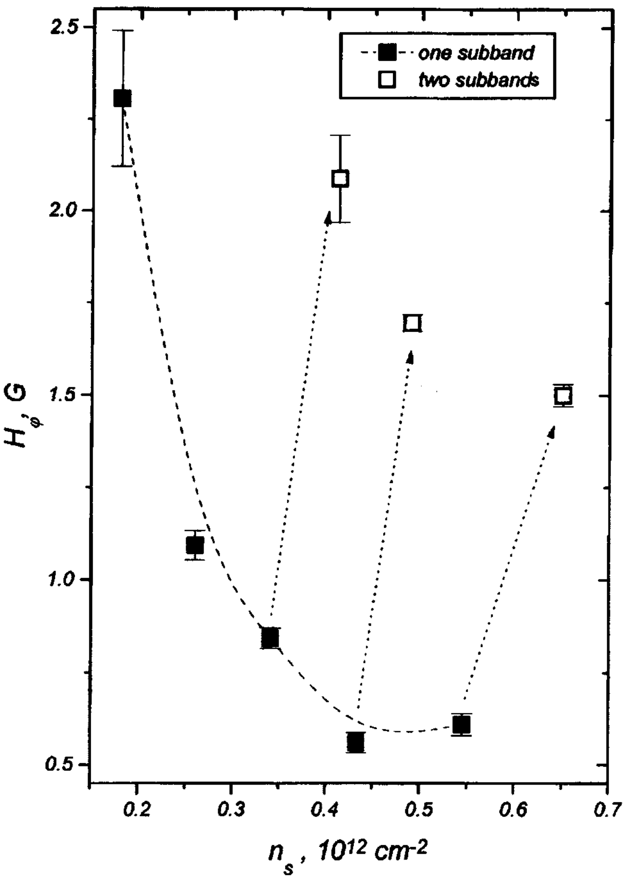

FIG. 6. Density dependence of the magnetic field characterizing the relaxation of the phase of the electron wave function for structures with one filled subband (filled symbols) and two filled subbands (open symbols).

It is easy to see that the appearance of carriers in the second subband results in a rapid increase in the value of H_ϕ (Fig. 6) and a decrease in the quantities characterizing the spin-orbit scattering, both H_s and H_{s1} (Figs. 5a, and 5b).

The increase in H_ϕ accompanying the filling of the second quantum-well subband was previously observed in Si-MIS structures[24] and explained by the need to take into account the increase in the density of states and, correspondingly, the decrease in the electron diffusion coefficient accompanying the filling of the second subband. The same conclusion was drawn in Ref. 25, where the problem of the magnetoconductivity associated with suppression of the weak localization of electrons that fill two subbands was solved theoretically. The question of the values of the characteristic magnetic fields for SO scattering of electrons in two subbands has still not been solved, either experimentally or theoretically.

If it is assumed, as done in Ref. 25, that when carriers appear in the second subband the characteristic relaxation times remain the same and only the diffusion coefficient changes, then the character of the variation of H_ϕ as the subbands are filled should be the same as that of H_s. One can see from Figs. 5 and 6 that in our case this is not so. We believe that it is necessary to take into account the difference in the diffusion coefficients of the electrons in the first and second subbands and also the characteristic phase and spin relaxation times. Just as in Ref. 25, we propose performing the averaging with a weight equal to the density of states in the first and second subbands, which is identical in the two-dimensional case. This approach reduces to a simple arithmetic averaging of H_ϕ and H_s, which are proportional to the relaxation rates of the phase and spin of electrons in the first and second quantum-well subbands. For H_ϕ we write

$$H_\phi=\frac{1}{2}\left(H_\phi^{(1)}+H_\phi^{(2)}\right). \tag{15}$$

Here and below the superscript in parentheses corresponds to the number of the subband. Let us see what the effect of such an averaging of the characteristic magnetic fields could be. Taking into account expression (9) and ignoring the difference in the properties of the carriers in the first and second subbands, we can write

$$H_\phi^{(i)}\propto\frac{1}{D^{(i)}\tau_\phi^{(i)}}\propto\left(n_s^{(i)}\right)^{-2}. \tag{16}$$

In the case of interest to us—low carrier density in the second subband $n_s^{(2)}\ll n_s^{(1)}$—we obtain from Eq. (16) $H_\phi^{(2)}\gg H_\phi^{(1)}$ and $H_\phi\cong H_\phi^{(2)}/2\gg H_\phi^{(1)}$. This means that despite the low density of carriers in the second subband, their contribution to phase relaxation can be large and, correspondingly, the appearance of a small number of carriers in the second subband should be manifested as a sharp increase in the characteristic magnetic field determined from analysis of the experiment, as shown in Fig. 6. The averaging represented by expression (15) is also valid for the magnetic fields H_s and H_{s1} characterizing the spin relaxation. In this case, however, as follows from expressions (13) and (14) and from the experimental curves (Fig. 5), the characteristic magnetic field increases rapidly with the carrier density $H_s^{(i)}\propto(n_s^{(i)})^2$. This means that for $n_s^{(2)}\ll n_s^{(1)}$ we obtain $H_s^{(2)}\ll H_s^{(1)}$ and $H_s\cong H_s^{(1)}/2$; i.e., the appearance of a small number of carriers in the second subband decreases the magnetic fields that characterize the spin-orbit relaxation by approximately a factor of 2. The same argument is also valid for H_{s1}. It is evident from the data presented in Fig. 5 that our experimental results correspond to the model described above.

We thank G. E. Pikus for helpful discussions.

This work was supported by the Russian Fund for Fundamental Research (Grant 95-02-04042a) and INTAS-RFFI (Grant 95-IN/RU-553).

---

1) Electronic mail: igs@nano.ioffe.rssi.ru

2) An empirical coefficient equal to 2 (Ref. 21) and initially absent in the theoretical expression (Ref. 20) is included in expression (9) (see the discussion for Fig. 15 in Ref. 21).

3) Here and below, the parameters $m/m_0 = 0.041$ and $\chi=14$ for the solid solution $In_{0.53}Ga_{0.47}As$ were used.

4) This approach is valid for the case of strong intersubband scattering of the carriers, which, as a rule, occurs in heterostructures with a two-dimensional electron gas.

[1] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[2] B. L. Al'tshuler and A. G. Aronov, Zh. Éksp. Teor. Fiz. **77**, 2028 (1979) [Sov. Phys. JETP **50**, 968 (1979)].
[3] B. L. Altshuler, A. G. Aronov, and P. A. Lee, Phys. Rev. Lett. **44**, 1288 (1980).
[4] S. Hikami, A. I. Larkin, and Y. Nagaoka, Progr. Theor. Phys. **63**, 707 (1980).
[5] B. L. Al'tshuler, A. G. Aronov, A. I. Larkin, and D. E. Khmel'nitskiĭ, Zh. Éksp. Teor. Fiz. **81**, 768 (1981) [Sov. Phys. JETP **54**, 411 (1981)].
[6] G. Bergmann, Phys. Rep. **107**, 1 (1984).
[7] R. J. Elliot, Phys. Rev. **96**, 266 (1954).
[8] J. E. Hansen, R. Taboryski, and P. E. Lindelof, Phys. Rev. B **47**, 16040 (1993).
[9] Zh. I. Alferov, A. T. Gorelenok, V. V. Mamutin, T. A. Polyanskaya, I. G. Savel'ev, and Yu. V. Shmartsev, Fiz. Tekh. Poluprovodn. **18**, 1999 (1984) [Sov. Phys. Semicond. **18**, 1247 (1984)].
[10] P. D. Dresselhaus, C. M. M. Papavassilion, R. G. Wheller, and R. N. Sacks, Phys. Rev. Lett. **68**, 106 (1992).
[11] W. Knap, C. Skierbiszewski, E. Litwin-Staszewska, F. Kobbi, A. Zduniak, J. L. Robert, G. E. Pikus, S. V. Iordanskii, V. Mosser, and K. Zekentes, Acta Phys. Polon. A **87**, 427 (1995).
[12] F. G. Pikus and G. E. Pikus, Phys. Rev. B **51**, 16928 (1995).
[13] G. Dresselhaus, Phys. Rev. **100**, 580 (1995).
[14] Yu. L. Bychkov and E. I. Rashba, J. Phys. C **17**, 6093 (1984).
[15] S. V. Iordanskiĭ, Yu. Ya. Lyanda-Geller, and G. E. Pikus, JETP Lett. **60**, 206 (1994).
[16] W. Knap, C. Skierbiszwski, A. Zduniak, E. Litwin-Staszewska, D. Bertho, F. Kobbi, J. I. Robert, G. E. Pikus, F. G. Pikus, S. V. Iordanskii, V. Mosser, K. Zelentes, and Yu. B. Lyanda-Geller, Phys. Rev. B **53**, 3912 (1996).
[17] I. G. Savel'ev, A. M. Kreshchuk, S. V. Novikov, A. Y. Shik, G. Remenyi, Gy. Kovács, B. Pôdör, and G. Gombos, J. Phys.: Condens. Matter. **8**, 9025 (1996).
[18] L. V. Golubev, A. M. Kreshchuk, S. V. Novikov, T. A. Polyanskaya, I. G. Savel'ev, and I. I. Saĭdashev, Fiz. Tekh. Poluprovodn. **22**, 1948 (1988) [Sov. Phys. Semicond. **22**, 1234 (1988)].
[19] A. M. Kreshchuk, S. V. Novikov, and I. G. Savel'ev, Fiz. Tekh. Poluprovodn. **26**, 1375 (1992) [Sov. Phys. Semicond. **26**, 771 (1992)].
[20] B. L. Altshuler, A. G. Aronov, and D. E. Khmelnitskii, J. Phys. C **15**, 7367 (1982).
[21] T. A. Polyanskaya and Yu. V. Shmartsev, Fiz. Tekh. Poluprovodn. **23**, 3 (1989) [Sov. Phys. Semicond. **23**, 1 (1989)].
[22] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[23] E. L. Ivchenko and G. E. Pikus, *Superlattices and other Heterostructures: Symmetry and Optical Phenomena,* Springer Ser. Sol. St. Sci., (Springer-Verlag, N. Y., 1995), Vol. 110.
[24] G. M. Gusev, Z. D. Kvon, I. G. Neizvestnyi, and V. N. Ovsuk, Solid State Commun. **46**, 169 (1983).
[25] Sh. Iwabuchi and Y. Nagaoka, J. Phys. Soc. Jpn. **50**, 1325 (1989).

Translated by M. E. Alferieff

Hall effect on inertial electrons in semiconductors

V. I. Kadushkin

Scientific-Research Technological Institute, 390011 Ryazan', Russia

(Submitted March 28, 1995; accepted for publication March 6, 1996)
Fiz. Tekh. Poluprovodn. **31**, 468–469 (March 1997)

A new effect, which is attributable to the Lorentz force acting on electrons in a semiconductor undergoing acceleration, is predicted. An expression is obtained for the Hall field and the Hall voltage is estimated for a real two-dimensional heterostructure. Possible schemes for intensifying the Hall field are analyzed for the example of two Hall elements, one of which is a voltage generator and the other is a load. [S1063-7826(97)01704-3]

The Tolman–Stewart experiment, in which a current pulse j associated with the inertia of free electrons was observed, is well known. In the presence of inertial separation of charges, an electric field of intensity E arises in a conductor. If such a conductor is placed in a magnetic field B, then it can be expected that a voltage, similar to the Hall effect, will appear as a result of the action of the Lorentz force on the inertial electrons.

A current j_x and field E_x given by

$$j_x=\sigma\,\frac{m}{e}\,\frac{d\nu_x}{dt}, \tag{1}$$

$$E_x=\frac{m}{e}\,\frac{d\nu_x}{dt}, \tag{2}$$

appear in a conductor moving with acceleration $d\nu_x/dt$, where $\sigma=en\mu$ is the conductivity, and μ is the mobility.[1] In a magnetic field $B=(0,0,B_z)$ a field $E_y=(1/ne)j_xB_z$ or

$$E_y=\frac{m}{e}\,\mu B_z\,\frac{d\nu_x}{dt} \tag{3}$$

is excited. The latter expression is equivalent to $E_y=E_x\mu Bz$.

The best object for observing this effect experimentally are two-dimensional electrons in the heterosystem n-Al$_x$Ga$_{1-x}$As/GaAs. In a 1×1-cm sample with $\mu\cong10^4$ cm^2/(V·s) in a 1-T field a signal $V_y\cong6\times10^{-11}$ V can be expected for $d\nu_x/dt\cong10$ m/s^2. This is easily accessible to modern measurement techniques.

Let us consider one possibility for intensifying the effect for the example of two Hall elements, one of which (I) is a Hall-field generator and the other (II) is the load. Figure 1 shows a diagram of the connections of the Hall elements I and II.

We have established that in a magnetic field B_z (whose direction in the figure is indicated by the symbol $\oplus$) a current $j_x^{(1)}$, field $E_x^{(1)}$, and Hall field $E_y^{(1)}$, which are given by expressions (1)–(3), are excited in the first Hall element I with the potential (Hall) contacts X_1-X_1 closed on the current contacts T_2-T_2 of the Hall element II. In the latter element, in addition to the primary field $E_x^{(2)}=E_x^{(1)}$ determined by expression (2), we have the field $E_y^{(1)}$ as well, so that the resulting field has two components — $E_x^{(2)}=E_x^{(1)}+E_y^{(1)}$. This is possible if the Hall element I is regarded as a voltage generator loaded on the Hall element II. In this case, the no-load regime, for which the condition $R(X_1-X_1)\ll R(T_2-T_2)$ must be satisfied, should be attained. Here R is the resistance between the corresponding contacts. In this case the field

$$E_y^{(2)}=(E_y^{(1)}+E_y^{(1)})\mu B_z \tag{4}$$

is excited in the Hall element II. Substituting $E_y^{(1)}=E_x^{(1)}\mu B_z$, we obtain

$$E_y^{(2)}=(1+\mu B_z)\mu B_zE_x^{(1)}. \tag{5}$$

It is evidently difficult to observe the effect directly. It is more realistic to perform experiments by vibrating the sample in a magnetic field. The useful signal $\mathscr{E}_y$ can then be separated from the interference $\mathscr{E}_y^*$ according to the quadratic dependence on the vibrational frequency ω (the interference is proportional to the first power of the vibrational frequency).

Indeed, for a given geometry of the experiment (see Fig. 1) in a magnetic field $B=(0,0,B_z)$ with the coordinate x varying $x=x_0\cos\omega t$, where ω is the frequency of driving oscillator loaded on a piezoelectric element, and x_0 is the vibrational amplitude of the element, we have from the relation (3)

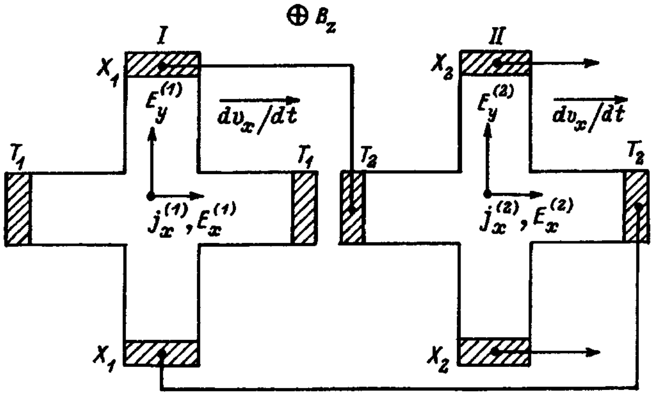

FIG. 1. Schematic diagram consisting of two elements I and II for intensifying the Hall field. The symbol $\oplus$ indicates the direction of the magnetic field B_z; the arrows indicate the direction of the acceleration dV_x/dt, the Hall fields $E_y^{(1)}$ and $E_y^{(2)}$, and the current densities $j_x^{(1)}$ and $j_x^{(2)}$.

$$\mathscr{E}_y = \frac{m}{e}\mu B_z l_y \omega^2 x_0 \cos\,\omega t, \tag{6}$$

where l_y is the distance between the Hall contacts $(X_1 - X_1)$ of the sample, i.e., $\mathscr{E}_y = E_y l_y$. The parasitic interference $\mathscr{E}_y^*$ arising in the connecting wires in accordance with Faraday's law is determined by the expression

$$\mathscr{E}_y^* = -B_z l_y^* \omega x_0 \sin\,\omega t, \tag{7}$$

where l_y^* is the effective length of the wires connecting the sample into the measuring circuit. Therefore, the useful signal $\mathscr{E}_y$ possesses features which distinguish it from the interference $\mathscr{E}_y^*$. The first feature is that it is proportional to ω^2, while $\mathscr{E}_y^* \sim \omega$. At the same time, $\mathscr{E}_y$ varies in-phase, overtime, and $\mathscr{E}_y^*$ varies out of phase with the voltage of the driving oscillator. It is important to note that the mass appearing in expressions (1)–(3) is the free-electron mass (see Ref. 2); the mobility μ, however, is determined by the effective mass.

This work was supported by the program ''Physics of solid-state nanostructures'' of the Ministry of Science and Technology of the Russian Federation, Project 2-016/2.

[1] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, translation of 1st edition, Pergamon Press, N. Y., 1960 [Russian orig., Nauka, Moscow (1982), p. 309].

[2] I. M. Tsidil'kovskihĭbĭ, Usp. Fiz. Nauk **115**, 321 (1975) [Sov. Phys. Usp. **18**, 161 (1975)].

Translated by M. E. Alferieff

Modeling of mass transfer under conditions of local gas-phase epitaxy through a mask

L. B. Proékt, M. A. Kaliteevskiĭ, V. B. Kantor, D. A. Piotrovskiĭ,
M. A. Sinitsyn, and B. S. Yavich

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

(Submitted May 29, 1996; accepted for publication September 16, 1996)
Fiz. Tekh. Poluprovodn. **31**, 470–474 (April 1997)

The diffusion mass transfer of reagents under the conditions of local gas-phase epitaxy through a mask has been studied by the method of numerical modeling. The possible effect of adsorption of reagents on the surface of the mask and their surface diffusion on the local growth rate is studied. The computational results are compared with published experimental data.

1. INTRODUCTION

An effective method for obtaining different semiconductor device structures is local gas-phase epitaxy (LGPE) through a mask.[1–5] The development of the local epitaxy technique makes it possible to apply new approaches in monolithic integration of GaAs and InP based semiconductor devices. The modern level of development of this technology makes it possible to implement growth regimes in which deposition of the grown material on the surface of the mask does not occur.[2,3]

In gas-phase epitaxy the process parameters are ordinarily chosen so that the growth rate would be limited by diffusion mass transfer of one component through the diffusion boundary layer to the surface of the substrate. In such regimes the concentration of the limiting component on the substrate surface is taken to be zero, and the boundary layer can be regarded as stationary (see Appendix).

It is well known that the growth rate in LGPE is higher than the growth rate on the open surface of the substrate in ordinary epitaxy performed under the same conditions[5,6] It is believed that this increase is due to the increase in the mass transfer through the gas phase and an additional flow of reagents over the surface of the mask. Mass transfer during local epitaxy has been modeled before,[7–9] but in the calculations the authors neglected the distortion of the concentration field of the limiting component in the diffusion boundary layer near the open surface. In general, this made it impossible to make a correct quantitative comparison of the contributions of volume and surface mass transfers to the increase in the growth rate.

Our objective was to study theoretically the effect of a mask on the growth rate of epitaxial layers during LGPE with allowance of the surface diffusion along the mask and disregarding it.

2. BASIC EQUATIONS

Let the mask consist of a periodic (period of the structure $2l$) system of strips of width $2a$ (Fig. 1a).

We assume that the growth rate is determined by the diffusion mass transfer of one component (for example, trimethyl gallium), whose concentration in the thickness of the boundary layer we denote by $C(x,z)$. For example, in the case of GaAs obtained by the GPE method this component is trimethyl gallium.

Let b be the thickness of the stationary boundary layer. The separation into a freely moving gas and a stationary boundary layer is somewhat arbitrary, since the velocity of the gas flow at the surface of a solid body varies continuously with increasing distance from the surface of the solid.[1] There arises in this connection the question of what to take as the thickness of the boundary layer in performing the modeling. This difficulty can be overcome as follows. As will be shown below, if the thickness of the boundary layer is much greater than the spatial period of the structure, then the computed growth rate does not depend on the thickness of the boundary layer. The real planar dimensions of the grown elements (several or tens of microns[5,6]) are much smaller than the thickness of the gas layer in which the diffusion mass transfer predominates over convective transfer (see Appendix).

We note that the mean-free path lengths of the molecules in GPE processes are much shorter than the characteristic dimensions of the figure of the mask (for example, under normal conditions the mean-free path length is of the order of 1 nm); this makes it possible to use a continuous medium as a model.

The distribution of the reagent concentration in the thickness of the boundary layer (Fig. 1b) is described by the solution of the two-dimensional Laplace equation

$$\frac{\partial^2 C}{\partial x^2}+\frac{\partial^2 C}{\partial z^2}=0. \tag{1}$$

As a result of the periodicity of the structure along the x axis, the following boundary conditions for the concentration can be imposed on the basis of symmetry considerations:

$$\left.\frac{\partial C}{\partial x}\right|_{x=0}=\left.\frac{\partial C}{\partial x}\right|_{x=l}=0. \tag{2}$$

We assume that the concentration at the boundary of the boundary layer equals the concentration C_0 of the limiting component in the vapor-gas mixture:

$$C|_{z-b}=C_0. \tag{3}$$

We also assume that the concentration at the open surface of the substrate equals 0 (i.e., the rate of adsorption and decom-

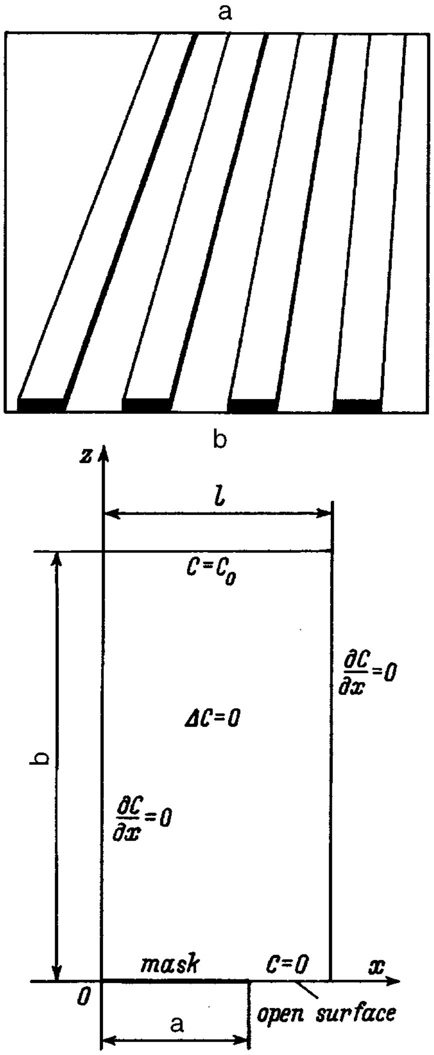

FIG. 1. a—Schematic diagram of the structure; b—diagram of the structure, formulation of the problem, boundary conditions.

position of the limiting component on the open surface is quite high). We further assume that the reagents which have reached the open surface become uniformly distributed along it.

We now consider mass transfer along the surface of the mask. We assume that the limiting reagent can be adsorbed on the surface of the mask without the formation of nucleation centers, and its surface concentration can be taken as

$$C^{\mathrm{surf}}=kC, \tag{4}$$

where C^{surf} is the surface concentration, and k is the coefficient of proportionality.

The motion of reagents along the surface of the mask can be described by the diffusion equation, taking into account the possibility of mass transfer between the volume and surface:

$$D^{\mathrm{bulk}}\frac{\partial C}{\partial z}\bigg|_{\substack{z=0\\0<x<a}}=D^{\mathrm{surf}}\frac{\partial^2 C^{\mathrm{surf}}}{\partial x^2}\bigg|_{\substack{z=0\\0<x<a}}, \tag{5}$$

where the term $D^{\mathrm{bulk}}\,\partial C/\partial z$ describes the diffusion flux from the boundary layer on the surface of the mask (D^{bulk} is the volume diffusion coefficient).

We take into account the finite (different from 1) probability of passage of a reagent particle through the mask-substrate boundary:

$$\frac{\partial C^{\mathrm{surf}}}{\partial x}\bigg|_{\substack{x=a\\z=0}}=gC^{\mathrm{surf}}\bigg|_{\substack{x=a\\z=0}}. \tag{6}$$

Here the coefficient g is proportional to the probability of passage of a reagent particle through the mask-open surface boundary. If all molecules reaching the interface cross over to the open surface of the substrate, then the boundary condition at the mask-substrate interface can be written in the form

$$C^{\mathrm{surf}}\bigg|_{\substack{z=0\\x=a}}=0. \tag{7}$$

Substituting expression (4) into Eq. (5), we obtain

$$\frac{\partial C}{\partial z}\bigg|_{\substack{z=0\\0<x<a}}=D\frac{\partial^2 C}{\partial x^2}\bigg|_{\substack{z=0\\0<x<a}}, \tag{8}$$

where $D=kd^{\mathrm{surf}}/D^{\mathrm{bulk}}$.

The relative growth rate R_g (i.e., the ratio of the growth on the mask substrate to the growth rate on the unmasked substrate with other parameters being equal) is given by

$$R_g=\frac{l}{l-a}\,\frac{\int_0^l(\partial C/\partial z)|_{z=b}dx}{l/b}. \tag{9}$$

We see that the relative growth rate R_g depends on the design of the structure (the parameters a and l), the thickness b of the boundary layer, and the parameters g and D characterizing the surface diffusion (we note that these parameters all have the dimension of length). To analyze the results, it is convenient to switch to the dimensionless parameters β, γ, and Δ, after normalizing b, g, and D to half the planar period l of the structure and after introducing the filling factor F_0 which characterizes the relative fraction of the open surface:

$$F_0=(l-a)/l,\quad \beta=b/l,\quad \gamma=g/l,\quad \Delta=D/l.$$

3. RESULTS AND DISCUSSION

The dependences of the relative growth rate on the boundary layer thickness β are shown in Fig. 2 for different filling factors F_0 and parameters characterizing the surface diffusion. One can see that in all cases this is a saturating dependence, and the saturation is reached when the thickness of the boundary layer is approximately ten times greater than the spatial period of the structure. In the case where the thickness of the boundary layer is much greater than the spatial period of the structure, the relative growth rate does not depend on the parameter β. This case is important for practical applications, and in what follows we shall investigate the behavior of the growth rate only for large values of β.

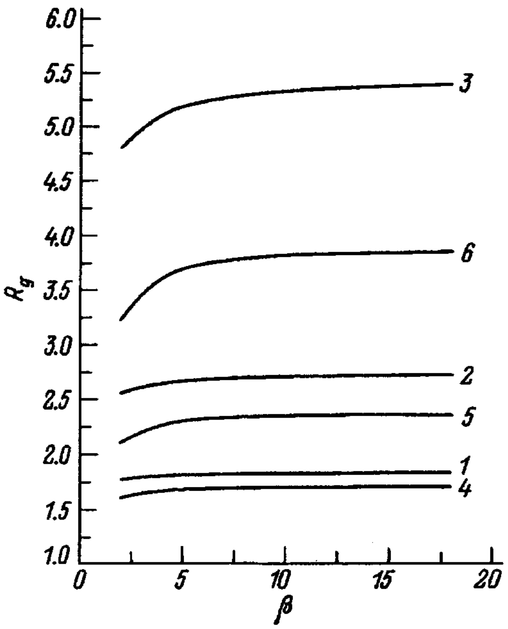

FIG. 2. Relative growth rate R_g versus relative thickness β of the boundary layer for the following values of the parameters: *1*—$F_0=0.55$, $\Delta=2.0$, $\gamma=1.0$; *2*—$F_0=0.36$, $\Delta=2.0$, $\gamma=1.0$; *3*—$F_0=0.18$, $\Delta=2.0$, $\gamma=1.0$; *4*—$F_0=0.55$, $\Delta=0$, $\gamma=0$; *5*—$F_0=0.36$, $\Delta=0$, $\gamma=0$; *6*—$F_0=0.18$, $\Delta=0$, $\gamma=0$.

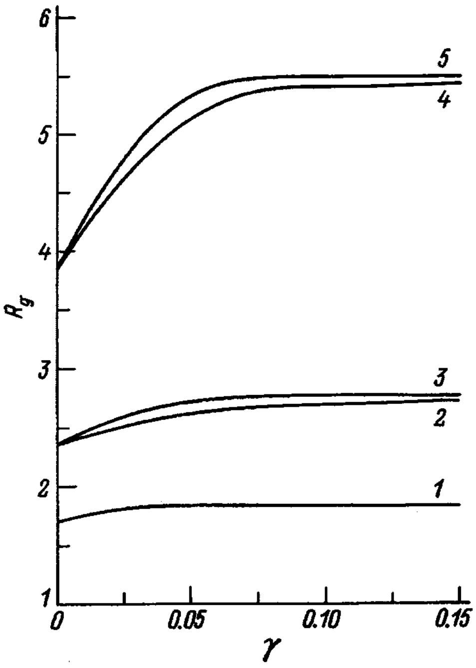

FIG. 4. Relative growth rate R_g versus the parameter γ with relative boundary-layer thickness $\beta=100$ and the following values of the parameters: *1*—$F_0=0.55$, $\Delta=2.0$; *2*—$F_0=0.36$, $\Delta=0.3$; *3*—$F_0=0.36$, $\Delta=2.0$; *4*—$F_0=0.18$, $\Delta=0.3$; *5*—$F_0=0.18$, $\Delta=2.0$.

Figure 3 shows the curves of the relative growth rate R_g versus the parameter characterizing the surface diffusion. The dependence has, naturally, a saturating form. The saturation corresponds to the case in which the surface mass transfer is so efficient that the volume concentration at the surface of the mask approaches zero.

Figure 4 shows the relative growth rate R_g versus the parameter γ, which characterizes the probability that a particle absorbed on the surface passes through the mask—open surface boundary. The value $\gamma=0$ corresponds to closing of the mass transfer channel along the surface of the mask. Large values of the parameter γ correspond to the absence of any barrier for particles diffusing along the mask at the boundary of the mask, and for this reason this dependence also has a saturating form. Saturation is reached for γ greater than 0.1.

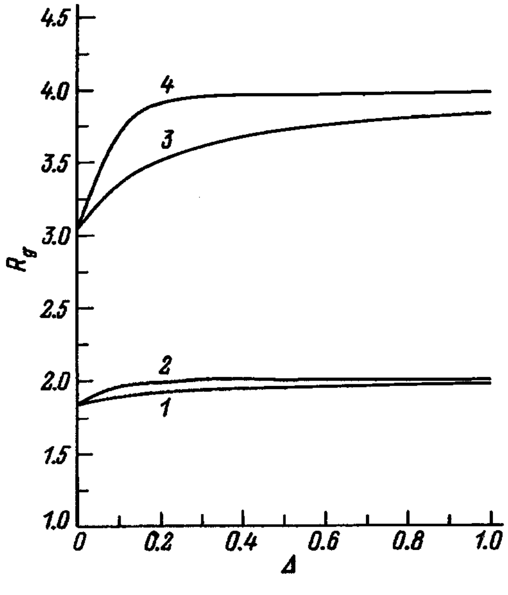

FIG. 3. Relative growth rate R_g versus the relative surface diffusion coefficient Δ for relative boundary layer thickness $\beta=100$ and the following values of the parameters: *1*—$F_0=0.5$, $\gamma=0.1$; *2*—$F_0=0.5$, $\gamma=1$; *3*—$F_0=0.25$, $\gamma=0.1$; *4*—$F_0=0.25$, $\gamma=1$.

In conclusion, we shall study the dependence of the relative growth rate R_g on the ratio of the dimensions of the mask and open surface. This growth rate can be set arbitrarily and controlled accurately in an experiment. Figure 5 shows the results of our calculations and the experimental data from Ref. 5. Curve *1* is the computational result when there is no surface diffusion, curve *3* is the computational result in the case of efficient surface diffusion (this curve practically reproduces the hyberbola $1/F_0$), and curve *2* is the computational result for an intermediate case. We see that the smaller the filling factor F_0, the more rapidly the surface diffusion increases the growth rate.

Comparing the experimental results from Ref. 5 (circles) with the computational data it can be concluded that in the system studied in Ref. 5 surface diffusion was apparently not important, since the experimentally determined values of the relative growth rate do not exceed the values of R_g calculated for the case of zero surface mass transfer. It should be noted that in Ref. 5 it is asserted that the increase in the relative growth rate R_g is due mainly to the additional mass transfer of the limiting component along the mask, although the authors do not support this assertion with any calculations.

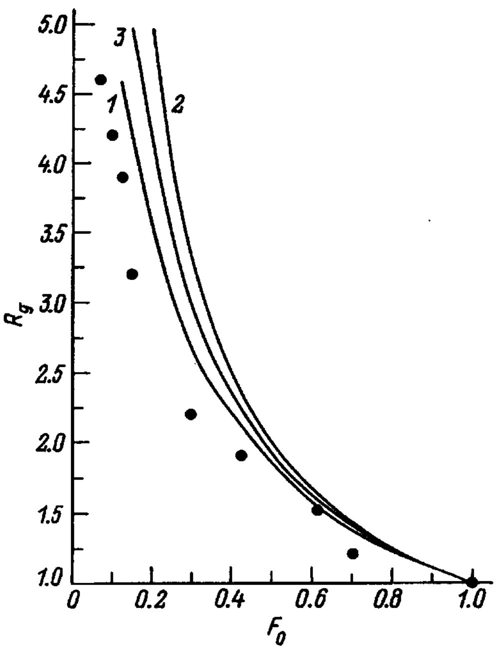

FIG. 5. Relative growth rate R_g versus the filling factor F_0 with relative boundary-layer thickness $\beta = 100$ and the following values of the parameters characterizing surface diffusion: *1*—$\Delta = 0.0$ (no surface diffusion); *2*—$\Delta = 1.0$, $\gamma = 1.0$ (efficient surface diffusion); *3*—$\Delta = 0.3, \gamma = 0.1$; circles—experimental data of Ref. 5.

The main results of the present work are as follows.

1. The relative growth rate in local gas-phase epitaxy through a mask was calculated numerically in the diffusion regime with and without surface diffusion.

2. The effect of the design of the structure, the thickness of the boundary layer, and surface diffusion along the mask on the relative growth rate was analyzed.

3. It was shown that for large boundary-layer thicknesses the relative growth rate (in the case of a periodic mask) does not depend on this thickness.

4. The calculations performed suggest that the increase in growth rate observed in local gas-phase epitaxy is due not to the additional flow of reagents over the surface of the mask but rather an increase in the diffusion flux from the gas phase.

We wish to thank R. A. Suris and A. Yu. Kaminskiĭ for helpful discussions.

4. APPENDIX

As a rule, gas-phase epitaxy processes are conducted in the diffusion regime, in which the growth rate is determined by diffusion of one of the components through a nonstationary boundary layer. This is a simplified model, since the velocity of the gas flowing around the solid surface varies with the distance from the surface continuously and not in a jump-like manner.[10] The character of this dependence is determined by the specific geometry of the reactor. Nonetheless, the concept of a ''stationary'' boundary layer near the surface can be used in the calculations.[10] In application to this problem, the characteristic thickness of the boundary layer is the distance at which the diffusion flux predominates over the convective flux. An obvious estimate of this distance is the ratio D^{bulk}/u_0, where u_0 is the characteristic velocity of the gas flow in the reactor.

TABLE I.

Pressure in the reactor P, atm	1	0.1	0.02
Volume diffusion coefficient D^{bulk}, cm^2/s	3	30	150
Gas flow velocity U_0, cm/s	10	100	500
Viscosity, cm^2/s	0.2	2	10
Boundary-layer thickness b, cm	0.9	1.2	0.7
D^{bulk}/u_0, cm	0.3	0.3	0.3

The quantities characterizing the boundary layer near the substrate holder in a reactor with a typical GPE process conducted at atmospheric and lower pressures are presented in Table I. The approximate formulas presented in Ref. 10 and the values of the volume diffusion coefficient presented in Ref. 8 were used in the calculation with the following assumptions: The carrier gas (hydrogen) flow rate is 100 cm ^{3}/s, the transverse cross-sectional area of the reactor is 10 cm^2, the linear dimensions of the substrate holder are of the order of 10 cm, and the process temperature is 630 °C, just as in Ref. 5. It is obvious that the thickness of the boundary layer, calculated according to the simplified formula from Ref. 10, is of the order of 1 cm, the same order of magnitude as the ratio D^{bulk}/u_0, which characterizes the thickness of the layer in which mass transfer of the reagent occurs mainly as a result of diffusion. In the case where the spatial period of the mask is much smaller than the thickness of the boundary layer (this case is realized in Ref. 5), the relative increase in the growth rate does not depend on the thickness of the boundary layer and an arbitrary thickness much greater than the period of the mask can be used in the calculations instead of the actual boundary-layer thickness, which, as a rule, is unknown.

[1] J. P. Duchemin, M. Bonnet, F. Kielsch, and D. H Huyghe, J. Cryst. Growth **45**, 181 (1978).
[2] K. Kamon, S. Takagish, and H. Mori, J. Cryst. Growth **73**, 73 (1985).
[3] K. Kamon, M. Shimazu, K. Kimura, M. Mihara, and M. Ishii, J. Cryst. Growth **77**, 297 (1986).
[4] G. B. Stringfellow, J. Cryst. Growth **68**, 111 (1984).
[5] Y. D. Galeuchet, P. Roentgen, and V. Graft, J. Appl. Phys. **68**, 560 (1990).
[6] T. Sasaki, M. Kitamura, and I. Mito, J. Cryst. Growth **132**, 435 (1993).
[7] D. G. Coronell and K. F. Jensen, J. Cryst. Growth **114**, 581 (1991).
[8] K. Yamaguchi, M. Ogasawara, and K. Okamoto, J. Appl. Phys. **72**, 5919 (1992).
[9] Yu. N. Makarov, M. S. Ramm, E. A. Subashieva, and A. I. Zhmakin, J. Cryst. Growth **145**, 271 (1994).
[10] H. Schlichting, *Boundary-Layer Theory,* McGraw-Hill, N. Y., 1968 [Russian trans., Nauka, Moscow, p. 134].

Translated by M. E. Alferieff

Photoconductivity of the germanium-doped solid solution p-$GaAs_{0.94}Sb_{0.06}$

T. Yu. Allen[1)] and T. A. Polyanskaya[2)]

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

A. A. Kopylov and A. A. Shakmaev

Electrical Engineering University, 197376 St. Petersburg, Russia

(Submitted August 20, 1996; accepted for publication September 16, 1996)
Fiz. Tekh. Poluprovodn. **31**, 475–477 (April 1997)

A new mechanism of impurity photoconductivity in semiconductors has been discovered. The form of the long-wavelength photoconductivity spectra observed in p-$GaAs_{0.94}Sb_{0.06}$: Ge is satisfactorily explained in terms of resonance ionization of impurity levels by phonons excited during absorption of infrared radiation. © *1997 American Institute of Physics.* [S1063-7826(97)01904-2]

We investigated the long-wavelength photoconductivity of epitaxial layers of the germanium-doped solid solution $GaAs_{0.94}Sb_{0.06}$ which were described in Ref. 1. The optical measurements were performed in the range 100–650 cm^{-1} with a LFS-1000 Fourier spectrometer and a R-118 vacuum helium cryostat. The radiation was focused on the sample with the aid of a brass cone. No impurity photoconductivity was observed at a temperature $T=80$ K. At $T=10$ K the photoresponse was recorded for samples with hole density no greater than 10^{17} cm^{-3}, which is lower than the critical charge carrier density corresponding to a metal–insulator transition in this material.[1]

The long-wavelength photoconductivity spectra which we obtained are shown in Fig. 1. Despite their complex structure, the spectra contain no characteristic features of impurity photoconductivity, such as a pronounced long-wavelength threshold. For Ge acceptor in GaAs, in accordance with the ionization energy data,[2] the ionization threshold is expected at $\hbar\omega=40.5$ meV. The photoconductivity spectra were also investigated with different electric field strengths in the sample in the range 0–60 V/cm. It should be noted that as the field strength increases, the shape of the spectra does not change appreciably. Only an absolute increase of the signal was observed, and saturation occurred at field strengths above 60 V/cm.

It is most interesting that all basic features of the absorption spectra correspond to the features of the absorption spectra associated with crystal lattice vibrations. This is clearly seen by comparing the spectra (Fig. 1, curves *2* and *3*) with the transmission spectrum (Fig. 1, curve *1*) of an epitaxial structure similar to the one which we investigated. The maxima in the photoconducitivity spectrum coincide with the two-phonon absorption maxima,[3] and a strong minimum is observed at the transverse optical (TO) phonon frequency. This indicates unequivocally that photoexcited optical phonons participate directly in activation of the photoconductivity. Efficient transfer of phonon energy to the impurity center is possible because of the closeness of the ionization energy of the acceptor ΔE_A and the energy of the optical phonons in GaAs $\omega_{TO}=33.5$ meV (Ref. 4), and the energy difference ΔE_A for the $^2P_{5/2}$–$^1S_{3/2}$ states for the Ge

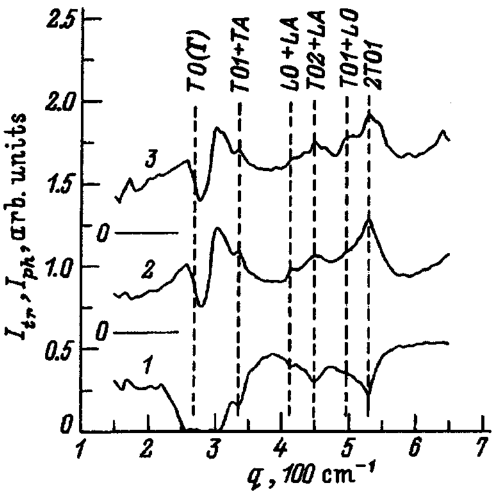

FIG. 1. Long-wavelength photoconductivity spectra of $GaAs_{0.94}Sb_{0.06}$: Ge epitaxial layers at $T=10$ K. *1* — Transmission spectrum $I_{tr}(q)$; *2*, *3* — photoconductivity spectra $I_{ph}(q)$. The reference point for curves *2* and *3* is indicated by the horizontal lines with the marker zero.

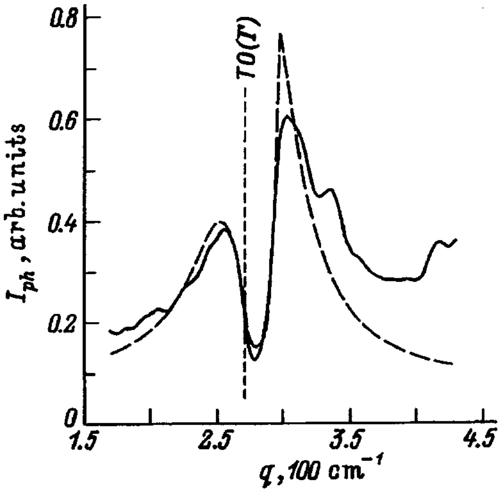

FIG. 2. Calculation of the phonon–impurity photoconductivity spectrum in the one-phonon approximation at $T=10$ K. Solid line — experiment, dashed line — calculation.

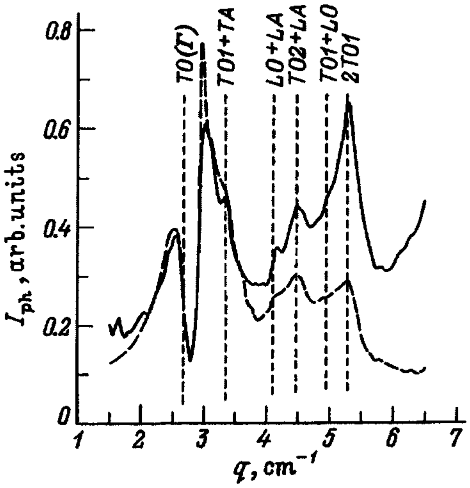

FIG. 3. Calculation of the phonon–impurity photoconductivity spectrum in the two-phonon approximation at $T=10$ K. Solid line — experiment, dashed line — calculation.

acceptor is approximately[5] 34 meV. The efficiency of energy transfer can also be increased by broadening the levels as a result of the disorder in the film.

To verify this assumption, we calculated the number of phonons excited in an epitaxial layer and the substrate under the action of the infrared radiation. Let us assume that the photocurrent is proportional to the phonon generation, and that all optical phonons created in the layer of thickness l participate with equal probability in ionization of the impurity centers. Here l is the phonon mean free path. To calculate the index of refraction, we employed the standard one-oscillator model

$$n(\omega)=\left(\varepsilon_0+\frac{(\varepsilon_0-\varepsilon_\infty)\omega_{\mathrm{TO}}^2}{\omega_{\mathrm{TO}}^2-\omega^2-i\omega\gamma}\right)^{1/2}, \tag{1}$$

where ε_0 and ε_∞ are the static and high-frequency permittivities, and γ and ω_{TO} are the damping constant and TO-phonon frequency. The contribution of free charge carriers was disregarded. Furthermore, for the calculations we employed the optical constants of GaAs, since the content of antimony in the epitaxial layer is low. As a result,

$$N_{\mathrm{phon}}(\hbar\omega)=I_0(1-R)[1-\exp(-2\omega n''l/c)]/\hbar\omega, \tag{2}$$

where I_0 is the intensity of the incident light, R is the reflection coefficient of the surface of the structure, and $n''=n''(\omega)$ is the imaginary part of the index of refraction (1).

In Fig. 2 the computational result obtained using Eq. (2) is compared with the observed photoconductivity spectrum. The best agreement between the curves is obtained for $\gamma=5$ cm^{-1} and $\omega_{\mathrm{TO}}=272$ cm^{-1} for layers of thickness $l=20$ μm. We used $\varepsilon_0=12.9$ and $\varepsilon_\infty=10.9$ (Ref. 6). We see that there is good agreement in the region of strong one-phonon absorption. The discrepancy between the curves in the short-wavelength region of the spectrum ($\omega>\omega_{\mathrm{TO}}$) is due to the contribution of two-phonon absorption, which can be taken into account empirically,

$$N_{\mathrm{phon}}(\hbar\omega)=I_0(1-R)[1-\exp(-2\omega n''/c+w\alpha_{\mathrm{2phon}})l]\hbar\omega, \tag{3}$$

where α_{2phon} is the two-phonon absorption coefficient, which was calculated as the difference between the measured value of the absorption coefficient of GaAs and the theoretical value of the one-phonon absorption coefficient $\alpha_{\mathrm{1phon}}=2\omega n''/c$, and w is an empirical weighting coefficient, which accounts for the contribution of different types of phonons in the ionization of the impurity states. The contribution of the difference two-phonon absorption was ignored, since its probability at $T\simeq10$ K is small. The best agreement between theory and experiment is obtained with $w\simeq10$ (Fig. 3). The fact that the value of the weighting coefficient was found to be greater than 1 can be explained by the greater efficiency of the interaction of impurity centers with phonons with a nonzero wave vector.

The discrepancy between the curves in the region $\hbar\omega>45$ meV could be due to several factors. First, the dispersion parameters of optical phonons are not known accurately for lattices of solid solutions. Second, we do not have sufficient information about the characteristics of the electron-phonon interaction at impurity centers. In addition, we disregarded the contribution of photoionization processes. An attempt to explain the observed photoconductivity spectra with the help of a bolometric effect does not make it possible to describe correctly the TO-phonon structure.

[1)] Present address: University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA. Electronic mail: tbilgild@utcvm.utc.edu
[2)] Electronic mail: pta@nano.ioffe.rssi.ru

---

[1] T. Yu. Allen and T. A. Polyanskaya, Fiz. Tekh. Poluprovodn. **31**, 405 (1997) [Semiconductors **31**, 405 (1997)].
[2] D. J. Ashen, P. J. Dean, D. T. J. Hurle, J. B. Mullin, and A. M. White, J. Phys. Chem. Sol. **36** 1041 (1975).
[3] V. Spitzer, in *Optical Properties of Semiconductors*, edited by R. Willardson and A. Beer [Russian translation], Mir, Moscow, 1970, p. 28.
[4] G. Dolling and WJ. L. T. Waugh, in *Lattice Dynamics*, edited by R. F. Wallis (Pergamon Press, 1965) p. 19.
[5] A. Baldereschi and N. O. Lipari, Phys. Rev. B **9**, 1525 (1974).
[6] Landolt-Bornstein, *Numerical Data and Functional Relationship in Science and Technology*. New Series: *Group III*, V. 17; subsol. a: *Physics of Group IV Elements and III-V Compounds*, edited by O. Madelung, Springer-Verlag (1982).

Translated by M. E. Alferieff

Propagation of a surface acoustic wave in a layered system containing a two-dimensional conducting layer

V. D. Kagan[1)]

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

(Submitted July 22, 1996; accepted for publication October 23, 1996)
Fiz. Tekh. Poluprovodn. **31**, 478–482 (April 1997)

A surface acoustic wave can propagate in a piezoelectric crystal above which lies a two-dimensional conducting layer. The electric fields excited by the wave in the crystal penetrate into the two-dimensional layer and excite dissipative currents in it, which results in absorption of the wave and change of its velocity. These characteristics are calculated for different configurations of the layered system taking into account not only the surface conductivity but also surface diffusion. When the layer has an elastic contact with the crystal, for the configuration allowance is made not only for the piezoelectric but also the deformation interaction of the sound wave with the electrons. © *1997 American Institute of Physics.* [S1063-7826(97)02004-8]

1. INTRODUCTION

Investigations of the properties of two-dimensional conducting channels, formed in semiconductor heterostructures, especially in the region of strong magnetic fields, where the quantum Hall effect regime is realized, have been exciting great interest recently. Very promising methods for studying the properties of a two-dimensional electron gas are acoustic methods, since they make it possible to study the rf conductivity of two-dimensional systems in a contact-free manner. These methods consist of investigating the effect of a conducting channel on the propagation of a surface acoustic wave in a piezodielectric crystal (i.e., on the absorption coefficient Γ and the sound velocity v). Analysis of the rf conductivity as a function of frequency, temperature, and magnetic field makes it possible to study the mechanisms for localization of two-dimensional electrons in the quantum Hall effect regime.

The first experimental studies[1–5] were performed on GaAs/AlGaAs structures with a two-dimensional electron gas. The rf conductivity in the quantum Hall effect regime in a $\mathrm{In}_x\mathrm{Ga}_{x-1}\mathrm{As/InP}$ structure was studied by acoustic methods in Ref. 6. In Refs. 2–5, the theoretical formulas for absorption and velocity of sound, extending to the two-dimensional case White's standard formula[7] for a three-dimensional medium, were used to analyze the experimental results. However, these formulas ignore the complicated geometry of a layered system. More complicated formulas, which take into account this geometry only partially, are given elsewhere[1] and in the theoretical paper,[8] which contains some errors that are noted below.

This paper consists of two sections. In Sec. 1 the formulas for the velocity and absorption coefficient of a surface acoustic wave are derived for the case of a layered system with a complex geometry. Generalizing all previous analyses, we take into account the fact that the current in the conducting channel consists of field and diffusion components. In Sec. 2, we examine the role of the deformation interaction between the surface acoustic wave and the electrons in the two-dimensional channel. Such an interaction is possible only when a mechanical contact is present between the medium, along which the surface acoustic wave propagates, and the conducting channel. An attempt to take into account of this interaction was made in Ref. 5, but we consider it to be unsuccessful. We derive the boundary condition for the equations of elasticity on a two-dimensional conducting layer. For the deformation interaction of a surface sound wave with electrons of a two-dimensional layer, this condition plays the role of boundary conditions of electrostatics for a piezoelectric interaction. The velocity and absorption of sound in the two cases are therefore expressed in terms of the same characteristics of the conducting layer.

2. CALCULATION OF THE ABSORPTION COEFFICIENT AND VELOCITY OF A SURFACE ACOUSTIC WAVE PROPAGATING IN A LAYERED SYSTEM WITH A COMPLEX GEOMETRY

The layered system consists of the half space, $x<0$, occupied by a piezoelectric crystal and a semiconductor placed above it in such a way that a vacuum gap of thickness a exists between them. The conductivity of the semiconductor is negligible. A two-dimensional conducting layer is built into the semiconductor at a distance d from its boundary. A surface acoustic wave (SAW) with wave vector k, frequency ω, and velocity ν, so that $\omega=\nu k$, propagates near the surface of the piezoelectric crystal. The electric fields excited by this wave exit from the piezoelectric crystal and penetrate into the semiconductor. These fields induce electric currents in the conducting layer, Joule heating occurs, and the energy of the sound wave dissipates. The absorption of the SAW and the change in its velocity in the presence of a layered structure are determined by the characteristics of the conducting layer. Because of the vacuum gap, the piezoelectric crystal is coupled with the semiconductor only electrically, there is no elastic contact, and therefore only the electrostatics problem need be solved. The electric fields are calculated by the method developed in Ref. 9. The surface wave consists of two coherent wave components with displacements: $u_i^{(\alpha)}(\mathbf{r},\,t)$, where α takes on the values 1 and 2. The waves propagate along the z axis and decay into the piezoelectric crystal at a damping rate γ_α proportional to k:

$$u_i^{(\alpha)}(\mathbf{r},t)=u_i^{(\alpha)}e^{ikz+\gamma_\alpha x},\quad u_i^{(\alpha)}\sim e^{-i\omega t}. \tag{1}$$

We assume that the z axis is the symmetry axis: The permittivity parallel to the axis is $\varepsilon_\parallel$ and the permittivity perpendicular to the axis is $\varepsilon_\perp$. The equation for the electric potential φ in the piezoelectric crystal is

$$\varepsilon_\parallel\frac{\partial^2\varphi^{(\alpha)}}{\partial z^2}+\varepsilon_\perp\frac{\partial^2\varphi^{(\alpha)}}{\partial x^2}=4\pi\beta_{l,ik}\frac{\partial^2 u_i^{(\alpha)}(\mathbf{r},t)}{\partial x_l\partial x_k}, \tag{2}$$

where $\beta_{l,ik}$ is the piezoelectric tensor. The solution of Eq. (2) consists of the solution of the problem of forced oscillations

$$\varphi_{\text{piez}}^{(i)}=\frac{4\pi\beta_{l,ik}\,(\partial^2/\partial x_l\partial x_k)\,(u_i^{(\alpha)}e^{ikz+\gamma_\alpha x})}{-\varepsilon_\parallel k^2+\varepsilon_\perp\gamma_\alpha^2} \tag{3}$$

and a free solution, which is the solution of the Laplace equation for an anisotropic medium is,

$$\varphi_{\text{piez}}^{(f)}=\varphi_p e^{ikz+\sqrt{\varepsilon_\parallel/\varepsilon_\perp}x},\quad \varphi_{\text{piez}}=\varphi_{\text{piez}}^{(i)}+\varphi_{\text{piez}}^{(f)}. \tag{4}$$

In the vacuum gap and in the semiconductor the dielectric constants are isotropic and equal to ε_0 and ε_s, respectively. The electric field potential is the solution of the Laplace equation

$$\varphi_0=(\varphi_0^+e^{kx}+\varphi_0^-e^{-kx})e^{ikz},\quad 0<x<a; \tag{5}$$

$$\varphi_s=(\varphi_s^+e^{kx}+\varphi_s^-e^{-kx})e^{ikz},\quad a<x<a+a; \tag{6}$$

$$\varphi_s^>=\varphi^>e^{-kx+ikz},\quad x>a+d. \tag{7}$$

Electrostatic boundary conditions, imposed on the tangential (along the z axis) components of the electric field and the normal components (along the x axis) of the electric induction, must be satisfied at all interfaces. The electric induction P in the piezoelectric crystal has the additional term

$$P_x=-\varepsilon_\perp\frac{\partial\varphi}{\partial x}+4\pi\beta_{x,ik}\frac{\partial}{\partial x}(u_k^{(\alpha)}e^{ikx+\gamma_\alpha x}). \tag{8}$$

The induction components have a jump on the two-dimensional conducting layer: Their difference is $4\pi\rho$, where ρ is the surface charge density.[8] We write for the charge density the continuity equation, in which, generalizing the standard analysis, we take into account the surface diffusion current together with the ohmic current:

$$\frac{\partial\rho}{\partial t}+\frac{\partial j}{\partial z}=0,\quad j_z=-\sigma\frac{\partial\varphi}{\partial z}\bigg|_{x=a+d}-D\frac{\partial\rho}{\partial z}, \tag{9}$$

where σ is the conductivity, and D is the diffusion coefficient.

The system of boundary conditions for the three planes ($x=0$, $x=a$, and $x=a+d$) is the complete system for the six amplitudes of the potential (4)–(7), making it possible to express them in terms of the elastic displacements in the surface wave.

In the piezoelectric crystal, the equations of the theory of elasticity include a term with an electric field. Substituting into it the potential (4) expressed in terms of the displacement, we obtain a system of equations that determines the complex dispersion equation for the surface wave. The dissipative currents in the conducting layer result in damping of the amplitude of the sound wave in accordance with $\exp(-\Gamma z/2)$. We determine Γ by the following procedure: We multiply the equations of the theory of elasticity by $u_i^{(\alpha)}e^{-ikz}$, sum over the signs of the components of the displacement i and over the branch numbers α, set $x=0$ (i.e., at the surface of the piezoelectric crystal), and take from these sums the imaginary part. We can then write

$$\Gamma=\chi^{(1)}k\frac{(2\pi\sigma/\varepsilon_s\nu)\,f_2(k)}{1+[D\,(k/\nu)+(2\pi\sigma/\nu\varepsilon_s)\,f_1(k)]^2}, \tag{10}$$

$$f_1(k)=1+e^{-2kd}$$
$$\times\frac{(\varepsilon_1+\varepsilon_0)(\varepsilon_s-\varepsilon_0)-e^{-2ka}(\varepsilon_1-\varepsilon_0)(\varepsilon_s+\varepsilon_0)}{(\varepsilon_1+\varepsilon_0)(\varepsilon_s+\varepsilon_0)-e^{-2ka}(\varepsilon_1-\varepsilon_0)(\varepsilon_s-\varepsilon_0)},$$

$$\varepsilon_1=\sqrt{\varepsilon_\perp\varepsilon_\parallel};$$

$$f_2(k)=\frac{8\varepsilon_0^2\varepsilon_s(\varepsilon_1+\varepsilon_0)e^{-2k(a+d)}}{[(\varepsilon_1+\varepsilon_0)(\varepsilon_s+\varepsilon_0)-e^{-2ka}(\varepsilon_1-\varepsilon_0)(\varepsilon_s-\varepsilon_0)]^2},$$

$\chi^{(1)}$ is the piezoelectric coupling constant, which is ordinarily small in the known piezoelectric crystal. We express it in terms of the amplitude of the surface wave

$$\chi^{(1)}=\frac{4\pi}{(\varepsilon_1+\varepsilon_0)\rho\omega^2\sum_\alpha u_i*(\alpha)u_i^{(\alpha)}}\text{Re}\sum_\alpha A_2^{(\alpha)}$$
$$\times\left[B^{(\alpha)}+\frac{\varepsilon_1-(\varepsilon_\perp\gamma_\alpha/k)}{\varepsilon_\parallel k^2-\varepsilon_\perp\gamma_\alpha^2}A_1^{(\alpha)}\right], \tag{11}$$

$$A_1=\beta_{l,ik}u_k^{(\alpha)}e^{-ikz}\frac{\partial^2}{\partial x_l\partial x_i}(e^{ikz+\gamma_\alpha x})\bigg|_{x=0},$$

$$A_2=\beta_{l,ik}u_k^{*(\alpha)}e^{-ikz}\frac{\partial^2}{\partial x_l\partial x_k}(e^{ikz+k\sqrt{(\varepsilon_\parallel/\varepsilon_\perp)}x})\bigg|_{x=0},$$

$$B=-(\beta_{x,ik}u_k^{(\alpha)}/k)e^{-ikz}\frac{\partial}{\partial x_i}(e^{ikz+\gamma_\alpha x})\bigg|_{x=0}.$$

The constant $\chi^{(1)}$ can be calculated only by means of unwieldy numerical calculations, which are determined by the structure of the surface acoustic wave. We shall try to relate $\chi^{(1)}$ with an experimentally measured quantity. In the case of the propagation of a wave in a piezoelectric crystal bordering only on the vacuum, it is possible to perform measurements of the velocity of the wave with (ν_0) and without (ν) surface metallization. In the former case, the electric fields are closed and completely ''switched off.'' The piezoelectric coupling constant is ordinarily determined as the relative difference in the velocities in these two cases, and in most studies[1–6] it is assumed to be equal to $\chi^{(1)}$. Our calculation showed, however, that there is some difference:

$$\nu^2=\nu_0^2[1+\chi^{(1)}+\chi^{(2)})], \tag{12}$$

$$\chi^{(2)}=\frac{4\pi}{\rho\omega^2\Sigma_\alpha u_i^{*(\alpha)}u_i^{(\alpha)}}\text{Re}\sum_\alpha\frac{(\widetilde{A}_1^{(\alpha)}-A_2^{(\alpha)})A_1^{(\alpha)}}{\varepsilon_\parallel k^2-\varepsilon_\perp\gamma_\alpha^2},$$

$$\widetilde{A}_1 = u_k^{*(\alpha)}\beta_{l,ik} e^{-ikz}\frac{\partial^2}{\partial x_l \partial x_i}\left(e^{ikz+\gamma_\alpha x}\right)\Bigg|_{x=0}.$$

It is difficult to calculate $\chi^{(2)}$, but it should be noted that it does not equal zero identically. In Ref. 6, Eq. (10) was used to analyze the experimental data on the absorption of surface acoustic waves in $LiNbO_3$ by a two-dimensional electron gas in a InGaAs/InP heterostructure.

In Ref. 1, the SAW damping coefficient was determined for a layered system in which the conducting layer borders on the vacuum $(d=0)$. Disregarding diffusion, the frequency dependence (10) is identical to that in Ref. 1, and the difference reduces only to the numerical value, already noted above, of the coupling constant.

In Ref. 8 the absorption coefficient was calculated by a method different from ours for a system without a vacuum gap $(a=0)$. The authors determined it as the ratio of the energy dissipation in the bulk, equal to the product of a unit area by the thickness of the conducting layer, to the flux density in the wave. This quantity does not have the dimension of Γ, and calculating it by the method of Ref. 8 does not give the frequency dependence (10) with $a=0$. The authors included the factor k and obtained the correct frequency dependence from dimensional considerations.

The absorption coefficient must in fact, be determined as the ratio of the energy dissipation, which they indicated, to the total energy flux in the volume, equal to the product of a unit area by the entire thickness of the piezoelectric crystal. For this definition of Γ, there arises in the denominator of the quantity employed in Ref. 8, an additional integration over x, which gives an additional factor $1/\gamma_\alpha$. This leads to the appearance of a factor k in the final formula for the absorption coefficient Γ.

The complex dispersion relation makes it possible to determine the velocity of a surface wave in the presence of a semiconductor and to compare it with ν_0. This difference is given by the real part of the sum whose imaginary part determines Γ:

$$\frac{\nu^2-\nu_0^2}{\nu_0^2} = \chi^{(2)} + M^{(1)} + M^{(2)} + M^{(3)}. \tag{13}$$

All quantities $M^{(i)}$are proportional to the constant $\chi^{(1)}$. Strange as it may be, the quantity $M^{(1)}$ does not depend on the characteristics of the conducting layer:

$$M^{(1)} = \chi^{(1)}\frac{(\varepsilon_0+\varepsilon_1)}{M_1+M_2}\{(\varepsilon_s+\varepsilon_0)[1-e^{-2k(a+d)}] + (\varepsilon_s-\varepsilon_0)(e^{-2kd}-e^{-2ka})\}, \tag{14}$$

$$M_1 = (\varepsilon_1+\varepsilon_0)(\varepsilon_s+\varepsilon_0),$$

$$M_2 = e^{-2ka}(\varepsilon_1-\varepsilon_0)(\varepsilon_s-\varepsilon_0) + e^{-2kd}(\varepsilon_1+\varepsilon_0)(\varepsilon_s-\varepsilon_0) - e^{-2k(a+d)}(\varepsilon_1-\varepsilon_0)(\varepsilon_s+\varepsilon_0),$$

$$M^{(2)} = 2\frac{D}{\nu}\Gamma, \tag{15}$$

$$M^{(3)} = \chi^{(1)}\frac{(1+4D^2k^2/\nu^2)f_2(k)}{f_1(k)\left\{1+\left[\frac{Dr}{\nu}+\frac{2\pi\sigma}{\varepsilon\nu_s}f_1(k)\right]^2\right\}}. \tag{16}$$

Experiments are often performed in a simpler system: The two-dimensional conducting channel is formed by layers which are produced by molecular epitaxy directly on the surface of a piezodielectric crystal.[2–5] Our formulas are completely suitable for describing such a system as well—one need only set $a=0$ and $d=0$. In addition, the following circumstance must be taken into account: In our analysis the upper half-space is filled by the semiconductor, while in the experiments of Refs. 2–5 the upper half-space is a vacuum, so that in Eqs. (10) and (13)–(16) ε_s is replaced by ε_0. Therefore, we obtain for Γ the following formula which generalizes the analysis given in Refs. 2–5, taking into account the surface diffusion:

$$\Gamma = \chi^{(1)}k\frac{4\pi\sigma/(\varepsilon_1+\varepsilon_0)\nu}{1+[Dk/\nu+4\pi\sigma/\nu(\varepsilon_1+\varepsilon_0)]^2}. \tag{17}$$

The equation for the velocity [Eqs. (13)–(16)] also simplify. Equation (15) is expressed in terms of Eq. (17). For the other terms we have

$$M^{(1)}=0, \quad M^{(2)} = \chi^{(1)}\frac{(1+4(D^2k^2/\nu^2)}{1+[D\,(k/\nu)+(4\pi\sigma/(\varepsilon_1+\varepsilon_0)u)]^2}. \tag{18}$$

3. CALCULATION OF THE ABSORPTION COEFFICIENT AND SOUND VELOCITY IN THE PRESENCE OF A DEFORMATION INTERACTION OF A SURFACE ACOUSTIC WAVE WITH A TWO-DIMENSIONAL ELECTRON GAS

In the experimental arrangement described above, a mechanical contact exists between the conducting layer and the solid, along which a surface wave propagates. Therefore, it is necessary to take into account the deformation interaction between the charge carriers and the sound wave. In a piezoelectric crystal this interaction is added to the piezoelectric interaction, but it also exists in nonpiezoelectric crystals. The effect of a sound wave on the electrons is manifested in the fact that it is necessary to include an additional current in Eq. (9) that depends on the corresponding constant E of the deformation potential:

$$j_z = \frac{\sigma}{e}\frac{\partial}{\partial z}\left(E\frac{\partial u_z}{\partial z}\right). \tag{19}$$

However, in the equations of the theory of elasticity the additional stress tensor is proportional to the electron density in the solid, and such a volume density equals zero in a crystal. It turns out that the action of the electrons in the sound wave is manifested not in volume forces but rather only as a change in the boundary condition. The boundary condition is derived in the same manner as in electrostatics—the jump in the normal component of the electric induction equals the surface charge density. We obtained the boundary condition in which the presence of a conducting layer is taken into account:

$$\frac{\lambda_{ixlm}}{2}\left(\frac{\partial u_l}{\partial x_m}+\frac{\partial u_m}{\partial x_l}\right)+\frac{E}{e}\,\frac{\partial \rho}{\partial z}\,\delta_{iz}\bigg|_{x=0}=0. \tag{20}$$

Here λ_{iklm} is the tensor of the elastic moduli. This boundary condition shows that the effect of the deformation interaction can be taken into account only by calculating completely the structure of the surface wave. We performed such a calculation for the simplest model—a Rayleigh surface wave in an anisotropic solid. The presence of a conducting layer on the surface of the solid requires that the standard theory[9] be modified by the boundary condition (20), the continuity equation (9), the Laplace equation for the potential, and electrostatic boundary conditions. It is well known that a Rayleigh wave propagates with velocity $\nu = c_t \zeta$, where ζ is a number which depends on the ratio c_t/c_l; here c_l and c_t are the velocities of the longitudinal and transverse sound waves, respectively. For the deformation damping of sound and the change in the velocity of the wave, because of the presence of the conducting layer, we obtain

$$\Gamma=\frac{E^2k^3}{\rho c_t^2 e^2}\,\frac{\sigma/\nu}{1+[Dk/\nu+4\pi\sigma/\nu(\varepsilon_1+\varepsilon_0)]^2}\,f(\zeta), \tag{21}$$

$$\frac{\nu^2-\nu_0^2}{\nu_0^2}=\frac{E^2k^2}{\rho c_t^2 e^2}\,\frac{(\sigma/\nu)\,[D(k/\nu)+4\pi\sigma/(\varepsilon_1+\varepsilon_0)\nu]}{1+[D\,(k/\nu)+4\pi\sigma/(\varepsilon_1+\varepsilon_0)\nu]^2}\,f(\zeta),$$

$$f(\zeta)=\frac{(1-\zeta^2)\sqrt{1-\zeta^2(c_t/c_l)^2}}{12\zeta^2-\zeta^4+2(c_t/c_l)^2-4(c_t/c_l)^2\zeta^2}.$$

We assume that the coupling constant $(E^2k^2/\rho c_l^2 e^2)$, which replaces χ in a piezoelectric crystal, is small. Comparison of Eq. (21) with Eqs. (17) and (18) shows that taking into account the deformation interaction instead of the piezoelectric interaction reduces mainly to a replacement of the coupling constant. For volume waves in a piezoelectric crystal , the coupling constants were added together; for a system where an acoustic contact exists between the conducting layer and the crystal, mixed terms are possible as well.

In Ref. 5 an attempt was made to take into account the deformation interaction of a sound wave with a conducting layer. However, the theory could not describe the dissipative currents in the two-dimensional layer. It can be shown that the final result in Ref. 5 does not contain, first of all, quantities which are proportional to the conventional deformation coupling constant E and, second, solids with the surface characteristics of the two-dimensional layer, which characterize a layer with ‘‘zero’’ thickness. The results presented in Ref. 5 show, however, that the absorption coefficient Γ is proportional to a quantity characterizing the dispersion dependence of the deformation potential on the energy, and since it is small in semiconductors, it has never been investigated. Moreover, this is the standard volume contribution to Γ and hence for a two-dimensional layer it is proportional to the thickness of the layer. For a ‘‘point-like’’ layer, this contribution equals zero. It is obvious that the contribution to Γ calculated in Ref. 5 is very small compared with Eq. (21). It can therefore be invoked to explain an experiment only as a small correction to the deformation absorption, similar to the well-known piezoelectric contribution (17) and (18).

4. CONCLUSIONS

It should be noted that the study of the rf conductivity of a two-dimensional electron gas by acoustic, contact-free methods has been found to be very informative for investigating the problems of localization of two-dimensional electrons in the quantum Hall effect regime.[6,10,11] Such investigations require a quantitative analysis of the experimentally observed dependences of the absorption coefficient of a surface acoustic wave (SAW) on the magnetic field, the SAW frequency, and the temperature. The formulas obtained by us were employed successfully to analyze the experiment performed by us. However, analysis of Refs. 2–5 showed that Eq. (17) (for the case $a=d=0$, disregarding diffusion), which the authors employed to analyze the experimental data, explains only qualitatively the experimental magnetic field dependences of the absorption coefficient and velocity of sound. It is possible that quantitative agreement could be achieved by examining a more complicated experimental geometry.

I wish to thank I. L. Drichko for helpful discussions and for assistance in writing this paper.

This work is supported by the Russian Fund for Fundamental Research Grant No. 95-02-0466-a.

1)Electronic mail: kagan@phk.pti.spb.su

---

[1] A. Schtenstrom, Y. J. Quan, M. F. Zu, H. P. Baum, M. Levy, and B. K. Sarma, Solid State Commun. **65**, 739 (1988).

[2] A. Wixforth, J. P. Kotthaus, and G. Weimann, Phys. Rev. Lett. **56**m, 2104 (1986); A. Wixforth, J. Scriba, M. Wassermeier, J. P. Kotthaus, G. Weimann, and W. Schlapp, Phys. Rev. **40**, 7874 (1989).

[3] R. L. Willet, M. A. Paalanen, R. R. Ruel, K. W. West, L. N. Pfeiffer, and D. J. Bishop, Phys. Rev. Lett. **65**, 112 (1990).

[4] R. Boulet, P. Coleridje, F. Guillon, M. D’Iorio, and A. Sachrajda, Can. J. Phys. **69**, 461 (1991).

[5] V. W. Rampton, K. McEnaney, A. G. Kozorezov, P. J. A. Carter, C. D. Wilkinson, M. Henin, and O. H. Hughes, Sem. Sci. Tech. **7**, 647 (1992).

[6] I. L. Drichko, A. M. D’yakonov, V. D. Kagan, A. M. Kreshchuk, G. D. Kipshidze, T. A. Polyanskaya, I. G. Savel’ev, I. yu. Smirnov, A. V. Suslov, and A. Ya. Shik, Fiz. Tekh. Poluprovodn. **29**, 1306 (1995) [Semiconductors **29**, 677 (1995)].

[7] D. L. White, J. Appl. Phys. **33**, 2547 (1962).

[8] A. L. Efros and Yu. M. Galperin, Phys. Rev. Lett. **64**, 1959 (1990).

[9] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, Pergamon Press, N. Y., 1986, 3rd English edition [Russian orig., Nauka, Moscow, 1987].

[10] I. L. Drichko, A. M. Diakonov, V. D. Kagan, A. M. Kreshchuk, G. D. Kipshidze, T. A. Polyznskaya, I. G. Sacel’ev, I. Yu. Smirnov, and A. V. Suslov, Phys. Low-Dim. Struct. **10/11**, 275 (1995).

[11] I. L. Drichko, A. M. D’yakonov, A. M. Kreshchuk, T. A. Polyanskaya, I. G. Savel’yev, I. Yu. Smirnov, and A. V. Suslov, Fiz. Tekh. Poluprovodn. **31**, 451 (1997) [Semiconductors **31**, 384 (1997)].

Translated by M. E. Alferieff

Injection heterolaser based on an array of vertically aligned InGaAs quantum dots in a AlGaAs matrix

A. E. Zhukov, A. Yu. Egorov, A. R. Kovsh, V. M. Ustinov, N. N. Ledentsov,
M. V. Maksimov, A. F. Tsatsul'nikov, S. V. Zaĭtsev, N. Yu. Gordeev, P. S. Kop'ev, and
Zh. I. Alferov

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

D. Bimberg

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

(Submitted October 29, 1996; accepted for publication October 31, 1996)
Fiz. Tekh. Poluprovodn. **31**, 483–487 (April 1997)

Arrays of vertically aligned InGaAs quantum dots in a AlGaAs matrix have been investigated. It is shown that increasing the band gap of the matrix material makes it possible to increase the localization energy of quantum dots relative to the edge of the matrix band, as well as the states of the wetting layer. The use of an injection laser as the active region makes it possible to decrease the thermal filling of higher-lying states, and thereby decrease the threshold current density to 63 A/cm^2 at room temperature. A model explaining the negative characteristic temperature section observed at low temperatures is proposed. The model is based on the assumption that a transition occurs from nonequilibrium to equilibrium filling of the states of the quantum dots. © *1997 American Institute of Physics.* [S1063-7826(97)02104-2]

Arrays of semiconductor quantum dots (QDs) are promising for applications in injection lasers. Specifically, it has been predicted that the threshold current density J_{th} will decrease substantially and the temperature dependence of J_{th} will be reduced.[1] The use of InGaAs vertically aligned quantum dots (VAQDs) in a GaAs matrix as the active region of a laser made it possible to decrease J_{th} to ~100 A/cm^2 at 300 K.[2] It was shown that one of the main factors determining the threshold current density in such lasers at room temperature is the filling of the states of the wetting layer and the GaAs matrix by carriers as a result of the relatively low carrier localization energy in quantum dots.[3]

The structural and optical properties of InGaAs QDs, placed in a GaAs matrix, as well as injection lasers based on them have been studied extensively.[4–6] We chose as the matrix material the solid solution $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ ($x \leqslant 0.3$). We have shown that the use of a wider-gap compound as a matrix (GaAs is replaced by the solid solution AlGaAs) makes it possible to increase substantially the energy gap between the states of the QDs and the wetting layer, as well as the band edge of the matrix material. The use of such structures as the active region of an injection laser makes it possible to decrease the threshold current density to 63 A/cm^2 at room temperature. A section of negative characteristic temperature has been observed at low temperatures and a model is proposed explaining the anomalous temperature dependence of the threshold current density.

A Riber 32P molecular-beam epitaxy system with a solid-state source of arsenic was used to grow the structures on GaAs (100) substrates. To avoid segregation and reevaporation of In, during deposition of the QDs and the 100-Å-thick $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ layer, directly next to the array of QDs, the substrate temperature was chosen to be 485 °C. The growth temperature of the rest of the structure was equal to 700 °C. Growth was conducted under As-stabilized conditions. The array of vertically aligned quantum dots was formed as a result of triple deposition of QD layers separated by AlGaAs layers with an effective thickness of 50 Å. $\mathrm{In}_{0.5}\mathrm{Ga}_{0.5}\mathrm{As}$ with an effective thickness of 12 Å was used as the material for the quantum dots.

When a strongly strained InGaAs layer is deposited on an AlGaAs (100) surface, at a critical thickness a morphological transition occurs from a two-dimensional (layerwise) to a three-dimensional (island-wise) growth regime. An array of InGaAs islands, lying on a thin (about two monolayers) InGaAs wetting layer, forms on the surface. Buried by the AlGaAs layer, these islands can be regarded as arrays of quantum dots in a matrix of a wider-gap material.

According to the data obtained by observing the diffraction pattern of high-energy electrons directly during the growth process, the critical thickness corresponding to island growth onset is 11 Å, and within the measurement error (± 0.5 Å) it does not depend on x, the molar fraction of AlAs in the matrix material.

Figure 1a shows the PL spectra of structures which contain InGaAs VAQD arrays and which differ by the molar fraction of AlAs in the $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ ($x = 0$, 0.15, 0.3) matrix. Photoluminescence was excited with an Ar$^+$ laser (100 W/cm^2) and detected with the aid of a cooled Ge photodiode. The detection temperature was equal to 10 K. The observed PL line is due to recombination of carriers localized in VAQDs.[2] Increasing the AlAs content in the matrix results in a small broadening of the line and a shift in the short-wavelength direction.

The PL excitation spectra for the cases $x = 0$ and 0.3 in the energy range above the VAQD radiation are shown in Fig. 1b. In Ref. 7 it was shown for the case of a GaAs matrix that the observed luminescence is associated with recombination in the wetting layer. The characteristic doublet shape of the ($WL1$, $WL2$) line could be due to the coexistence of

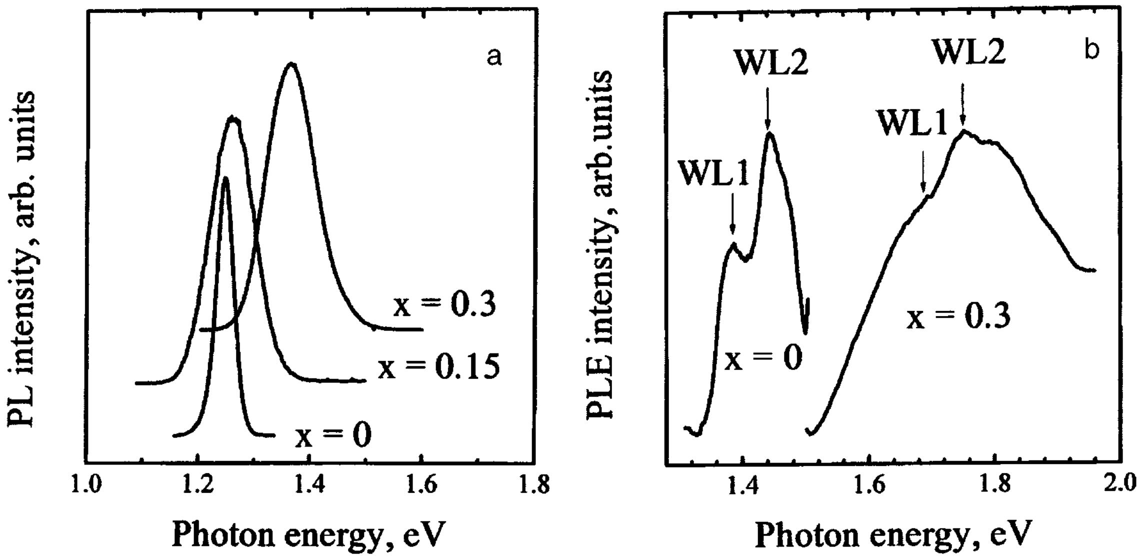

FIG. 1. Photoluminescence (a) and PL excitation (b) spectra at 10 K for structures containing arrays of InGaAs VAQDs in a $Al_xGa_{1-x}As$ matrix. The numbers on the curves denote the molar fraction of AlAs in the matrix material.

sections of a two-dimensional coating of InGaAs which differ somewhat in thickness, similarly to the case of InGaAs VAQDs in a GaAs matrix.

The positions of the luminescence maxima for radiation due to VAQDs and the wetting layers ($WL1$) as well as the edge luminescence of the matrix for structures containing arrays of $InGaAs/Al_xGa_{1-x}As$ ($x=0$ and 0.3) VAQDs are summarized in Table I. As one can see, when the band gap of the matrix increases by 370 meV, the short-wavelength shift of the VAQD line is only 120 meV. At the same time, the shift of the wetting layer line exceeds 300 meV. Therefore, increasing the Al content in the matrix increases the energy separation between the PL peak of the quantum dots and the peaks of the wetting layer and matrix.

The experimental results obtained can be interpreted as follows: Increasing the AlAs content in the matrix increases the energy barriers for electrons (holes) localized both in the wetting layer and in the quantum dots. However, as a result of the lower carrier localization energy and therefore the greater penetration of the wave functions into the barriers in the wetting layer than in the QDs, the perturbation of the energy levels under the action of a change in the height of the barriers is greater in the wetting layer. A similar behavior has been observed when arrays of InAs quantum dots were placed into an external AlGaAs/GaAs quantum well.[8] Although the luminescence data do not make it possible to distinguish unequivocally the contributions of the perturbation of the electron and hole levels, for carriers of both types the energy separation between the level in a QD and in a wetting layer and between a level in a QD and the band edge of the matrix apparently increases. This in turn results in lower filling of the states of the wetting layer and matrix with carriers at elevated temperatures.

TABLE I. Energy position of the luminescence lines of structures containing $InGaAs/Al_xGa_{1-x}As$ ($x = 0$ and 0.3) VAQD arrays.

Matrix material	VAQD	Wetting layer Line position, eV	Matrix
GaAs	1.246	1.388	1.514
$Al_{0.3}Ga_{0.7}As$	1.364	1.689	1.883

The dependence of the width of the PL line of the sample containing an array of $InGaAs/Al_{0.15}Ga_{0.85}As$ VAQDs on the observation temperature is shown in Fig. 2. The dependence shown is nonmonotonic in nature: As the temperature increases in the range 10–150 K, the linewidth decreases, and an increase of the linewidth is observed at high temperatures. This behavior is uncharacteristic of the luminescence of bulk material or quantum wells, where a regular increase of the linewidth with increasing observation temperature is typically observed. The reasons for the observed

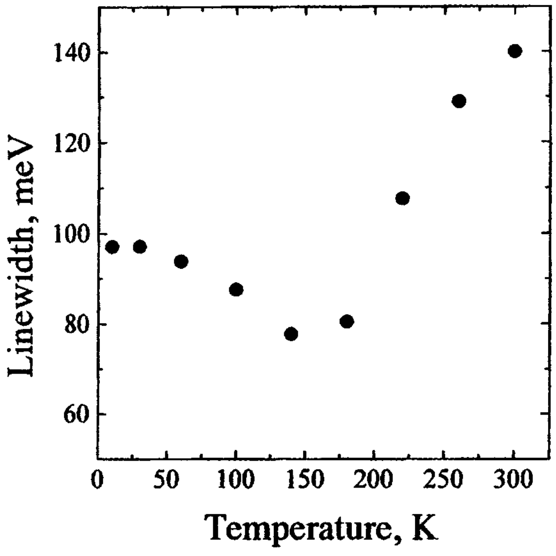

FIG. 2. Observation temperature dependence of the width of the PL line due to emission from $InGaAs/Al_{0.15}Ga_{0.85}As$ VAQDs.

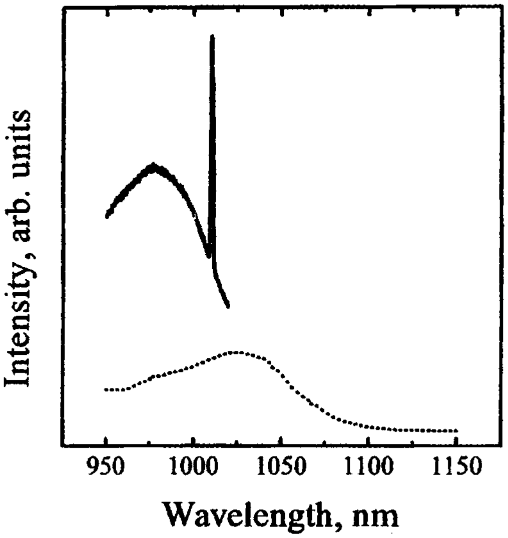

FIG. 3. Electroluminescence spectra near the threshold (65 A/cm^2) (solid line) and PL with a low excitation level (1 W/cm^2) (dotted line) of structures in InGaAs/Al$_{0.15}$Ga$_{0.85}$As VAQDs at 300 K.

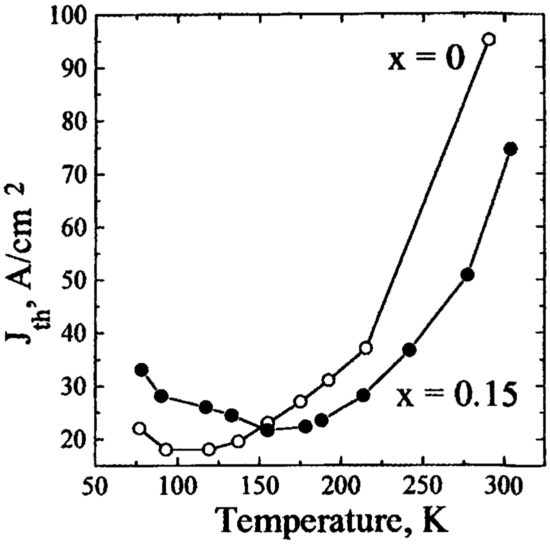

FIG. 4. Temperature dependence of the threshold current density for injection lasers based on an array of InGaAs/Al$_x$Ga$_{1-x}$As VAQDs. The numbers on the curves designate the molar fraction of AlAs in the matrix material.

anomalous dependence will be discussed below.

We investigated injection lasers with an active region based on InGaAs/AlGaAs VAQDs. The standard geometry of a AlGaAs/GaAs heterostructure laser with separate confinement was employed. An array of In$_{0.5}$Ga$_{0.5}$As/Al$_{0.15}$Ga$_{0.85}$As ($N=3$) VAQDs was placed at the center of an undoped 0.4-μm-thick Al$_x$Ga$_{1-x}$As ($x=0.5-0.15$) waveguide layer. The Al$_{0.5}$Ga$_{0.5}$As emitter layers were 1.5 μm thick. Si and Be were employed as the n- and p-type dopants. The laser diodes were formed by depositing and melting in AuGe/Ni/Au and AuZn/Ni/Au metallization to a n^+-GaAs substrate and a p^+-GaAs contact layer, respectively. Lasers with four sheared faces were employed. The measurements were conducted at temperatures in the range 77–300 K with 5-kHz pulsed pumping with 1-μs pulses.

The electroluminescence spectrum of the laser structure based on InGaAs/AlGaAs VAQDs near threshold at 300 K is presented in Fig. 3. The PL spectrum (300 K) with a low excitation level is also presented. The lasing line lies within the PL band, due to recombination in the QDs, near the maximum of the PL. Therefore, generation occurs through the ground state of the VAQDs.

The temperature dependence of the threshold current density is shown in Fig. 4. The corresponding data for a laser based on an array of InGaAs VAQDs in a GaAs matrix are shown for comparison. As one can see from Fig. 4, the values of J_{th} in the high-temperature range ($\geqslant$150 K) for a laser based on InGaAs/AlGaAs VAQDs are appreciably lower than in the case of a GaAs matrix. At room temperature the threshold current density is 63 A/cm^2, which, to the best of our knowledge, is the lowest value for all types of lasers with an active region with dimension less than 2 (quantum wires or quantum dots). We assume that the decrease of J_{th} in a laser employing a QD array in AlGaAs, as compared with the case of QDs in GaAs, is a direct consequence of the increase in the energy splitting between a QD level and a level of the wetting layer (matrix band edge).

A decrease of the threshold current density with increasing observation temperature is observed at low temperatures (77–150 K); i.e., the laser possesses a negative characteristic temperature ($T_0 \approx -200$ K). This behavior is unusual for lasers with four sheared faces, since in the case at hand it cannot be assumed that the effective strip width changes with temperature, as could happen in strip lasers.

The variance in the dots sizes results in a broadening of the aggregate density of states of the QD array. Assuming that the electron (hole) distribution over the states is described by a Fermi–Dirac function, an increase of the temperature at a given injection level results in increased filling of the higher-lying states corresponding to smaller dots as a result of the decrease in the filling of larger dots. Therefore, as the temperature increases, a higher injection current is required in order to maintain the required maximum gain. Therefore, the negative characteristic temperature cannot be explained on the basis of a quasiequilibrium carrier distribution.

In the model which we propose, the absence of a quasiequilibrium filling of the QD states at low temperatures is regarded as the reason for the appearance of the section of negative characteristic temperature. To achieve quasiequilibrium, there must exist a mechanism for carrier transport between dots. Tunneling or thermal transfer of carriers from QDs into the wetting layer or matrix followed by trapping in a different dot serves as such a mechanism.

According to transmission electron microscopy data,[9] the typical distances between neighboring islands exceed 300 Å. Therefore, the contribution of tunneling is negligible, and the condition for achieving a quasiequilibrium distribution is a short thermal transfer time compared with the radiative recombination time in QDs. The ejection of carriers is strongly suppressed at low temperatures as a result of the thermal activation nature of such processes. In Ref. 10 it was shown that there exists a limiting temperature (T_B) which separates the region of equilibrium ($T > T_B$) and nonequilibrium ($T < T_B$) filling of the QD states.

In the latter case, the filling of the QD states is determined not by the Fermi distribution but rather by the trapping probability of an electron (hole) in a dot of a definite size. Since approximately the same number of carriers is injected into dots of different sizes, the region of nonequilibrium filling is characterized by a wider carrier distribution over QD states than in the case where quasiequilibrium is achieved. In the latter case, the position of the Fermi quasilevel in the low-temperature limit is the upper limit of the filled states, while in the nonequilibrium case the higher-lying QD states are also filled.

Therefore, an identical gain will be achieved in the case of a quasiequilibrium distribution at a lower injection current density (J_{th}^{eq}) than in the temperature range where the Fermi distribution is of no consequence ($J_{th}^{\mathrm{noneq}} > J_{th}^{\mathrm{eq}}$). When the observation temperature passes through the limiting temperature T_B, the width of the PL line of a QD array must therefore decrease, and the threshold current density must decrease from J_{th}^{noneq} to J_{th}^{eq}. This is the reason for the observed negative characteristic temperature in a QD-based laser at low temperatures (Fig. 4), and it also makes it possible to explain the observed nonmonotonic temperature dependence of the width of the PL lines of an array of VAQDs (Fig. 2). We note the excellent agreement between the temperature range corresponding to a decrease of the linewidth and the region of negative characteristic temperature, in agreement with the proposed model.

An increase in the energy splitting between the electron (hole) level in a QD and the conduction (valence) band edge of the matrix decreases the probability of thermal transfer of carriers. Therefore, the limiting temperature T_B increases in the case where a wider gap material is used as the matrix material. We believe that this explains why the region of negative T_0 is more pronounced in the temperature dependence of the threshold current density of a laser based on InGaAs/AlGaAs VAQDs than in the case of InGaAs/GaAs VAQDs.

In summary, in the present work we investigated the characteristics of an injection laser with an active region based on an array of vertically aligned InGaAs quantum dots formed in an AlGaAs matrix. It was shown that using an AlGaAs solid solution as the matrix material makes it possible to increase substantially the energy separation of the QD states from the band edge of the matrix material and from states in the wetting layer as compared with the case of a GaAs matrix; this lowers the threshold current density at room temperature (63 A/cm^2). A model based on the assumption of a transition from nonequilibrium to equilibrium filling of the states of the quantum dots was proposed to explain the section of negative characteristic temperature observed at low temperatures.

This work was supported by the INTAS program (Grant 94-1028) and the Russian Fund for Fundamental Research (Grant 96-02-1784).

[1] Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).

[2] V. M. Ustinov, A. Yu. Egorov, A. R. Kovsh, A. E. Zhukov, N. N. Ledentsov, M. V. Maksimov, A. F. Tsatsul'nikov, N. Yu. Gordeev, S. V. Zaitsev, Yu. M. Shernyakov, N. A. Bert, P. S. Kop'ev, Zh. I. Alferov, N. N. Ledentsov, J. Behrer, D. Bimberg, A. O. Kosogov, P. Werner, and U. Gosele, in *Proc. 9th Int. Conf. on MBE*, (Aug. 5–9, 1996, Malibu, CA, USA).

[3] N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. Schmidt, R. Heitz, J. Bohrer, D. Bimberg, V. M. Ustinov, V. A. Shehukin, A. Yu, Egorov, A. E. Zhukov, S. V. Zaitsev, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, P. Werner, U. Gosele, and I. Heidenreich, Solid–Status Electron. **40**, 785 (1996).

[4] P. M. Petroff and S. P. DenBaars, Superlat. Microstruct. **15**, 15 (1994).

[5] J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre, and O. Vatel, Appl. Phys. Lett. **64**, 196 (1994).

[6] S. V. Zaitsev, N. Yu. Gordeev, Yu. M. Sherniakov, V. M. Ustinov, A. E. Zhukov, A. Yu. Egorov, M. V. Mazimov, P. S. Kop'ev, Zh. I. Alferov, N. N. Ledentsov, N. Kirstaedter, and D. Bimberg, in *Proc. 9th Int. Conf. on Superlattices, Microstructures and Microdevices* (14–19 July, 1996, Liege, Belgium).

[7] M. V. Maksimov, A. Yu. Egorov, A. E. Zhukov, N. N. Ledentsov, V. M. Ustinov, A. F. Tsatsul'nikov, A. O. Kosogov, P. S. Kop'ev, and Zh. I. Alferov, in *Abstracts of Reports at the 2nd All-Russia Conference on Semiconductor Physics* (26 Feb.—1 March 1996, Zelenogorsk), Vol. 1, p. 101.

[8] A. R. Kovsh, M. A. Odnoblyudov, A. Yu. Egorov, A. E. Zhukov, V. M. Ustinov, A. F. Tsatsul'nikov, M. V. Maksimov, N. N. Ledentsov, and P. S. Kop'ev in *Abstracts Int. Symp. ''Nanostructures: Physics and Technology''* (St. Petersburg, 1996) p. 286.

[9] S. Ruvimov, P. Werner, K. Scheerschmidt, U. Gosele, J. Heydenreich, U. Richter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, P. S. Kop'ev, and Zh. I. Alferov, Phys. Rev. B **51**, 14 766 (1995).

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

Translated by M. E. Alferieff

Neutron-activation analysis of the impurity composition of gallium arsenide based semiconductor structures

A. G. Dutov, V. A. Komar, S. V. Shiryaev, and L. A. Smakhtin

Institute of Solid-State and Semiconductor Physics, Bellorussian Academy of Sciences,
220072 Minsk, Belarus

(Submitted March 12, 1996; accepted for publication May 22, 1996)
Fiz. Tekh. Poluprovodn. **31**, 488–491 (April 1997)

The impurity composition of gallium arsenide slabs has been investigated using the highly sensitive method of neutron-activation analysis. The content of Se, Cr, Ag, Fe, Zn, Co, and Sb in the slabs as well as variations of the content of these impurities for different batches of slabs were determined. It is shown that neutron-activation analysis can be used, together with layerwise chemical etching, to study the volume distribution of impurity elements in gallium-arsenide-based semiconductor structures. © *1997 American Institute of Physics.* [S1063-7826(97)02204-7]

1. INTRODUCTION

High purity ($10^{15}-10^{14}$ atoms/cm^3) of the constituent layers and films of a semiconductor device structure,[1,2] which is very difficult to ensure considering the processes which occur during the growth and processing of gallium arsenide wafers,[3–5] takes an important place among the basic requirements which gallium-arsenide-based semiconductor structures must meet, such as, minimum content of different defects, reproducibility of the carrier density distribution along the growth axis, and other requirements. At the same time, information about extraneous impurities (type, density, character of distribution) is important for understanding the physical–chemical processes occurring in the semiconductor structures and for the fabrication technology of devices based on them.

In this study we looked at improving the methods for analyzing gallium arsenide based semiconductor structures. Using the methods of neutron-activation analysis (NAA) which have been developed, the impurity composition was investigated for several batches of gallium arsenide wafers, as well as the impurity distribution in semiconductors formed on these wafers.

2. SAMPLES AND METHOD OF ANALYSIS

In purely instrumental neutron-activation analysis of gallium-arsenide-based samples, substantial problems appear as a result of the high, specific, post-irradiation activity of the samples. This activity is due to the long-lived radioisotopes formed from matrix elements (Table I), which necessitates a long post-irradiation holding period for the samples and makes it difficult to determine many elements which have short-lived radioisotopes. For this reason, as a rule, different methods of radiochemical post-irradiation separation of the analyzed elements are used in the analysis of gallium-arsenide-based samples.[6,7]

In our work we employed simpler procedures to perform the analysis that makes it possible to substantially decrease at the same time the noise due to the high activity of the elements. Commercial gallium arsenide wafers and semiconductor structures formed on these wafers under test production conditions served as the experimental objects.

The prepared gallium-arsenide samples were irradiated with a thermal-neutron flux $\Phi = 10^{13}\ \mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}$ in the vertical channel of a VVR-Ts nuclear reactor. The mass of the samples was 50–250 mg. The irradiation time was 100 h. After ‘‘cooling’’ for 30–40 days, the samples were decomposed in a mixture of nitric and hydrofluoric acids with sublimation of arsenic fluoride in order to decrease the activity of $^{74}A_5$, the most long-lived matrix radioisotope; this made it possible to increase substantially the sensitivity of the determination of the elements as compared with the purely instrumental method of NAA.

A DGDK-160-type semiconductor Ge(Li) detector, combined with a LR4900V multichannel pulse-height analyzer, was used to measure the γ-ray spectra of the samples.

In the investigation of the semiconductor structures, the part of the gallium arsenide wafer (~ 1 cm^2) with structure deposited over the entire surface area was irradiated twice. The first irradiation (~ 20 h) is required in order to activate impurities which have long-lived radioisotopes; the second irradiation (1 min) was performed immediately before the measurements in order to activate elements with short-lived radioisotopes as well as matrix elements, whose radioactivity was used to monitor the thickness of the layers being etched off.

A method was developed for gradual removal of thin, plane-parallel layers on an irradiated gallium-arsenide slab, by chemical etching. The etchant consisted of the mixture $NH_4OH-H_2O_2-H_2O$, with polishing properties, in different ratios, depending on the required rate of etching of the sample and the thickness of the layers removed. The technique used to remove the layers was the same as the technique described in Ref. 8 for etching silicon wafers. The thickness of the layers removed was monitored according to the activity of ^{72}Ga and ^{76}As in the etched-off layer and a standard layer whose weight was known accurately. A small weighed quantity (~ 0.2 mg) of single-crystalline gallium arsenide, irradiated together with the sample, was used as a standard.

TABLE I. Principal radioisotopes contributing to the activity of GaAs irradiated with reactor neutrons.

Nuclear reaction and production of radioscopes	Half-life	Activation cross section,[b]	Specific activity after irradiation, Bq/mg
^{69}Ga$(n,\gamma)^{70}$Ga	20 min	1.4	1.85×10^{8}
^{71}Ga$(n,\gamma)^{72}$Ga	141 h	5.0	1.52×10^{8}
^{69}Ga$(n,2n)^{68}$Ga	68 min	0.55	2.8×10^{4}
^{69}Ga$(n,p)^{69}$Zn	14.0 h	0.024	1.1×10^{2}
^{71}Ga$(n,p)^{71}$Zn	3.9 h	0.05	2.6×10^{1}
^{69}Ga$(n,\alpha)^{66}$Cu	5.1 min	0.105	4.0×10^{3}
^{71}Ga$(n,\alpha)^{68}$Cu	30 s	–	–
^{75}As$(n,\gamma)^{76}$As	26.5 h	4.5	2.0×10^{8}
^{75}As$(n,2n)^{74}$As	18 days	0.55	2.8×10^{4}
^{75}As$(n,p)^{75}$Ge	82 min	0.118	5.0×10^{1}
^{75}As$(n,\alpha)^{72}$Ga	14.1 h	0.123	5.2×10^{1}

Note: Irradiation time 12 h, dose $\Phi=2.5\times10^{13}$ cm$^{-3}\cdot$s^{1}.

3. RESULTS OF ANALYSIS AND DISCUSSION

The results of an analysis of four different batches of single-crystalline gallium arsenide wafers are presented in Table II. In determining zinc and selenium, allowance is made for the fact that when the samples are irradiated in a nuclear reactor, the radioisotopes ^{65}Zn and ^{75}Se, which were used to determine the zinc and selenium, formed not only from the impurity elements themselves, but also from the matrix elements (As, Ga). Specifically, ^{75}Se can form from arsenic along two possible paths:

$$^{75}\mathrm{As}(n,2n)^{74}\mathrm{As}\xrightarrow{\beta^-}74\mathrm{Se}(n,\gamma)^{75}\mathrm{Se}$$

and

$$^{75}\mathrm{As}(n,\gamma)^{76}\mathrm{As}\xrightarrow{\beta^-}76\mathrm{Se}(n,2n)^{75}\mathrm{Se}.$$

For the conditions of our analysis, the quantity of the isotope ^{75}Se formed from arsenic was comparable to the amount of ^{75}Se formed from impurity selenium in the reaction

$$^{74}\mathrm{Se}(n,\gamma)^{75}\mathrm{Se},$$

so that the selenium concentration was determined relative to a sample which had the lowest selenium content.

The contribution of the isotope ^{65}Zn formed from matrix gallium in the reaction

$$^{69}\mathrm{Ga}(n,\gamma)^{66}\mathrm{Cu}\xrightarrow{\beta^-}{}^{66}\mathrm{Zn}(n,2n)^{65}\mathrm{Zn}$$

to the impurity zinc concentration was, according to our estimates, very small and did not exceed 0.5%.

Among the elements found in the gallium arsenide wafers (Table II), the highest content was that of the iron, zinc, and chromium impurity metals. The source of these impurities is probably the technological equipment employed for growing and processing single-crystalline gallium arsenide. The negligible batch-to-batch variance of the content of these impurities is within an order of magnitude. A larger variance was observed in the selenium, antimony, and silver contents. Differences in the concentration of these elements for different batches of wafers are also clearly observed. Of all experimental samples, the purest samples were those from the batch ''C'', in which not only the content of the impurity elements was lowest, but the sample-to-sample variance in the content of impurities was smallest.

Neutron-activation analysis combined with layerwise chemical etching of irradiated samples provide greater pos-

TABLE II. Results of neutron-activation analysis of single-crystalline gallium arsenide wafers.

	Elemental content, at./cm^3						
Sample	Se ($\times10^{14}$)	Cr ($\times10^{15}$)	Ag ($\times10^{13}$)	Fe ($\times10^{16}$)	Zn ($\times10^{15}$)	Co ($\times10^{13}$)	Sb ($\times10^{13}$)
S1	0.97	0.65	0.43	0.90	0.83	0.25	0.92
S6	0.81	0.35	0.76	1.34	0.47	0.20	1.81
S81	0.99	0.26	0.64	1.70	0.70	0.20	2.40
S82	0.95	0.80	0.81	1.36	0.80	0.41	2.65
V18	0.91	1.16	0.68	0.81	1.11	0.33	2.46
V61	1.22	0.51	1.34	1.82	1.59	0.32	3.22
V95	6.05	0.93	1.19	2.51	3.92	0.57	3.36
V121	0.96	0.66	0.53	0.71	0.59	0.24	2.47
V221	7.48	0.63	0.32	1.24	1.49	0.26	2.38
V251	3.64	1.84	1.26	2.04	1.31	0.52	5.84
V261	4.61	1.81	0.48	2.25	1.07	0.72	8.11
B1	6.33	0.33	0.44	1.43	1.44	0.29	1.88
B2	3.15	0.13	0.65	1.52	0.81	0.25	1.23
B31	8.75	0.34	0.46	2.53	1.11	1.12	2.90
B41	6.58	0.11	0.46	1.02	0.72	0.18	0.93
B61	1.66	0.83	0.41	0.95	0.38	0.22	1.25
M32	0.89	3.10	0.12	1.63	0.84	0.33	1.81
M42	1.25	3.15	2.18	1.12	0.94	0.56	16.50
M52	0.64	0.67	2.48	1.43	1.49	0.34	20.05
M62	0.74	1.68	1.14	1.48	4.30	0.36	12.30

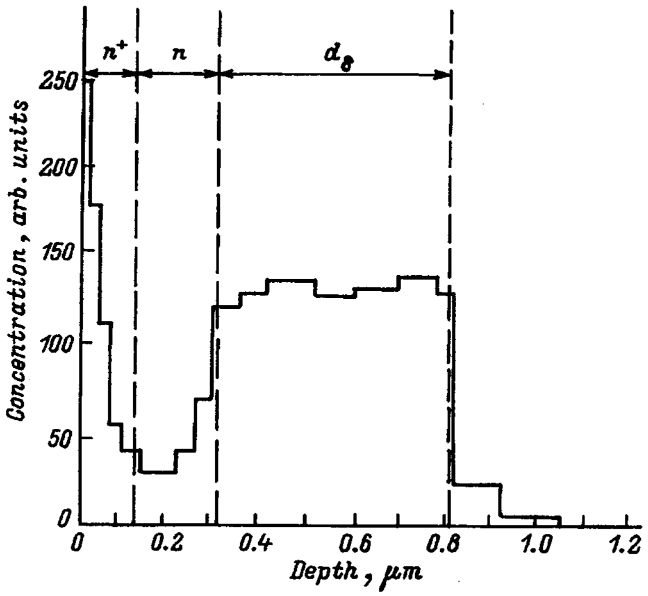

FIG. 1. Chromium distribution in a SAG-2BK structure. The distribution was obtained by the NAA method combined with layerwise chemical etching of the sample.

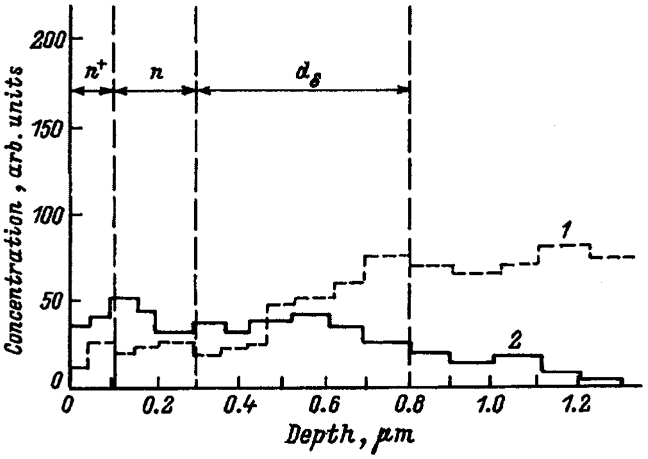

FIG. 2. Copper and sodium distributions in the SAG-2BK structure. The distribution was obtained by the NAA method combined with layerwise chemical etching of the sample. *1* — Copper, *2* — sodium.

sibilities for monitoring the impurity composition. Figure 1 shows the chromium distribution in a SAG-2BK structure, whose buffer layer (d_δ) was doped with chromium. As one can see from the distribution obtained in performing the successive operations, chromium diffuses from the buffer layer into the working layers of the semiconductor structure, and its density at the surface of the n^+ layer is very low and comparable to the density in the buffer layer.

The distribution of sodium and copper metal impurities (Fig. 2), which were not specially introduced into the structure but rather enter it in the course of different technological operations as uncontrollable impurities, was also investigated. Having higher mobility and electrical activity, they can appreciably affect the parameters of the semiconductor structures. The manner in which they are distributed is therefore of particular interest.

As one can see from the distribution of these elements over the volume of the SAG-2BK structure in Fig. 2, the copper content is typically higher in a single-crystalline gallium-arsenide substrate and the total copper content in the semiconductor structure is determined by the content of this element in the initial gallium-arsenide slabs. For sodium, the directly opposite picture is observed. Most of the sodium enters the semiconductor structure when its working layers are formed.

In summary, the results presented above show that the methods of neutron-activation analysis proposed above, first of all, make it possible to reliably monitor the impurity composition of gallium-arsenide slabs according to an entire series of elements at the level $10^{12}-10^{16}$ atoms/cm^3 and, second, they make it possible to obtain a detailed picture of the impurity distribution in the device. These methods also allow addtional monitoring and further improvement of the semiconductor device fabrication technology.

[1] B. V. Luft, V. A. Perevodchikov, and L. N. Voslikova, in *Physicochemical Methods of Treatment of Semiconductor Surfaces* [in Russian], Radio and svyaz', Moscow (1982), p. 136.
[2] T. G. Yugova, A. V. Govorkov, and A. S. Bruk, Élektron. Tekhnika, Materialy **2**, 24 (1984).
[3] D. Brice, J. Mat. Sci. **2**, 131 (1967).
[4] W. Kern, J. Electrochem. Soc. **109**, 200 (1962).
[5] E. Bujdoso, in *Proc. Conf. Appl. Phys. Chem. Methods*, Budapest (1963), Vol. 2, p. 262.
[6] K. Kudo, N. Hishinuma, J. Radioanal. Chem. **3**, 369 (1969).
[7] K. Kudo, N. Hishinuma, J. Radioanal. Chem. **5**, 331 (1970).
[8] A. G. Dugov, V. A. Komar, S. V. Shiruaev, K. E. Lobanov, V. V. Zabrodskaya, and L. P. Mozheĭko, Zh. Anal. Khim. **49**, 76 (1994).

Translated by M. E. Alferieff

Effect of laser radiation on the electronic density of states of an ⟨insulator gallium arsenide⟩ interface

L. N. Vozmilova

Scientific-Research Institute of Semiconductor Devices, 634045 Tomsk, Russia

V. I. Gaman, V. M. Kalygina, and A. V. Panin

V. D. Kuznetsov Siberian Physicotechnical Institute at Tomsk State University, 634050 Tomsk, Russia

T. P. Smirnova

Institute of Inorganic Chemistry, Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted April 17, 1996; accepted for publication June 4, 1996)
Fiz. Tekh. Poluprovodn. **31**, 492–497 (April 1997)

The effect of annealing with pulsed laser radiation with wavelengths of 0.69 and 308 μm on the capacitance-voltage and conductance-voltage characteristics and the density of surface states at the insulator–(n, p)-GaAs interface as a function of the radiation energy density has been investigated. © *1997 American Institute of Physics.* [S1063-7826(97)02304-1]

The high density of surface electronic states SESs at a GaAs-insulator interface gives rise to a weak dependence of the surface potential φ_s on the gate voltage U and prevents realizing a MIS transistor based on gallium arsenide. Attempts to lower the SES density reduce to treating GaAs in passivating solutions,[1–5] treatment with a rf plasma,[6,7] deposition of films with different composition on the semiconductor surface,[8–10] and matching the insulator material.[11,12] Passivating treatments make it possible to decrease the SES density and, in the opinion of the authors, to ''detach'' the Fermi level at the GaAs surface. However, the effect is reversible: After some time the surface returns to the initial state[2,13] and the characteristic features in the behavior of MIS structures for an alternating signal, which are due to the high density of SESs and the large set of relaxation times, are observed again.

Pulsed laser annealing (PLA),[14] combined with matching of the insulator, could be an effective method for controlling the GaAs-insulator interface. This paper discusses the results of an investigation of the effect of PLA on the capacitance-voltage characteristics (FVCs) and conductance-voltage characteristics (SVCs) of metal-insulator-GaAs structures.

1. EXPERIMENTAL PROCEDURE

The samples were fabricated on the basis of epitaxial layers of electronic (doping level $N_d = 1\times 10^{15} - 2\times 10^{16}$ cm^{-3}) and hole (doping level $N_a = (1-2)\times 10^{16}$ cm^{-3}) gallium arsenide, which were grown on n^+ and p^+-type substrates, respectively, and oriented in the (100) plane. Films prepared by the plasma chemical method—silicon oxynitride ($Si_xN_yO_z$), silicon dioxide (SiO_2), boron nitride (BN), double layers ($BN+Si_xN_yO_z$), and also anodic oxide films—were employed as the insulator. Before the insulator was deposited by the plasma-chemical method, after degreasing, the semiconductor substrates were etched in NH_3 plasma for 2–3 min. Insulator films of thickness $d = 115 - 190$ nm were grown in a tunnel-type reactor[15] with substrate temperature $T_s = 200$ °C (BN) and 400 °C ($Si_xN_yO_z$, SiO_2). The anodic oxide was deposited in a static-voltage regime. An electrode to the insulator (controlling electrode) with fixed area $S = 2\times 10^{-3}$ cm^2 was obtained after PLA by thermal vaporization of NiCr through appropriate templates.

The wavelengths of the laser radiation $\lambda = 0.69$ μm (pulse duration $\tau = 1\times 10^{-3}$ s) and $\lambda = 0.308$ μm ($\tau = 3\times 10^{-8}$ s) were chosen from the condition that the insulator be transparent. This made it possible to treat the semiconductor-insulator interface directly. The pulsed laser annealing was conducted on the insulator side before the controlling electrode was deposited. The radiation energy density W was varied from 0 to 15–20 J/cm^2 for radiation with $\lambda = 0.69$ μm and up to 100–150 mJ/cm^2 in the case $\lambda = 0.308$ μm.

The capacitance-voltage and conductance-voltage characteristics were measured in the frequency range $f = 10^3 - 10^6$ Hz at room temperature. The error in determining the capacitance C at a frequency of 1 MHz did not exceed ± 0.1 pF in the bias voltage range $U = \pm 5$ V and ± 0.35 pF for higher values of U. The active component of the total conductance G was determined to within ± 1 μS.

2. EXPERIMENTAL DATA

The capacitance-voltage and conductance-voltage characteristics of MIS structures exhibit a strong frequency dependence up to 10^6 Hz, irrespective of the type of conduction in GaAs and the insulator material. A somewhat weaker frequency dependence of the FVCs and SVCs is observed for samples with BN. However, the boron nitride layers which are synthesized at a low temperature possess inadequate electrical strength. Because of this circumstance, to perform measurements up to voltages of 40–50 V, a silicon oxynitride film was deposited on top of BN. For all MIS structures investigated, the maximum capacitance is less than the design capacitance C_d of the insulator (see Table I). When the same insulator is used, the modulation of the capacitance and

TABLE I. Effect of laser radiation on the density of electronic states at the interface.

Sample No.	Structure	C_d, pF	$N_{a,d}$ 10^{15} cm^{-3}	N_{tt}, 10^{12} cm$^{-2}\cdot$eV^{-1} before PLA	N_{tt}, 10^{12} cm$^{-2}\cdot$eV^{-1} after PLA	N_t, 10^{11} cm$^{-2}\cdot$eV^{-1} before PLA	N_t, 10^{11} cm$^{-2}\cdot$eV^{-1} after PLA	W, J/cm^2
1	M$-$Si$_x$N$_y$O$_z$$-$$p$-GaAs	75	3.0	30	2.8	4.8	5.4	10
2	M$-$Si$_x$N$_y$O$_z$$-BN-$$n$-GaAs	76	1.5	62	39	1.4	5.7	8
3	M$-$anodic oxide$-$$p$-GaAs	81	2.0	7.5	4.3	4.8	7.7	8
					2.6		7.4	10
					6.9		6.4	12

Note: M—metal.

the maximum conductance are observed at negative voltages for structures based on p-type GaAs and at positive voltages for samples based on n-type GaAs.

Irrespective of the type of conductivity in GaAs and the wavelength of the radiation, after pulsed laser annealing the FVCs and SVCs shift along the voltage axis, and the maximum conductance $G_{\max}$ and also the minimum capacitance $C_{\min}$ and maximum capacitance $C_{\max}$ of the MIS structures change. Figure 1 shows the FVCs and SVCs for a sample based on hole-type GaAs before (curve *1*) and after (curves *2–5*) the action of laser radiation with $\lambda = 0.69$ μm. For $W \leqslant 6-8$ J/cm^2 the FVCs and SVCs shift to lower negative voltages, the slope of the capacitance-voltage characteristics increases on the capacitance modulation section, and $C_{\max}$ increases (curves *2* and *3*). At the same time, the maximum on the conductance-voltage characteristic becomes narrower, and its position on the voltage axis corresponds to lower values of $C_{\max}$ (curves *2* and *3*). As W increases, the reverse effect is observed: The FVCs and SVCs shift to higher negative voltages, $C_{\max}$ decreases, and dC/dU decreases (Fig. 1, curves *4* and *5*).

After PLA with $\lambda = 0.69$ μm and $W \leqslant 8$ J/cm^2, the dispersion of the FVC also decreases (Fig. 2). Similar data were obtained for MIS structures based on electronic gallium arsenide. The only difference is that after laser annealing with $W \leqslant 8$ J/cm^2 the FVCs and SVCs shift to lower positive voltages.

Decreasing the wavelength and duration of the laser pulse has no effect on the changes in the electric characteristics of the experimental samples. Once again, there exists a critical energy density W_c, which determines the effect of the laser radiation. After laser annealing with $\lambda = 0.308$ μm and $W < W_c \simeq 9.1$ mJ/cm^2, the minimum and maximum capacitances of the MIS structures increase (Fig. 3, curves *1–3*) and the frequency dispersion of the FVCs and SVCs decreases. In the case $W > W_c$ laser annealing has the opposite effect (Fig. 3, curves *4* and *5*). As the density of the

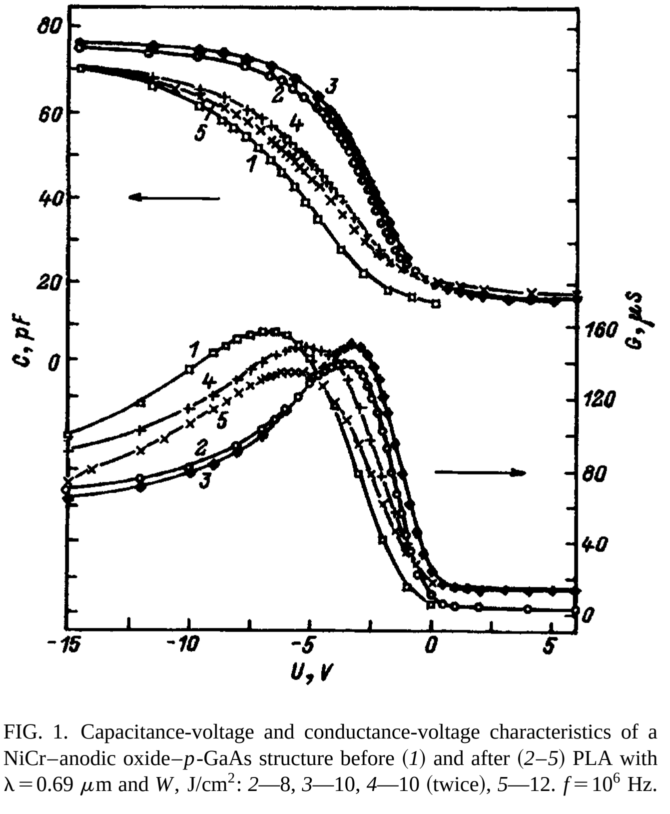

FIG. 1. Capacitance-voltage and conductance-voltage characteristics of a NiCr–anodic oxide–p-GaAs structure before (*1*) and after (*2–5*) PLA with $\lambda = 0.69$ μm and W, J/cm^2: *2*—8, *3*—10, *4*—10 (twice), *5*—12. $f = 10^6$ Hz.

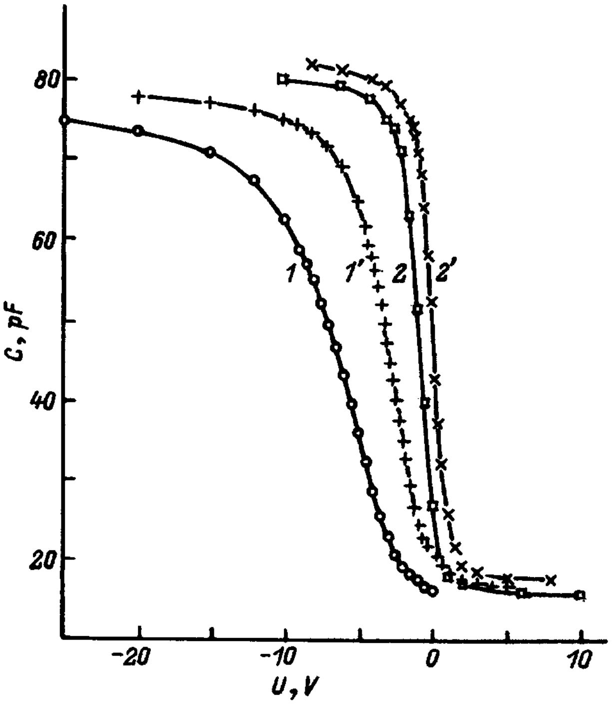

FIG. 2. Capacitance-voltage characteristics of NiCr–anodic oxide–p-GaAs structure before (*1*, *2*) and after (*1′*, *2′*) PLA with $W = 10$ J/cm^2. f, Hz: *1*, *1′*—10^6; *2*, *2′*—10^3.

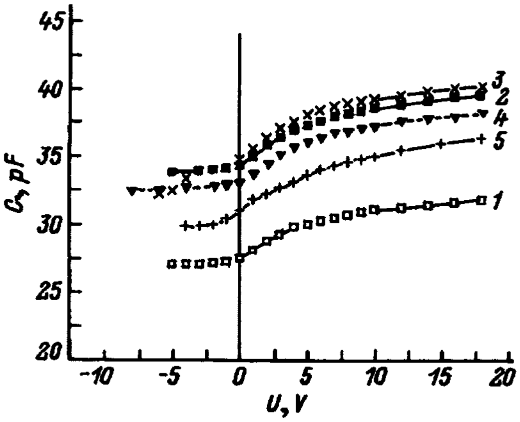

FIG. 3. Capacitance-voltage characteristics of NiCr-anodic oxide-*n*-GaAs structure before (*1*) and after (*2–5*) PLA with λ = 0.308 μm and *W*, J/cm^2: *2*—4.6, *3*—9.1, *4*—14.6, *5*—27.4. $f = 10^6$ Hz.

energy incident on the sample increases further, an inversion of the FVCs and SVCs is observed. For *n*-GaAs samples with a double insulator, the maximum conductance and modulation of the capacitance after laser annealing are observed for negative potentials on the controlling electrode (Fig. 4, curve *3* and *4*). A similar inversion of the FVCs and SVCs after the action of radiation with $\lambda = 0.69$ μm and $\tau = 1 \times 10^{-3}$ s on the $Si_xN_yO_z$-*n*-GaAs interface for $W > 15-20$ J/cm^2 was observed in Ref. 14.

The change in the electrical characteristics produced in GaAs-based MIS structures by PLA is not a short-time effect. Measurements performed on samples subjected to laser action immediately after annealing exhibit approximately the same results a year and three years later.

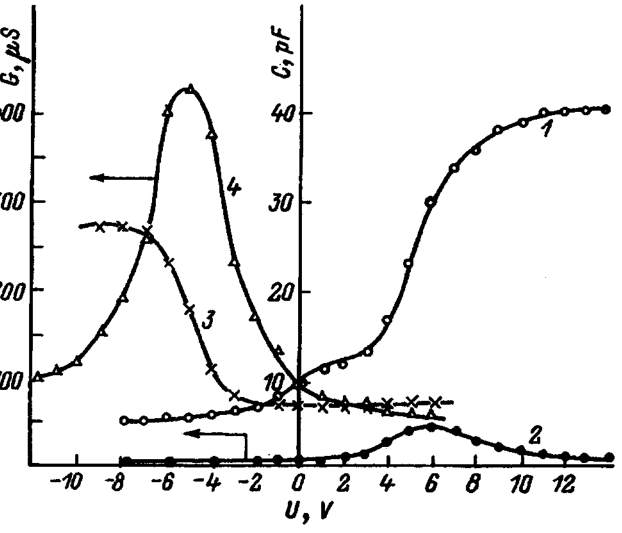

FIG. 4. Capacitance-voltage and conductance-voltage characteristics of a NiCr–$Si_xN_yO_z$–BN–*n*-GaAs structure before (*1, 2*) and after (*3, 4*) PLA with λ = 0.308 μm and $W > W_c$. $f = 10^6$ Hz.

3. DISCUSSION OF THE EXPERIMENTAL DATA

In analyzing the experimental data, we note that $C_{\max}$ does not equal to C_d. This indicates that the accumulation regime is not realized in the voltage interval studied, and the appearance of a section of capacitance modulation and a conductance maximum is due to charge exchange on the fast surface states, which have enough time to respond to the measuring signal at the given frequency.[16,17] In this case the maximum value of the capacitance can be less than the capacitance of the insulator. The quantity $C_{\max}$ corresponds to the low-frequency and $C_{\min}$ corresponds to the high-frequency capacitance of the MIS structure. As the voltage increases ($U > |U_0|$), a transition occurs from the high-frequency value of the capacitance of the MIS structure

$$C_{hf} = \frac{C_d C_{dl}}{C_d + C_{dl}} = C_{\min} \tag{1}$$

to the low-frequency value

$$C_{lf} = \frac{C_d (C_{dl} + C_t)}{C_d + C_{dl} + C_t} = C_{\max}, \tag{2}$$

where U_0 is the voltage corresponding to the onset of the capacitance modulation section,

$$C_t = S e^2 N_t (F_s) \tag{3}$$

is the low-frequency differential capacitance limit due to charge exchange on fast surface states whose energy equals the Fermi energy (F_s) at the semiconductor surface, N_t is the energy density of fast SESs, e is the electron charge, and C_{dl} is the capacitance of the depleted layer.

Since at voltages $U < |U_0|$ the capacitance ($C_{\min}$) of MIS structures is virtually independent of the voltage (if there is no nonequilibrium depletion), it can be assumed that this section of the FVCs corresponds to strong depletion or weak inversion. The capacitance of the depleted layer is determined by the expression[17]

$$C_{dl} = S \sqrt{\frac{e \varepsilon \varepsilon_0 N_{d,a}}{2 |\varphi_s(U)|}}, \tag{4}$$

where ε is the relative permittivity of the semiconductor, ε_0 is the permittivity of free space, and $|\varphi_s(U)|$ is the absolute value of the surface potential at the voltage U.

The surface potential $\varphi_s(0)$ at $U = 0$ before and after the action of laser radiation can be estimated with the aid of Eqs. (1) and (4). It is well known that in GaAs the value of $\varphi_s(0)$ is determined by the total density of surface states N_{tt}, whose energy levels lie near the Fermi level at the surface (F_s)[17] and which can be conventionally divided into fast and slow levels. The calculation for NiCr–anodic oxide–*n*-GaAs structures (Fig. 3) was performed for the following values of the parameters: $S = 2 \times 10^{-3}$ cm^2, anodic oxide permittivity $\varepsilon_d = 8.5$, and $N_d = 5 \times 10^{15}$ cm^{-3}. Before laser annealing $\varphi_s(0) = 0.81$ V and it decreases to 0.67 V after the action of laser radiation with $W = 9.1$ mJ/cm^2. Annealing of the structures at $W = 27.4$ mJ/cm^2 increased $\varphi_s(0)$ to 0.97 V. In the calculations it was assumed that PLA does not change the charge carrier density in the surface region of GaAs.

The variation of the height of the potential barrier at the surface of gallium arsenide after PLA correlates with the decrease in the total density of surface states for low values of W and the increase in N_{tt} for $W > W_c$ (see Table I). The values of N_{tt} were determined by the method described in

Ref. 17. However, pulsed laser annealing can actually change not only $\varphi_s(0)$ but also the impurity density in the surface region of the semiconductor. This effect is especially important in the case of annealing of structures with $W > W_c$.

According to Eqs. (2) and (3), the increase and decrease in the maximum capacitance after PLA with $W < W_c$ and $W > W_c$ are explained, correspondingly, by the change in the density of fast surface states. The expression

$$(G_t/\omega)_{\max} \simeq 0.4 e^2 N_t(F_s) S, \tag{5}$$

where $(G_t/\omega)_{\max}$ is the maximum value of the conductance, due to charge transfer on the fast surface states and normalized to the frequency ω, was used to estimate N_t (see Table I). The values of G_t were calculated according to the formulas presented in Ref. 17. As follows from the table, the density N_t of fast surface states is higher after PLA in the case $W < W_c$ and lower in the case $W > W_c$.

The displacement of the FVCs and SVCs to $U = 0$ and also the increase in the slope of the capacitance-voltage characteristics (Fig. 1) as a result of PLA at $W < W_c$ is explained by the change in the surface potential φ_s due to the decrease in the total density N_{tt} of surface states.[18] This is confirmed by the estimates of N_{tt} presented in Table I. It follows from the data in the table that N_t is, on the average, approximately two orders of magnitude lower than the total density of SESs, i.e., the slow surface states make the main contribution to pinning of the Fermi level at the surface.

Consequently, as W increases, the density of fast SESs and the total density of surface states vary in phase opposition. This is probably attributable to their different nature. Since $N_t \ll N_{tt}$, the resulting action of the PLA on the electrical characteristics of MIS structures is determined mainly by the change in the slow surface states (with the exception of $C_{\max}$).

We shall use the electron-deformation-thermal (EDT) model to analyze the possible mechanism responsible for the change in the density of SESs under pulsed laser annealing. This model was proposed in Refs. 19 and 20 to explain the change in the density of defects in semiconductors under the action of laser radiation. This model assumes a multifactor action of PLA on the semiconductor surface: excitation of the electronic subsystem, increase in the temperature, and appearance of deformation as a result of thermal expansion of the surface region of the semiconductor and because of the photostriction effect. One or another factor could dominate, depending on the energy density W.

In our experiments, laser radiation energy densities which do not exceed the melting threshold of GaAs ($W < W_m$) were used. We take into account the fact that for the indicated wavelengths the energy of a laser photon is higher than the band gap in GaAs; this gives rise to interband transitions and results in a sharp growth of the density of nonequilibrium charge carriers in the absorption region.

For $W < W_c$ the heating and deformation of the semiconductor surface are apparently weaker than in the case of electronic excitation, which is mainly localized on biographic defects.[19,20] Their density is higher in a thin surface layer of the semiconductor and directly at the GaAs-insulator interface. The energy stored in the electronic subsystem relaxes through several processes, including the electron-phonon interaction. We assume that the energy released with recombination or trapping of nonequilibrium charge carriers is sufficient for multiplication of defects responsible for fast surface states N_t, whose energy density is comparatively low ($N_t \simeq 10^{11}$ cm$^{-2} \cdot$ eV).

The locally excited states of a defect which arise at the surface exist for a comparatively long time,[21] during which a rearrangement of the atomic structure of the primary defects is possible. As a result, the density of defects of the second type decreases. These defects play the role of slow surface states whose energy density exceeds that of SESs by two orders of magnitude, as is indicated by the decrease in the frequency dispersion of the electrical characteristics of the experimental structures (Fig. 2). Annihilation of defects of the second type for $W < W_c$ proceeds more rapidly than multiplication of defects of the first type. Ultimately, the total density of surface states decreases as W increases (see Table I, sample 3). Therefore, under PLA with $W < W_c$ two processes occur simultaneously: multiplication of defects responsible for N_t (their initial density is low) and annihilation of defects which give rise to slow surface states (their initial density is high). The effect of the initial density of defects in GaAs crystals on the effect of pulsed laser annealing was discussed in Ref. 22. An estimate of the power P_c consumed for $W = W_c$ shows that $P_c = 8.0 \times 10^3$ W$\cdot$s^{-1} in the case $\tau = 10^{-3}$ s and $P_c = 2.3 \times 10^5$ W$\cdot$s^{-1} for $\tau = 4 \times 10^{-8}$ s. Therefore, increasing the pulse duration for $W < W_c$ increases the efficiency of annihilation of the defects. In Ref. 23 a similar effect is explained by the prolonged existence of states with a high density of nonequilibrium charge carriers.

Comparing the Auger spectra of the insulator-GaAs interface before and after laser action with $W < W_c$ (see Figs. 5a and 5b) does not show any appreciable changes in the elemental distribution profiles. The interface remains sharp after PLA with $W < W_c$. This indicates a weak interpenetration of the components of GaAs and the insulator. It can be assumed, therefore, that under laser annealing with $W < W_c$ the processes developing at the interface and on the gallium arsenide surface are of an athermal character.

Thermal effects and deformation play a larger role when $W > W_c$. According to the EDT model, the activation energy of defect formation decreases,[19,20] which results in a higher density of defects and a higher total density of surface states.

For $W > W_c$ thermal annealing of defects, which are responsible for fast surface states, probably occurs, and the density of such states decreases with increasing W. But since $N_t \ll N_{tt}$, a decrease of N_t does not change the character of the behavior of the total density of surface states with increasing laser radiation energy and N_{tt} increases for $W > W_c$. This gives rise to a shift of the FVCs and SVCs to higher voltages, decreases the slope of the FVCs, and increases the frequency dispersion of the electrical characteristics.

Quenching of the photoluminescence signal was observed in Ref. 21 for approximately the same values of W and higher values. The authors attribute this behavior to gen-

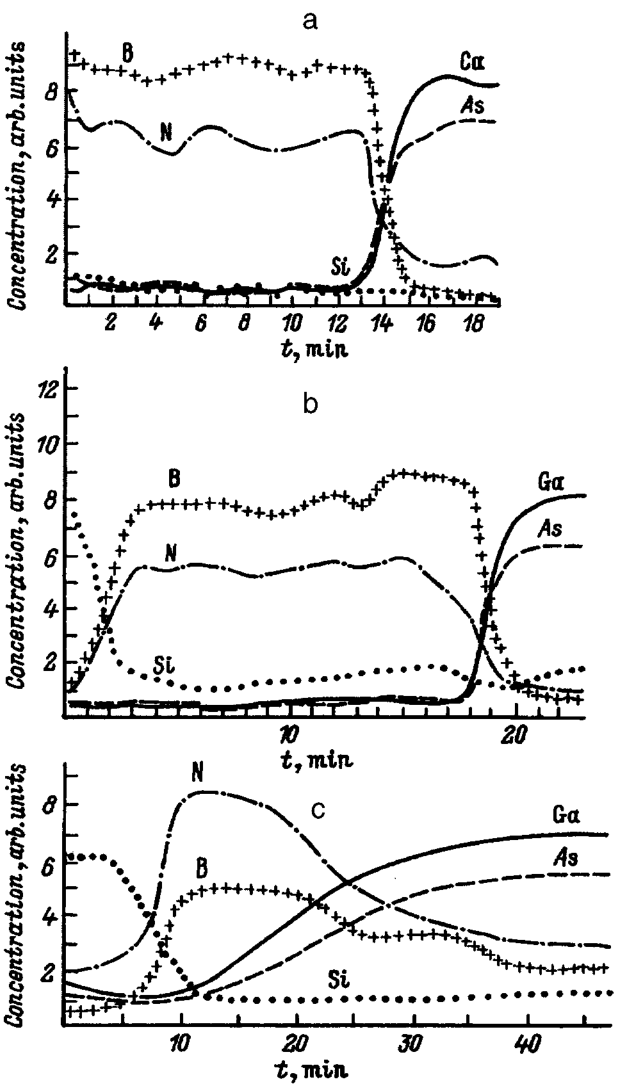

FIG. 5. Elemental distribution profiles at the $Si_xN_yO_z$–BN–GaAs interface before (a) and after PLA with $W<W_c$ (b) and $W>W_c$ (c).

eration of defects which are nonradiative-recombination centers.

The experimental results for PLA with $W=15-20$ J/cm^2 ($\lambda=0.69$ μm, $\tau=10^{-3}$ s) and W equal to several tens of mJ/cm^2 ($\lambda=0.308$ μm, $\tau=3\times10^{-8}$ s) are explained under the assumption that under these conditions doping of the surface region of GaAs with an acceptor impurity occurs as a result of high mass-transfer rates. As a result, the type of conductivity in the semiconductor changes. This change accounts for the observed course of the capacitance-voltage and conductance-voltage characteristics after ILA (Fig. 4). This assumption is confirmed by the change in the elemental distribution profiles at the insulator-GaAs interface (Fig. 5), which are substantially different from the corresponding data before laser annealing and after the action of laser radiation with $W<W_c$. Strong mixing of the components of the semiconductor and insulator is observed for large values of W. This results in an appreciable smearing of the interface. The surface layer of GaAs becomes enriched with boron, which plays the role of an acceptor impurity.[24] It should be noted that thus far we have observed the above-mentioned effect (change in the type of conductivity) only for n-GaAs–based MIS structures.

In summary, the electrical characteristics of MIS structures based on gallium arsenide can be effectively controlled up to inversion of the type of conductivity in the surface region of the semiconductor, by matching the conditions of PLA (wavelength of the laser radiation, pulse duration, and energy density). For $W<W_c$ the character of the changes in the FVCs and SVCs does not depend on the insulator material and is determined by the wavelength λ and the pulse duration τ of the laser radiation. An advantage of this method is that the post-PLA parameters of the characteristics remain unchanged when the samples are stored for several years under ambient atmosphere conditions.

[1] B. I. Bednyĭ, N. V. Baĭdus', T. V. Belich, and I. A. Karpovich, Fiz. Tekh. Poluprovodn. **26**, 1383 (1992) [Sov. Phys. Semicond. **26**, 775 (1992)].
[2] B. N. Bessolov, A. F. Ivankov, E. V. Konenkov, and N. V. Lebedev, Pis'ma Zh. Tekh. Fiz. **21**, 46 (1995) [*sic*].
[3] S. I. Kirillova, V. E. Primachenko, and V. A. Chernobaĭ, Poverkhnost', No. 12, 80 (1994).
[4] Berkovitz, V. N. Bessolov, T. V. L'vova, and B. V. Tsarenkov, Fiz. Tekh. Poluprovodn. **25**, 1406 (1991) [Sov. Phys. Semicond. **25**, 847 (1991)].
[5] M. S. Carpenter, M. R. Melloch, and T. E. Dangan, Appl. Phys. Lett. **53**, 60 (1988).
[6] S. Cassette, F. Plais, and J. Olver, Surf. Interf. Anal. **16**, 1 (1991).
[7] T. Sugino, T. Yamada, K. Matsuda, and J. Shiraafuju, Appl. Sur. Sci. **56–68**, pt A, 311 (1992).
[8] J. R. Waldrop and R. W. Grant, J. Vac. Sci. Technol. B **6**, 1015 (1988).
[9] K. Kjyanagy, S. Kasar, and H. Hasegava, Jpn. J. Appl. Phys. **32**, 502 (1993).
[10] S. A. Chambers and V. S. Sundaram, J. Vac. Sci. Technol. B **9**, 2256 (1991).
[11] *Multilayer thin-film structure*, Matsumoto Yosinari; nippon denki k. k. Patent 60–223134, Japan, Patent April, 19, 1984, No. 59–74957, published Nov. 7, 1985, MKI N 01 L21/314, C23C16/30.
[12] G. Yu. Bagratishvilli, R. N. Dzhanelidze, N. I. Kurdiana *et al.*, Élektron. Tekhn., Ser. 2. Poluprovodnikovye Pribory **67**, 31 (1972).
[13] M. Yoichi, W. Kazumi, and W. Yoshinori, Appl. Phys. Lett., **61**, 2993 (1992).
[14] V. P. Voronkov, V. M. Kalygina, S. Yu. Mulenkov, E. I. Oborina, E. G. Sal'man, and T. P. Smirnova, Fiz. Tekh. Poluprovodn. **26**, 1120 (1992) [Sov. Phys. Semicond. **26**, 626 (1992)].
[15] T. P. Smirnova, L. V. Khramova, and I. K. Yashkin, Neorg. Mater. **28**, 1414 (1992).
[16] V. N. Ovsyuk, *Electronic Processes in Semiconductors with Space Charge Regions* [in Russian], Novosibirsk (1984).
[17] V. I. Gaman, N. N. Ivanova, V. M. Kalugina, and E. B. Sudakova, Izv. Vyssh. Uchebn. Zaved., Fiz. **35**, 99 (1992).
[18] V. I. Gaman, V. M. Kalygina, A. V. Panin, and T. P. Smirnova, Poverkhnost', No. 5, 18 (1995).
[19] V. I. Emel'yanov and P. K. Pashkarov, Poverkhnost', No. 2, 77 (1990).
[20] P. K. Kashkarov and V. Yu. Timoshenko, Poverkhnost', No. 6, 5 (1995).
[21] A. V. Zoteev, P. K. Kashkarov, and V. F. Kiselev, Poverkhnost', No. 5, 97 (1993).
[22] I. S. Belichev, P. K. Kashkarov, and V. Yu. Timoshenko, Vest. MGU, Ser. 3, **30**, 77 (1989).
[23] H. Okigawa, T. Nakayama, K. Talayama, and N. Itoh, Commun. **49**, 347 (1984).
[24] N. Ainsbrugh and Y. Wissman [Eds.], *Gallium Arsenide in Semiconductors* [Russian trans.], Moscow (1988).

Translated by M. E. Alferieff

Radiative cooling under the conditions of magnetoconcentration

A. I. Liptuga, V. K. Malyutenko, and V. I. Pipa

Institute of Semiconductor Physics, Ukrainian National Academy of Sciences, 252650 Kiev, Ukraine

L. V. Levash

Institute of Physics, Ukrainian National Academy of Sciences, 252028 Kiev, Ukraine

(Submitted May 12, 1996; accepted for publication June 26, 1996)
Fiz. Tekh. Poluprovodn. **31**, 498–502 (April 1997)

The results of theoretical and experimental studies of the time dependences of radiative cooling produced by a pulsed semiconductor source of negative interband luminescence are presented. The kinetics of cooling and subsequent temperature relaxation of the passive element, which depend on the power and duration of the luminescence pulse, are studied. It is shown that the luminescence excitation method employed (the magnetoconcentration effect) makes possible both radiative cooling and heating of the passive element; this is promising for practical applications. The results of a theoretical study are in qualitative agreement with the experimental data. The maximum cooling achieved in the experiment at room temperature is 0.62 K. © *1997 American Institute of Physics.* [S1063-7826(97)02404-6]

One manifestation of negative luminescence of semiconductors[1,2] is radiative cooling (RC).[3] This effect can be described as follows. If nonequilibrium electron and hole densities are maintained below the equilibrium values by some method, then the radiative-recombination intensity drops below the intensity of the equilibrium radiation.[4] As a result, in the region $\omega > E_g/\hbar$ (E_g is the band gap) absorption dominates over emission, i.e., a radiation energy flux from the outside toward the surface of the semiconductor appears. Therefore, the semiconductor absorbs the thermal radiation from the surrounding bodies in this part of the spectrum and can thereby lower their temperature.

Radiative cooling was first observed in Ref. 3. An InSb crystal was employed as the cooling (active) element under the conditions of magnetocondensation, which is one of the most effective and well-studied methods for exciting negative luminescence. It is realized in a semiconductor placed in a magnetic field **H** and in a pulling electric field **E** oriented perpendicular to the magnetic field. Bipolar depletion appears near one of the wide faces of the plate, which is parallel to **H** and **E** and which is characterized by a low rate of surface recombination, and therefore the radiation from this face is suppressed in the spectral region, $\omega > E_g/\hbar$. When the polarity of the controlling voltage is reversed, this surface becomes enriched with nonequilibrium carriers, which results in excess emission.

Since the cooling element is heated when a current passes through it, its thermal emission in the rest of the spectrum ($\omega < E_g/\hbar$) increases, which in itself results in heating of the cooled (passive) element optically coupled with it. Therefore, the magnitude of the radiative cooling depends not only on the power of the negative luminescence, but also on the duration of the luminescence pulse. The analysis of this effect performed in Ref. 3 was based on the assumption that the change in the temperature of the cooled layer is proportional to the luminescence power; this assumption is valid if the Joule heating of the active element is ignored. The main laws of the radiative cooling, taking into account the thermal radiation of the current-heated cooling element, were investigated theoretically in Ref. 5.

In this present study we investigate experimentally and theoretically the time and field dependences of the radiative cooling produced by semiconductors under conditions of magnetoconcentration. The kinetics of cooling and subsequent temperature relaxation of the cooled layer are studied. The possibilities of practical applications of the radiative cooling, making it possible to cool and heat a passive element, are discussed.

1. EXPERIMENTAL PROCEDURE

The radiative cooling was studied at $T = 300$ K using InSb and CdHgTe semiconductor crystals as the active element. The crystals were made in the form of $(0.015-0.02)\times 3\times 8$-mm^3 thin plates fitted with ohmic contacts. The semiconductor was cemented to a substrate, which was secured to a copper radiator with the aid of a heat-conducting Paste. The radiator was built into the casing of a vacuum capsule so that the part of the radiator to which the semiconductor crystal was secured was located inside the capsule. The cooled element was placed in the capsule directly near the surface of the semiconductor with a low surface recombination rate (at a distance $\simeq 1.5$ mm). An optical filter was placed between the cooling and cooled elements. The measuring capsule was placed between the poles of a magnet; the magnetic field could be regulated from 0 to 20 kOe. The duration t_i of the controlling electric field pulses was varied from 30 to 500 ms.

The choice of the semiconductor materials employed as the active element is dictated primarily by the possibility of obtaining a high negative-luminescence power. The latter power with the same bipolar depletion of the semiconductor increases with decreasing band gap. The maximum possible modulated radiation flux P_0 is reached with complete bipolar depletion of the surface region of the semiconductor. Ac-

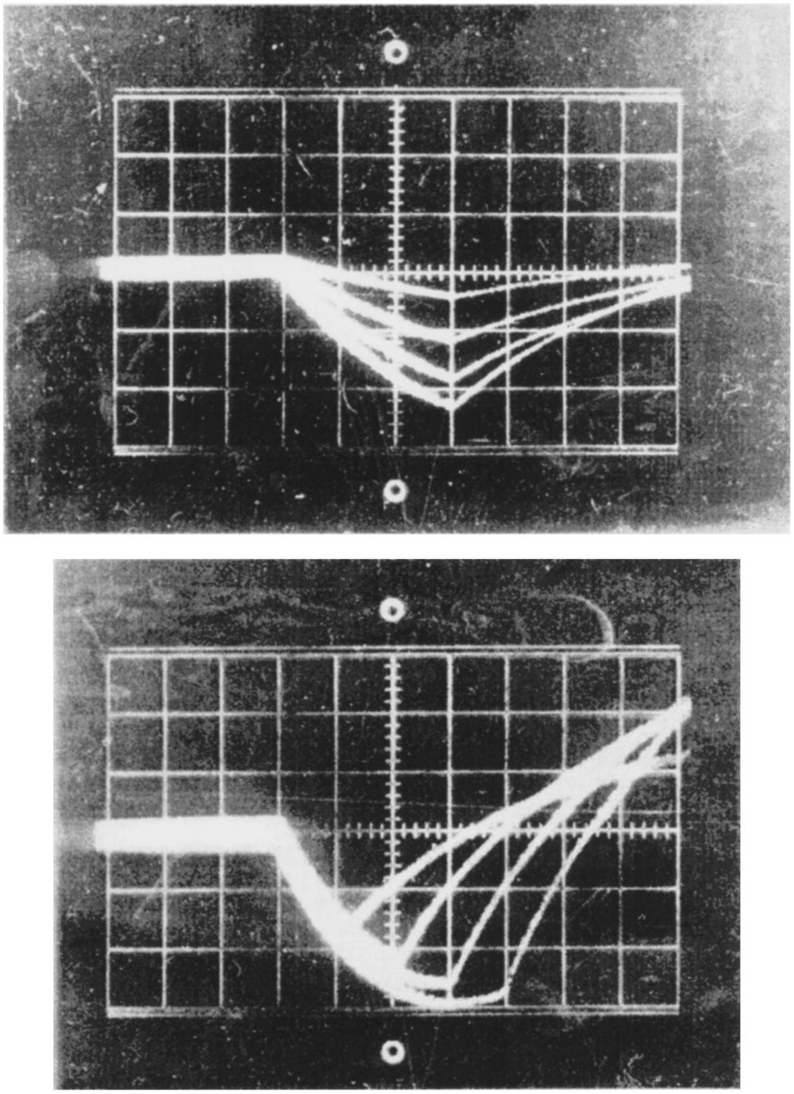

FIG. 1. Oscillograms of the electrical signal of a passive element (temperature of the element). The magnetic field intensity $H=15$ kOe. a—Duration of the negative-luminescence pulse 300 ms, voltage, V: 0.4, 0.8, 1.2, 1.6 (oscillograms top to bottom); cooling element—InSb; b—voltage 3.2 V, durations of the negative-luminescence pulses, ms: 100, 200, 300, 400 (oscillograms from left to right); cooling element—$Cd_{0.19}Hg_{0.81}Te$.

cording to Ref. 2, as E_g increases, P_0 decreases according to the law $E_g^3\exp(-E_g/kT)$ with $E_g>kT$.

A lithium tantalate single crystal, consisting of a thin $0.015\times 3\times 8$-mm^3 plate with metal electrodes deposited on the wide faces was used as the cooled element. The plate was secured in a fixed position, without touching other elements of the capsule, with the aid of thin contact wires. One of the wide faces of the cooled element (turned toward the semiconductor plate) was blackened. The cooled pyroelectric element, which served as a temperature sensor, possessed a high temperature sensitivity (at a level of 0.01 K) and was inserted into the amplifying circuit with a high-resistance input.[6]

2. RESULTS AND DISCUSSION

The variation of the temperature $T(t)$ of the cooled layer under conditions when the luminescence pulse duration t_i is much greater than the characteristic cooling time of the layer was calculated in Ref. 5. We shall investigate here the time dependence $T(t)$ for a pulse of finite duration. For $0<t<t_i$ the temperature $T(t)$ of the cooled layer is described by the equation

$$\frac{dT}{dt}+\frac{T-T_0}{\tau}=\frac{T_0 t}{\tau\tau_s}-W. \tag{1}$$

Here $T_0=T(0)$ is the initial temperature of the cooled element and the cooling semiconductor (state of equilibrium), τ is the cooling time of the layer due to the thermal emission of the layer, τ_0 is a constant determining the Joule heating of the semiconductor [it is assumed that the temperature T_s of the semiconductor increases according to the law $T_s=T_0(1+t/\tau_s)$, where $\tau_s\sim I^{-2}$ for a semiconductor with a current I that is uniform over the cross section], W is proportional to the negative-luminescence power,[4]

$$\tau=\frac{\rho C d}{4\sigma T_0^3(1-R)},\quad W=\frac{P}{\rho C d}, \tag{2}$$

where ρ, C, and d are the density, specific heat, and thickness of the layer, σ is the Stefan–Boltzmann constant, R is the reflection coefficient of the semiconductor surface, averaged over the entire spectrum $(0\leqslant\omega<\infty)$, and P is the power of the negative luminescence. The expression for τ is given for the case corresponding to the experimental condi-

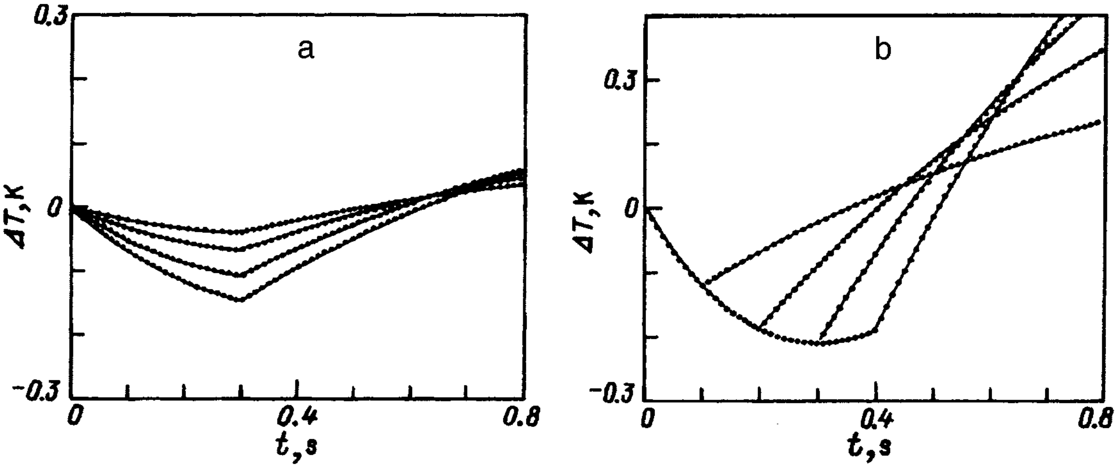

FIG. 2. Kinetics of the temperature of the cooled element (theory) as a function of the power (a) and duration (b) of the negative-luminescence pulses. a—(P, W/cm^2, τ_s, s): (0.0016, 800), (0.0026, 550), (0.0046, 360) (curves top to bottom), $\tau = 0.74$ s; b—$P = 0.01$ W/cm^2, $\tau_s = 100$ s, $\tau = 0.74$ s, t_i, ms: 100, 200, 300, 400 (curves from left to right).

tions when the front (turned toward the conductor) surface is blackened and the back surface is metallized.

The time t_1 over which a photon deficiency is produced can be estimated as the bipolar depletion time of the semiconductor surface layer with a thickness of the order of α^{-1} (α is the interband absorption coefficient). Under magnetoconcentration conditions $t_1 \simeq (\alpha V)^{-1}$, where $V = \mu_n \mu_p HF/c$ is the transverse drift velocity of electron-hole pairs, μ_n and μ_p are the electron and hole mobilities, and c is the speed of light. After the voltage is switched off, the recombination radiation is restored to the equilibrium value over a time t_2 of the order of the recombination time. Under the experimental conditions $t \simeq 10^{-9}$ s, $t_2 \simeq 10^{-8}$ s, and the thermal constants τ, $\tau_s \gg t_1, t_2$ ($\tau \simeq 1$ s and, according to our estimates τ_s, varies in the range 100–1000 s, depending on the controlling voltage), so that the luminescence pulse can be assumed to be square.

For $\Delta T = T - T_0$ we obtain the following expression from Eq. (1) for $t \leqslant t_i$:

$$\Delta T = \tau(W + T_0/\tau_e)(e^{-t/\tau} - 1) + T_0 t/\tau_s . \qquad (3)$$

After the controlling voltage is switched off, the temperature of the cooled element starts to decrease mainly as a result of contact with the substrate. The change in the semiconductor temperature can be ignored for the short time interval after $t = t_i$, during which the temperature of the thin cooled layer relaxes. In this case the temperature $T(t)$ is described by Eq. (1), where the right side is now equal to $T_0 t_i/\tau\tau_s$. The solution for $t \geqslant t_i$ has the form

$$\Delta T = \tau(W + T_0/\tau_s)(e^{-t/\tau} - e^{(t_i - t)/\tau}) + T_0 t_i/\tau_s . \qquad (4)$$

The experimental temperature dependences $\Delta T(t)$ as a function of the magnitude of the controlling voltage and pulse duration (oscillograms) are shown in Figs. 1a and 1b, respectively. The theoretical time dependences $\Delta T(t)$, calculated for different powers and durations of the luminescence pulse, are illustrated in Figs. 2a and 2b. The value $\tau = 0.74$ s, obtained from Eq. (2) for the parameters of lithium tantalate[7] $\rho = 7$ g/cm^3 and $C = 0.42$ J/g·K at $T = 300$ K and $R = 0.36$, was used for the calculation. The Joule heating parameter τ_s of the semiconductor was estimated from the temperature relaxation of the passive element after the negative luminescence pulse is switched off. When varying the luminescence power W, which depends linearly on the applied voltage U, the voltage dependence $\tau_s(U) \sim U^{-2}$ was used.

The characteristic features of the experimental observations of radiation cooling are determined by the fact that when a controlling voltage was applied to the cooling element, the temperature of the element always increased (Joule heating). The increase in the temperature of the semiconductor depends on the duration and amplitude of the controlling voltage and on the amount of energy lost by the crystal as a result of heat transfer to the substrate. As a result, during a time corresponding to the pulse duration, two processes compete with one another—negative luminescence (spectral range $E_g/\hbar < \omega < \infty$, cooling) and nonequilibrium thermal radiation (the range $0 < \omega < E_g/\hbar$, heating). Naturally, the action of negative luminescence should result in a decrease of the temperature of the cooled element and an increase in the intensity of the thermal radiation should result in heating of the element, i.e., the cooling element possessed the properties of a radiative cooler and heater simultaneously. We note that the onset of the negative-luminescence corresponds in time to the leading edge of the controlling pulse. The heat flux from the active element, however, increases as the heating of the semiconductor by the passing current increases.

A massive absorbing filter, transmitting radiation in the spectral region $\omega > E_g/\hbar$, was placed between the semiconductor crystal and the cooled element. This made it possible to decrease substantially the effect of the thermal radiation of the active element on the cooled element, although this effect could not be completely eliminated. The maximum decrease in the temperature of the cooled element under the experimental conditions also depends on the heat flux to this element as a result of heat conduction through the metal con-

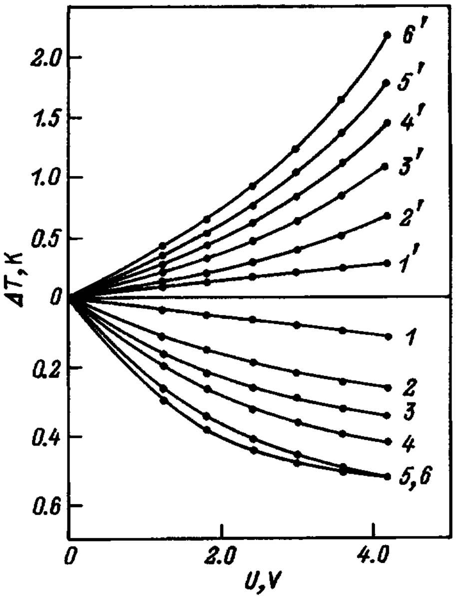

FIG. 3. Experimental field dependence of the cooling temperature (*1–6*) and heating temperature (*1′–6′*) of the cooled element for different luminescence pulse durations, ms: *1, 1′*—30; *2, 2′*—100; *3, 3′*—200; *4, 4′*—300; *5, 5′*—400; *6, 6′*—500. $H=20$ kOe, cooling element—InSb.

tacts of the holder (the capsule was evacuated and there was no convection).

The change (maximum value achieved over the pulse time) in the temperature of the cooled element is plotted in Fig. 3 as a function of the voltage applied to the active element (InSb crystal). The parameter of the curves is the duration t_i of the electric field pulse. We see that for a short pulse duration (curves *1*, *1′*, *2*, *2′*) the voltage dependence of the change in the temperature of the passive element is nearly linear. For long durations, heating of the InSb lattice has an effect (a small fraction of the thermal radiation from the semiconductor penetrates through the filter), which results in the sublinearity of the cooling curves (*3–6*) and superlinearity of the heating curves (*3′–6′*). An increase in the voltage (Fig. 1a) and the pulse duration (Fig. 1b) results in appreciable competition between the thermal radiation and the negative luminescence. As a result, the cooled element cools for some period of time and then starts to heat up. After the controlling voltage is switched off and, correspondingly, in the absence of negative luminescence, the rate of increase of the temperature of the cooled element increases with increasing amplitude or duration of the voltage pulse.

For the controlling-field values employed by us the luminescence power is a linear function of the voltage U. In the case of short pulses ($t_i \ll \tau$) it follows from Eq. (3) that $\Delta T \simeq -W t_i$, i.e., it does not depend on the heating parameter τ_s of the cooling element, and ΔT increases linearly with the voltage, as is observed experimentally (Fig. 3, curves *1, 1′*). For long pulses $\Delta T(t)$ reaches a minimum for $t_{\min} < t_i$ (see bottom curves in Figs. 1b or 2b). From Eq. (3) we find that $\Delta T_{\min}$ is determined by the expression

$$\frac{\Delta T_{\min}}{T_0} = -\frac{\tau}{\tau_s}\left[\frac{W\tau_s}{T_0} - \ln\left(1+\frac{W\tau_s}{T_0}\right)\right]. \tag{5}$$

Hence, for $W\tau_s/T_0 \ll 1$ we obtain

$$\Delta T_{\min} \simeq -W^2 \tau_s \tau / 2T_0 . \tag{6}$$

Therefore, for $W \sim U$ and $\tau_s \sim U^{-2}$ the function $\Delta T_{\min}(U)$ saturates. We see this tendency on curves *5* and *6* (Fig. 3), which correspond to the longest pulses.

The duration and amplitude of the controlling pulse can be chosen in such a way that the temperature of the cooled element does not rise above the equilibrium value after the voltage is switched off (Fig. 1a). This latter circumstance is very important in the case of practical applications of the effect described above.

It was found that the field dependences of the radiative cooling are approximately the same for the InSb and CdHgTe crystals used as the cooling elements, while the maximum values of ΔT differ substantially. The greatest decrease in temperature of the cooled element is observed in the case where a CdHgTe crystal ($E_g \simeq 0.14$ eV) is used as the active element and equals 0.62 K. The somewhat lower values which we obtained for $\Delta T_{\min}$ from theoretical calculations (Fig. 2b) could stem from the fact that the theory ignored the effect of a filter on radiation transfer and heat transfer through the metal contacts-holders.

Our studies showed that the radiative cooling has practical applications. It is important that in the present situation the active semiconductor element can function as a cooler and heater, depending on the polarity of the controlling voltage (see Fig. 3). This is especially important when it is used as an active element in a temperature stabilization device. We developed such a device and used it successfully to stabilize the frequency of a small quartz oscillator. Analysis of the operation of the device shows that because the active element switches on and off rapidly, it has advantages over analogs (for example, thermoelectric devices). We note that to improve the performance of this device, especially to decrease the time of the radiative effect on the controlled object, the cooling elements should be made of thin (high-resistance) semiconductor films fabricated from narrow-gap materials.

[1] C. C. Bolgov, V. K. Malyutenko, and V. I. Pipa, Pis'ma Zh. Tekh. Fiz. **5**, 1444 (1979) [Sov. Tech. Phys. Lett. **5**, 610 (1979)].
[2] S. S. Bolgov, V. K. Malyutenko, and V. I. Pipa, Fiz. Tekh. Poluprovodn. **17**, 208 (1983) [Sov. Phys. Semicond. **17**, 134 (1983)].
[3] P. Berdahl in *Proc. of the 18th Int. Conf. on the Physics of Semiconductors*, Stockholm (1986), p. 1595.
[4] V.K. Malyutenko. Semicond. Sci. Technol. **8**, S390 (1993).
[5] V. I. Pipa, Fiz. Tekh. Poluprovodn. **29**, 526 (1995) [Semiconductors **29**, 272 (1995)].
[6] V. F. Kosorotov, L. S. Kremenchutskiĭ, V. B. Samoĭlov, and L. V. Shchedrin, *Pyroelectric Effect and Its Practical Applications* [in Russian], Naukova dumka, Kiev (1989).
[7] M. P. Shaskol'skiĭ [Ed.], *Acoustic Crystals* [in Russian], Nauka, Moscow (1982).

Translated by M. E. Alferieff

Effect of ultrasonic treatment on deformation effects and the structure of local centers in the substrate and in the contact regions of *M*/*n* − n^+-GaAs structures (*M*=Pt, Cr, W)

I. B. Ermolovich, V. V. Milenin, R. V. Konakova, L. N. Primenko, I. V. Prokopenko, and V. L. Gromashevskiĭ

Institute of Semiconductor Physics, Ukrainian National Academy of Sciences, 252650 Kiev, Ukraine

(Submitted August 15, 1994; accepted for publication July 1, 1996)
Fiz. Tekh. Poluprovodn. **31**, 503–508 (April 1997)

The effect of ultrasonic treatment on the physicochemical, structural, and electrical properties of Pt, Cr, W/$n-n^+$-GaAs structures has been studied. It is shown that the ultrasonic treatment produces spatial and chemical ordering of the contact GaAs region. This decreases the reverse currents in diode structures with a Schottky barrier. A possible mechanism of the effect of ultrasonic treatment on the structural and chemical reorganization in a $M/n - n^+$-GaAs contact is discussed. © *1997 American Institute of Physics.* [S1063-7826(97)02504-0]

1. INTRODUCTION

The values and stability of the electrical parameters of metal-semiconductor (MS) contacts depend on the characteristic features of the physical-chemical state of the interface. Variations of the microscopic bonds between atoms at the interface result in a sharp variation of the electronic properties, operational characteristics, reliability, and longevity of macroscopic MS junctions. Controllable actions on the defect structure of a semiconductor in the contact region could make it possible to vary the interdiffusion of the contact materials and the character of the interphase reactions, i.e., form an interface to suit our purpose. Data are now available on the effectiveness of the influence of ultrasonic treatment (UST) on the microdefect structure of semiconductors: diffusion and association of point defects, dissociation of complex centers, and others,[1–4] which can strongly influence the formation and parameters of Schottky barriers.

In the present work we studied the effect of UST on the physicochemical properties of technically important heterostructures $M/n - n^+$-GaAs formed by metals with different chemical activities (Pt, Cr, W).

2. SAMPLES AND EXPERIMENTAL PROCEDURE

The MS structures were fabricated by electron-beam evaporation of Pt, Cr, and W in a 10^{-4} Pa vacuum on a $n-n^+$-GaAs substrate with (100) surface orientation and heated to temperatures of 200, 180, and 260 °C, respectively. The metallization layers were no more than 800 Å thick. The structures fabricated were subjected to UST in a longitudinal pulse mode with frequency in the range 0.2–14 MHz and power 6 W/cm^2 at the transmitting transducer. The treatment time was 1 h.

The physicochemical state of the heterostructures was analyzed using the following methods. The deformation effects in the $M/n-n^+$-GaAs systems were studied by the x-ray diffraction method.[5] The uniformity of the bending of the system was checked on the basis of measurements of the peak intensities of the Bragg reflections and the radii of curvature, measured at intermediate points of translation of the sample. In so doing, we assumed that the radius of curvature $R>0$ for a concave surface and $R<0$ for a convex surface.

The spectrum of local states in the band gap of the substrate and near the heterojunction of the structures was determined from the photoluminescence (PL) spectra measured on the substrate and metal sides, respectively. The luminescence was recorded in the spectral range 0.54–2.0 eV at 77 K. Excitation was performed with light from a PZh-100 incandescent lamp with $h\nu>2.0$ eV [the absorption coefficient was 10^5 cm^{-1} (Ref. 6)]. A PbS photoresistor served as a PL detector. The PL measured on the metal side was approximately ten times weaker than that on the substrate side, since the metal layer served as a neutral attenuator for the excitation and the luminescence light.

The data on the spatial distribution of the components of the metal-GaAs contacts in the direction normal to their surface were obtained with the aid of Auger spectroscopy with layerwise etching of the samples using 1-keV Ar ions. The Auger spectra were measured in a differentiated mode in the region of the LMM transitions in Cr, Ga, and As and *MNN* transitions in Pt and W. The spectra were recorded with analyzer resolution of 0.6%, primary electron beam energy $E=3$ keV, and a modulation amplitude of 5 V.

The atomic electron densities were calculated by the method of elemental sensitivity coefficients,[7] using the most intense lines in the Auger spectra.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The normalized elemental distribution profiles reflecting the variation of the chemical composition of the $M/n-n^+$-GaAs interphase boundaries before and after UST are shown in Fig. 1.

We see that the main process occurring at the interface under the conditions chosen for preparing the structures is interdiffusion of the contact components, which destroys the stoichiometric composition in the surface layer of the semiconductor. There are no pronounced effects associated with the chemical interactions between the metals and GaAs, even under the most favorable conditions—with deposition of Pt, which are characterized by high chemical activity.[8]

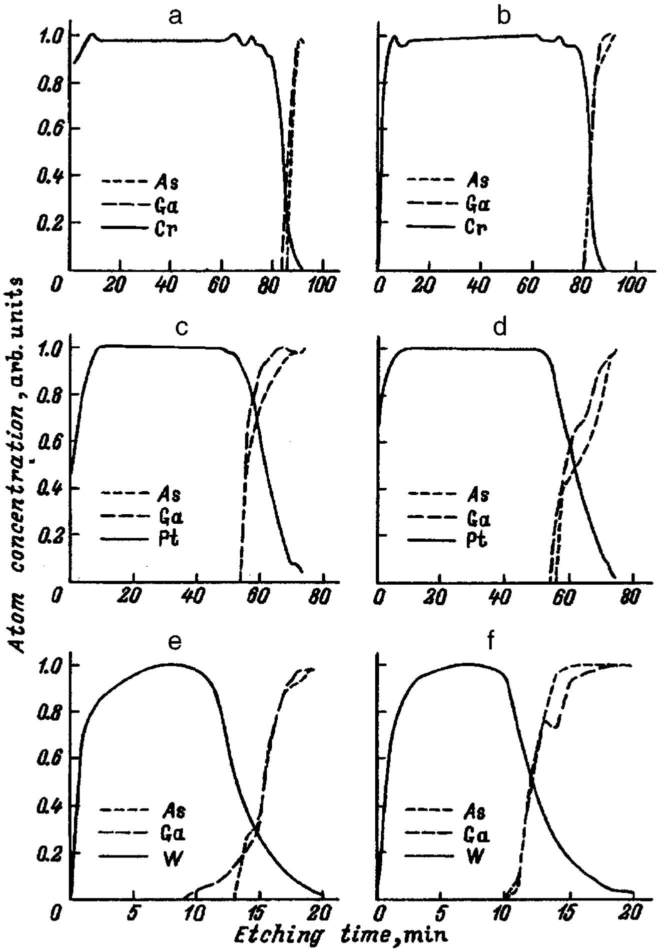

FIG. 1. Normalized depth distribution of the components of contacts before (a, c, e) and after (b, d, f) ultrasonic treatment. a, b—Cr/GaAs; c, d—Pt/GaAs; e, f—W/GaAs.

Ultrasonic treatment does not influence appreciably the penetration depth of metal atoms in GaAs or the variation of the stoichiometric ratio of the components in the surface region of the semiconductor. The observed characteristic features in the Ga and As distributions in contacts with Pt and W are most likely etching artifacts, which are especially pronounced on the ''tails'' of the profiles.[9]

Ultrasonic treatment under the conditions employed by us therefore does not influence the redistribution of the atomic components in the experimental contact pairs. The degree of smearing of the boundary, however, is determined by the chemical activity of the metal and the initial conditions under which the structures were prepared: The highest degree of smearing occurs in the case of Pt and the smearing is much weaker for Cr and W. For the latter two metallizations, the diffusion of the semiconductor components is more nonstoichiometric due to the large penetration depth of Ga in the metal layers. The observed interactions at the interface of the MS contacts should result in variations in the structure and evolution of the boundary states in GaAs.

Typical PL spectra of the experimental structures before and after UST are shown in Fig. 2. The spectra were measured on the metallized layer side (Figs. 2a–2e) and on the substrate side (Figs. 2f–2j). Wide overlapping bands with $h\nu_{\max}=0.78$, 1.02, 1.20, and 1.33 eV, whose intensity ratios are different in different samples, are observed in the PL spectra of the substrates. The PL spectra obtained on the side of the metallized layers of the initial structures (Figs. 2a–2e) contain a larger number of bands than the spectra obtained on the substrate side, where the bands are poorly resolved. This shows that the density of defects is higher in the contact regions than in the bulk. We note that the shape of the PL energy spectrum of the contacts is close to that of the PL spectra of GaAs epitaxial layers.[10] Metallization transforms the spectrum, but it is not possible to establish a correlation between the variations and the nature of the deposited metal. Ultrasonic treatment intensifies PL as a whole for the sub-

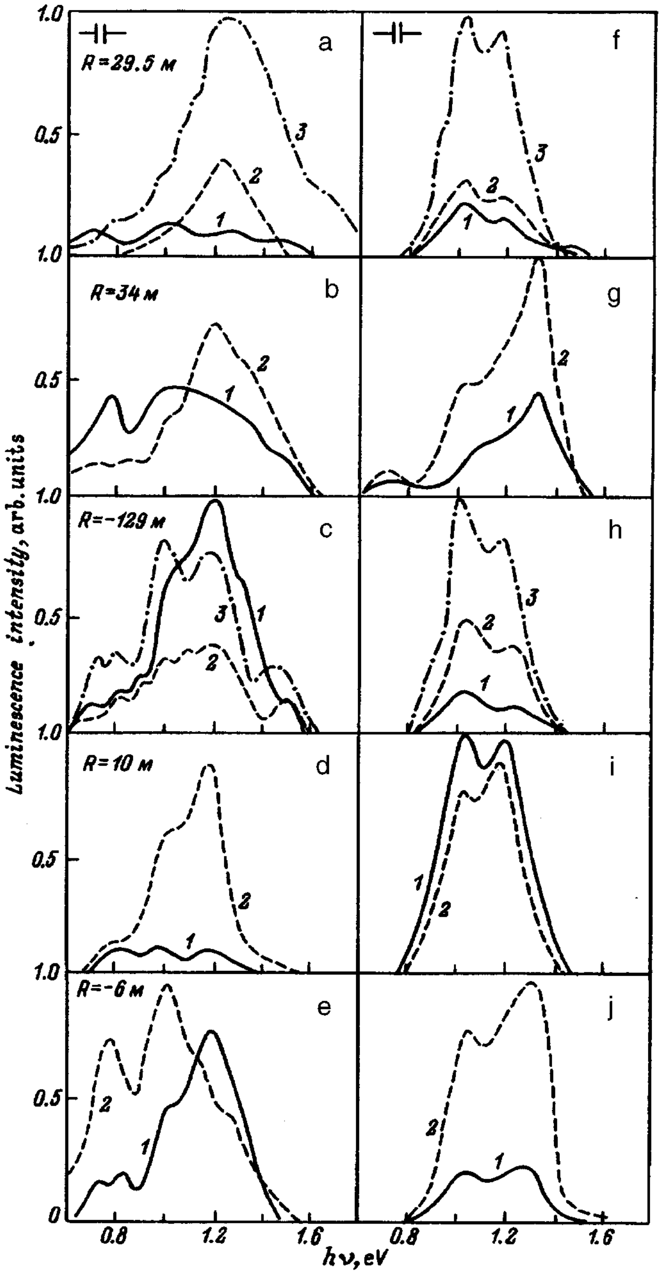

FIG. 2. Luminescence spectra of $M/n-n^{+}$-GaAs structures at 77 K: M=Pt (a, b, c, f, g, h); M=Cr (d, i); M=W (e, j) on the metal side—a–e; on the substrate side—f–j. 1—Initial, 2 and 3—after UST-1 and UST-2, respectively.

TABLE I. Radii of curvature $R$and deformations ε of contacts before and after ultrasonic treatment (UST).

Type of metal-conductor contact	Before UST		After UST	
	R,m	ε	R,m	ε
Pt/GaAs				
1	29.5	5×10^{-6}	34	4.4×10^{-6}
2	29.5	5×10^{-6}	-	-
3	10.2	1.4×10^{-5}	-	-
4	34	4.4×10^{-6}	29.5	5×10^{-6}
5	23	6.5×10^{-6}	-	-
6	25.8	5.8×10^{-6}	-	-
7	29.5	5×10^{-6}	-	-
8	-129	1.2×10^{-6}	-129	1.2×10^{-6}
9	∞	-	∞	-
10	18.8	8×10^{-6}	-	-
11	22.9	6.5×10^{-6}	22.9	6.5×10^{-6}
W/GaAs				
$1'$	-6.9	2.2×10^{-5}	-6.6	2.35×10^{-5}
$2'$	-6.5	2.3×10^{-5}	-	-
$3'$	-6.0	2.5×10^{-5}	-	-
Cr/GaAs				
$1''$	10.0	1.94×10^{-5}	-	-

strates and the contact regions and it results in band narrowing and redistribution of the band intensities, which are especially pronounced in the PL spectra obtained on the metallized-layer side. This means that the ultrasonic wave substantially weakens the nonradiative recombination channel, homogenizes the structure as a result of the large ordering of the spatial localization of defects and impurities in the lattice, and also changes the spectrum of localized states, which is most effective in the contact region.

Although there is no single point of view concerning the microstructure of the centers responsible for the observed radiative transitions, it is believed that the luminescence centers are complexes, whose common constituent are vacancies.[6] On the other hand, the form of the PL spectra of epitaxial structures subjected to moderate deformations[11] correlates well with the data obtained in the present work: absent or weak edge band and quite intense long-wavelength PL bands. All this indicates that deformation effects, which occur during the formation of a contact,[5] in the restructuring of local centers in the contact layers of a semiconductor, including also with UST, must be taken into account.

The radii of curvature R of structures, whose PL spectra are shown in Fig. 2, are presented in Table I. As follows from these data, in the case of Pt under standard conditions of metallization, the mechanical stresses vary from sample to sample even for the same film thickness: In most cases, in addition to the absence or compensation for the stresses in a metal/GaAs contact (R=0), a tensile stress with deformations in the range $\varepsilon=510^{-6}-1.4\times10^{-5}$ is present in the semiconductor $(R>0)$, together with a compressive stress $(R<0)$ with $\varepsilon=1.2\times10^{-6}$. The sign of the deformations $R<0$ for samples with tungsten metallization and $R>0$ for samples obtained with deposition of Cr.

Ultrasonic treatment does not influence appreciably the deformation parameters of the structures; i.e., the appearance and propagation of elastic waves does not change the density of the structural defects in the crystal that could be recorded according to the level of the residual deformations. However, sharp variations, which correlated with the sign of the curvature of the investigated structures, in the structure of the PL spectra are recorded. As the stresses decrease (Figs. 2c and 2e), the most intense band becomes the 1.02-eV band and under tensile stresses the 1.2-eV band is strongest (Figs. 2a, 2b, and 2d).

In accordance with Ref. 12, the external stresses change the chemical potential of the vacancies as compared with the corresponding value in the undeformed crystals, a consequence of which is the appearance of directed diffusion fluxes of vacancies. In the presence of compressive stresses, a state of supersaturation with respect to vacancies is realized. This supersaturation relaxes in time to a new equilibrium state of the system because of the vacancy losses at a free surface or because of dislocations or the formation of complexes and clusters of point defects. The picture is reversed when the crystal is stretched: Complexes or clusters of point defects, which include vacancies, decompose into point-like components. The degree of supersaturation or undersaturation with vacancies depends on the concentration of dopants and background impurities, which can participate in the formation of impurity-vacancy complexes—radiative and nonradiative recombination centers. Indeed, it has been found that compressive stresses give rise to the appearance of complex centers of emission for the 1.02-eV band—donor-acceptor pairs of the type $(V_{\mathrm{Ga}}-D)$.[13–15] At the same time, predominant centers of emission in the presence of tensile stresses are isolated acceptors $\mathrm{Cu}_{\mathrm{Ga}}$ (1.02-eV band).[16] The presence of the band with $h\nu_{\max}=1.33$ eV, which is attributed to $V_{\mathrm{As}}\mathrm{Cu}_{\mathrm{Ga}}V_{\mathrm{As}}$ complexes, points to the presence of background Cu impurity in the experimental samples.[17] At low (room) temperatures, the association of defects into complex centers and dissociation of complex centers are retarded as a result of the extremely low diffusion coefficient of vacancies [$\sim 10^{-20}$ cm^2/s (Ref. 18)]; i.e., there are virtually no vacancy fluxes. Stimulated diffusion of vacancies and of impurities, including those from the metallized layer, occurs under the action of the ultrasonic wave, since the diffusion activation energy decreases: $E=E_0-\sigma_{U_t}V$, where E_0 is the equilibrium diffusion activation energy, σ_{U_t} is the ultrasonic stress, V is the activation volume, and the complex-formation process (dissociations, depending on the sign of the deformation of the metal/GaAs contact) is more vigorous. The increase in the PL intensity as a whole over the entire spectral region and the narrowing of the bands, which could be associated with an increase in the mobility of structural defects and impurities resulting in their spatial and chemical ordering, also attest to such a mechanism of the action of UST.

This is also confirmed by the character of the change in the PL from the backside of the heteropairs, which are strongly disordered by mechanical grinding at the stage of preparation of GaAs substrates for epitaxy. The increase in the intensity of the luminescence bands in this case is most likely due to the ordering of the dislocation structure. The vacancy fluxes, stimulated by ultrasonic deformation, ab-

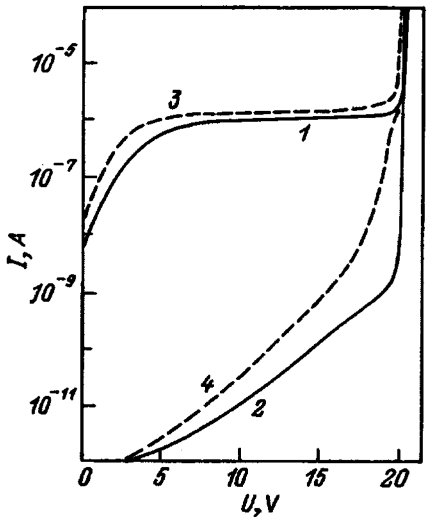

FIG. 3. Reverse current-voltage characteristics of Pt/$n - n^+$-GaAs diodes before (*1*) and after (*2*) UST. Reverse current-voltage characteristics of $M/N - n^+$-GaAs diodes before (*3*) and after (*4*) UST.

sorbed (or gettered) by dislocations, result in dislocation climb and ordering, changing the ratio of the radiative and nonradiative recombination rates.

The observed characteristic features of the restructuring of the defect-impurity structure of a metal/GaAs contact under the action of UST are manifested not only through the radiative and nonradiative recombination channels. Figure 3 shows the changes in the reverse currents of the Schottky diodes before and after UST. The decrease in the reverse current by two orders of magnitude is due to the decrease in the density of centers responsible for the recombination-generation component of the current.[19,20]

The use of UST in metal/GaAs contacts with opposite sign of the deformation thus makes it possible to determine more accurately the physicochemical structure of the luminescence centers—isolated point defects of the structure or their complexes, depending on the dynamics of the corresponding PL band intensity.

The investigations performed by us are also of interest for understanding the nature of defects participating in the stabilization of the Fermi level at the GaAs surface accompanying condensation of metal atoms on it. The deposition of metals perturbs the atomic and electronic subsystems of the semiconductor, which is accompanied by a deformation of the surface layer of the semiconductor which influences the defect composition and the conditions for the formation of defect complexes in it and therefore the energy position of local centers, whose real physicochemical structure is determined by many technological factors which are difficult to control. This objective circumstance makes it difficult to understand the evolution of the boundary electronic states. At the same time, in Ref. 21 a correlation was observed in the position of the PL peaks and the height of the Schottky barrier; this confirms the importance of taking into account the spectrum of local states in the contact regions of metal/GaAs contacts.

[1] V. L. Gromashevskiĭ, V. V. Dyakin, and E. A. Sal'kov, Ukr. Fiz. Zh. **29**, 550 (1984).

[2] V. L. Gromashevskiĭ, V. V. Dyakin, and N. S. Sayats, in *Abstracts of Reports at the 12th All-Union Conference on Acoustoelectronics and Quantume Acoustics* [in Russian], Kiev, 1986, Part II, p. 35.

[3] A. P. Zdebskiĭ, V. L. Korchnaya, T. V. Torchinskaya, and M. K. Sheĭnkman, Pis'ma Zh. Tekh. Fiz. **12**, 76 (1986).

[4] A. P. Zdebskiĭ, M. I. Lisyanskoĭ, N. B. Luk;yanchikova, and M. K. Sheĭnkman, Pis'ma Zh. Tekh. Fiz. **13**, 1009 (1987) [Sov. Tech. Phys. Lett. **13**, 421 (1987)].

[5] Yu. A. Tkhorik and L. S. Khazan, *Plastic Deformation and Misfit Dislocations in Heteroepitaxial Systems* [in Russian], Naukova Dumka, Kiev (1983), p. 176.

[6] V. I. Gavrilenko, A. M. Grekhov, D. V. Korbutyak, and V. G. Litovchenko, *Handbook of the Optical Properties of Semiconductors* [in Russian], Naukova dumka, Kiev (1987), p. 353.

[7] L. E. Davis, N. C. McDonald, P. W. Palmberg, G. T. Piach, and R. E. Weber, *Handbook of Auger Spectroscopy,* Physical Electronics Juchel-tries, Minnesota (1976), p. 195.

[8] V. J. Kumar, Phys. Chem. Sol. **36**, 535 (1975).

[9] O. D. Protopopov, Obzory po élektron. tekhn. Ser. 7, 10 (117), 74 (1985).

[10] Zh. I. Alferov, V. M. Andreev, and D. Z. Garbuzov, Fiz. Tekh. Poluprovodn. **2**, 1441 (1968) [Sov. Phys. Semicond. **2**, 1204 (1968)].

[11] O. K. Gorodbnichenko, *High-Molecular Plasticity and Strength of Epitaxial Gallium Arsenide Structures* [in Russian], Vishcha Shkola, Kiev (1993), p. 108.

[12] J. P. Hirth and J. Lothe, *Theory of Dislocations*, Wiley, N. Y., 1982, 2nd edition [Russian trans., Atomizdat, Moscow (1972), p. 412].

[13] Queisser and C. S. Fuller, J. Appl. Phys. **37**, 4895 (1966).

[14] E. W. Williams, Phys. Rev. **168**, 922 (1968).

[15] S. Y. Chiang and G. L. Pearson, J. Luminesc. **10**, 313 (1979).

[16] A. S. Popov and A. Jakimova, Phys. Status Solidi A **51**, PK17 (1979).

[17] N. S. Averkiev, T. K. Ashirov, and A. A. Gutkin, Fiz. Tekh. Poluprovodn. **15**, 1970 (1981) [Sov. Phys. **15**, 1145 (1981)].

[18] G. B. Abdullaev and G. D. Dzhafarov, *Atomic Diffusion in Semiconductor Structures* [in Russian], Atomizdat, Moscow (1980), p. 221.

[19] A. A. Akopyan, Z. S. Gribnikov, R. V. Konakova, Yu. A. Tkhorik, and Yu. M. Shvarts, Fiz. Tekh. Poluprovodn. **9**, 1799 (1975) [Sov. Phys. Semicond. **9**, 1182 (1975)].

[20] S. A. Grusha, F. Dubetski, M. Dubovinski, V. I. Zimenko, K. A. Ismaĭlov, R. V. Konakova, V. S. Lysenko, I. P. Osiyuk, O. V. Snitko, Yu. A. Tkhorik, V. M. Faĭnberg, and L. S. Kazan, Dokl. Akad. Nauk Ukr. SSR (A) **1**, 53 (1989).

[21] R. E. Vitturo, M. L. Slude, and L. J. Brillson, Phys. Rev. Lett. **57**, 487 (1986).

Translated by M. E. Alferieff

Transitory switching-on of microplasmas at subthreshold voltages

V. N. Dobrovol'skiĭ, I. E. Pal'tsev, and A. V. Romanov

Kiev State University, 252127 Kiev, Ukraine

Submitted April 22, 1996; accepted for publication May 22, 1996)
Fiz. Tekh. Poluprovodn. **31**, 509–510 (April 1997)

The kinetics of the establishment of a current through a *p–n* junction with microplasmas, upon application to the junction of a voltage less than the threshold voltage for switching-on of microplasmas under stationary conditions, has been investigated. A transitory switching-on of microplasmas and a current an order of magnitude higher than the stationary current were observed. An explanation is proposed for the observed effect. The explanation is based on the previously discussed strong effect of nonuniform heating of a *p–n* junction on the avalanche current. © *1997 American Institute of Physics.* [S1063-7826(97)02604-5]

In the presence of microplasma breakdown of a *p–n* junction, a region of current instability is present on the reverse branch of the current-voltage characteristic (IVC) of the junction. If the voltage U on the *p–n* junction lies in this region, then the current I through the junction possesses a component that varies randomly in time.[1] This component appears for the following reason. Microplasma channels are present in the space charge region (SCR) of the *p–n* junction. These channels are sections where the electric field is stronger than elsewhere in the SCR. In the presence of instability the electric field in a channel is strong enough for avalanche multiplication of electrons and holes to occur there, and outside the channel the electric field is too weak for such an occurrence. An electron (hole) which has randomly entered a channel triggers an avalanche there and increases the current — it switches on a microplasma. The current flowing through the microplasma channel is weak and the number of electrons and holes in the microplasma is small. Therefore, the probability for all electrons and holes to escape randomly from the channel and for the avalanche current to cease — the microplasma is switched off — is high.

At voltages U less than the lower limit of the region of instability avalanche multiplication does not occur in a microplasma channel. For U above the upper limit of this region the number of electrons and holes in the channel is large and a stationary avalanche-multiplication current is established through the channel. We denote by U_l the voltage corresponding to the lower limit of the region of instability of the lowest-voltage microplasma in the *p–n* junction. This is the threshold voltage. For $U< U_l$ all microplasmas are switched off, and for $U> U_l$ one or more of the microplasmas are switched on. This is the known behavior of microplasmas.

We have discovered the possibility of a different behavior of silicon *p–n* junctions — transitory switching-on of microplasmas for subthreshold values of U.

1. The *p–n* junctions employed were similar to those in Refs. 2 and 3 . They differed by the presence of the microplasmas. The experimental sample consisted of a crystal 1.8-mm-diam and 0.25-mm-thick disk. Measurements were also performed on KD 243 diodes. In these measurements, the crystals were enclosed in a plastic case.

Curve *1* in Fig. 1 is the I–V characteristic of one of the *p–n* junctions. The curve was taken from the screen of an oscillograph with the voltage U applied across the *p–n* junction which increased from 0 to 1330 V in 1 min. A sawtooth voltage with a frequency of 500 Hz and an amplitude of 6 V was summed with this voltage. The noise sections on the I–V characteristic are the regions of current instability.

The current oscillogram 2 was obtained with a constant voltage $U_0< U_l$ switched on. The rise time of the voltage was less than 1 μs; the values of U_0 and U_l are indicated in the figure. The initially high current drops over a time t to the value I_0 realized on the I–V characteristic (curve *1*) at $U=U_0$. Noise sections are observed in the oscillogram. The number of such sections is the same as on the I–V characteristic. The current values of these sections are close to the values of the corresponding sections of the I-V characteristic. As U_0 decreased, the initial current and the number of noise sections decreased.

2. The effect of nonuniform longitudinal heating of microplasma-free *p–n* junctions on the impact-ionization current was investigated on the basis of Refs. 2 and 3 . It was found that even weak, but sharply nonuniform, heating can substantially decrease the impact-ionization current — the heating can nearly switch off the impact-ionization current. The strong effect of nonuniform heating can be explained by the action of the mechanism studied theoretically in Ref. 4 .

Using the results of Refs. 2 –4 , the observed change in the current in a *p–n* junction with microplasmas can be explained as follows. When a voltage $U_0$is applied, four microplasmas are switched on in the junction. Impact ionization in the microplasma channels gives rise to a high current. As the channels heat up, the mechanism of Ref. 4 at first decreases and then completely stops the impact-ionization currents that flow through them — it switches off the microplasmas. Before each microplasma is switched off, the current through its channel is weak and the number of electrons and holes in it is small. Therefore, the probability that the electrons and holes escape simultaneously from the channel is high, which accounts for the characteristic noise.

Comparisons of the currents on the noise sections of curves *1* and 2 shows that the switching-off starts with the highest-voltage microplasmas. This sequence agrees with the mechanism that follows from Refs. 2 and 4 — the effect of the nonuniform heating on the impact-ionization current in-

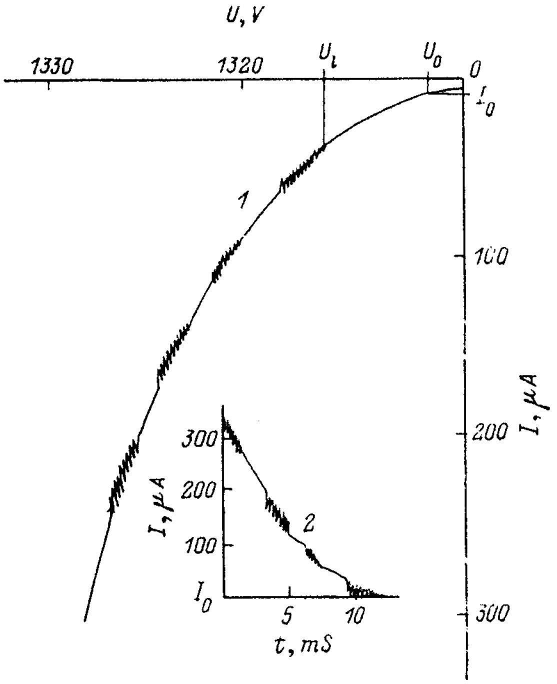

FIG. 1. Current-voltage characteristic of a *p–n* junction with microplasmas (1) and an oscillogram of the current (2).

creases with increasing avalanche-breakdown voltage.

The microplasma channels are heated by about 1.5 K. This value was estimated as follows. The breakdown voltage U_b of the *p–n* junction was determined from the I–V characteristic. On the basis of the impurity density dependences (presented in Ref. 5) of U_b and the thickness of the SCR at breakdown, the length of a microplasma is 0.1 mm. The diameter of microplasmas is ordinarily of the order of their length.[1] Assuming that the diameter equals the length, the thermal resistance of the microplasma channels[1] and then the degree of heating were calculated. The heating is several times weaker than that occurring when the impact-ionization current was switched off in a microplasma-free *p–n* junction with $U_b \simeq 400$ V.[3]

When a voltage U_0 greater than the instability voltages of all microplasmas was switched on, the initial current was greater than the current I_0 on the I–V characteristic (Fig. 1, curve *1*). The current without noise sections then dropped to I_0. The absence of noise is explained as follows. Heating decreases the impact-ionization currents that flow through the microplasma channels and around them. However, because of the large value of U_0, the current does not switch off the microplasmas or the noise that accompanies this process.

A memory effect occurred just as in a microplasma-free *p–n* junction.[3]. The voltage U_0 was switched off and then again switched on after some time had passed. With such repeated switching on, a high initial current (curve *2*) was observed only with a time interval of the order of a second or more.

The observed effects remained virtually the same in the range 77–300 K for the same energy of the first pulse.

3. The explanation presented above depends on the mechanism[4] of the strong effect of nonuniform heating of a *p–n* junction on avalanche breakdown. On the other hand, the microplasmas can be switched off as a result of a slight increase in U_b due to a heating-induced (uniform or nonuniform) decrease of the electron and hole impact-ionization coefficients.[1] The following considerations indicate that this does not happen.

To switch off microplasmas, U_b, must increase by an amount greater than the difference in the voltages at which the highest-voltage microplasma exists and U_l. Calculations according to the formulas in Ref. 1 show that at room temperature heating by not less than 10 K can produce such an increase in U_b. At low temperatures the heating should reach hundreds of degrees. Both are less than the estimated heating of the channel.

The above-described characteristic features of the behavior of microplasmas must be taken into account when making predictions of the reliability of semiconductor devices with the aid of microplasma cathode-ray curve tracers.[6]

---

[1] I. V. Grekhov and Yu. N. Serezhkin, *Avalanceh Breakdown of p–n Junctions in Semiconductors* [in Russian], Énergiya, Leningrad (1980).

[2] V. N. Dobrovol'skiĭ and A. V. Romanov, Fiz. Tekh. Poluprovodn. **26**, 1361 (1992) [Sov. Phys. Semicond. **26**, 763 (1992)].

[3] V. N. Dobrovol'skiĭ, A. V. Romanov, and S. B. Gryaznov, Fiz. Tekh. Poluprovodn. **29**, 1453 (1995) [Semiconductors **29**, 757 (1995)].

[4] V. N. Dobrovol'skiĭ and S. B. Gryaznov, Fiz. Tekh. Poluprovodn. **26**, 1366 (1992) [Sov. Phys. Semicond. **26**, 766 (1992)].

[5] S. M. Sze, *Physics of Semiconductor Devices*, Wiley Interscience, N. Y., 1981, 2nd edition [Russian trans., Mir, Moscow (1964), part 1].

[6] R. V. Konakova, P. Kordosh, Yu. A. Tkhorik, V. I. Faĭnberg, and F. Shtofanik, *Prediction of the Reliability of Semiconductor Avalanche Diodes* [in Russian], Naukova dumka, Kiev, 1986.

Translated by M. E. Alferieff