# Effect of local vibrations on the H and D atom densities at a Si surface

I. P. Ipatova and O. P. Chikalova-Luzina

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

K. Hess

*Beckman Institute, Illinois State University, Urbana, Illinois USA*

The equilibrium surface densities of passivating adatoms for a silicon crystal in equilibrium with
$H_2$ or $D_2$ gas are calculated. The difference in the surface densities of H and D adatoms is
determined by the difference in their local surface vibrations. The equilibrium deuterium surface
densities are an order of magnitude higher than the hydrogen surface densities. © *1999
American Institute of Physics.* [S1063-7826(99)01909-2]

In this paper we calculate the surface densities of hydrogen H (or deuterium D) atoms on a silicon surface in equilibrium with $H_2$ (or $D_2$) gas. The chemical potentials of H (or D) on a surface and the chemical potentials of H (or D) in gas are required to describe the equilibrium. In addition to the electronic contribution, the chemical potential of a H (or D) adatom possesses a vibrational contribution, which is determined by the dynamics of a crystal lattice with impurities. The chemical potential of H (or D) in gas is known from the thermodynamics of diatomic gases. The condition that the chemical potential of an adatom is equal to the chemical potential of H (or D) in gas at thermodynamic equilibrium makes it possible to find the equilibrium H (or D) surface density. It is shown that the surface density of D adatoms is an order of magnitude higher than that of H adatoms. For simplicity, an unreconstructed (100) silicon surface with the symmetry of a simple square lattice is considered. We assume that an adatom with mass $M^H$ (or $M^D$) lies above a surface silicon atom with mass $M$ and is bound with the silicon atom by a force constant $\gamma$. For low adatom density, in the nearest-neighbors approximation the following characteristic equation can be obtained for determining the frequencies of local vibrations of adatoms:[1]

$$1 - \frac{\gamma \omega^2}{\omega^2 - \omega_0^2} G(\omega^2) = 0, \qquad (1)$$

where $\omega_0^2 = \gamma / M^{(H,D)}$, and $G(\omega^2)$ is the diagonal element of the Green's functions matrix for a semi-infinite crystal.[2]

In the numerical calculations of the frequencies of the local vibrations from Eq. (1) it was assumed that the force constant $\gamma$ is, to a high degree of accuracy, equal to the effective force constant for central forces in a Si–Si bond at the surface, i.e., $\gamma = 8.8 \times 10^4$ dynes/cm. The experimentally measured values of the frequencies of surface vibrations of adatoms are used below: $\omega_{\text{loc}}^{(H)} = 3.96 \times 10^{14} \, \text{s}^{-1}$ and $\omega_{\text{loc}}^{(D)} = 2.83 \times 10^{14} \, \text{s}^{-1}$ (Ref. 4), which are somewhat different from the frequencies calculated from Eq. (1).

The following chemical potentials for the H and D adatoms are obtained from the vibrational part of the partition function of a crystal with adatoms written in terms of the frequency distribution function:

$$\mu_{\text{surf}}^{(H,D)} = T \ln n^{(H,D)} + 3T \ln \left[ 2 \sinh \frac{\hbar \omega_{\text{loc}}^{(H,D)}}{2T} \right] + \varepsilon_{(H,D)}. \quad (2)$$

Here $n^{(H,D)}$ are equilibrium H or D adatom surface densities. Since the binding energies of the isotopes are the same, i.e., $\varepsilon_H = \varepsilon_D$, the difference in the chemical potentials for H and D arises because of the difference in the frequencies of their local vibrations, $\omega_{\text{loc}}^{(H)} \neq \omega_{\text{loc}}^{(D)}$.

The chemical potentials for H and D gas are obtained from the condition of equilibrium for the dissociation reaction $H_2 = 2H$ (or $D_2 = 2D$) and from the chemical potential of a molecule consisting of two identical atoms[5]

$$\mu_{\text{gas}}^{(H,D)} = \frac{1}{2} \Bigg[ T \ln P_{(H_2, D_2)} - c_p T \ln T - \zeta_{(H_2, D_2)} T + \frac{\hbar \omega^{(H_2, D_2)}}{2} + \varepsilon_0 \Bigg]. \qquad (3)$$

Here $P_{H_2}$ and $P_{D_2}$ are the pressures of the corresponding gases, $c_p$ is the specific heat at constant pressure,

$$\zeta_{H_2} = \ln \left[ \frac{I_{H_2}}{\hbar^5} \left( \frac{M_{H_2}}{2\pi} \right)^{3/2} \right],$$

$$\zeta_{D_2} = \ln \left[ \frac{I_{D_2}}{\hbar^5} \left( \frac{M_{D_2}}{2\pi} \right)^{3/2} \right]$$

are, respectively, the chemical constants of hydrogen and deuterium gases; $I_{H_2}$ and $I_{D_2}$ are the moments of inertia and $\omega^{(H_2)}$ and $\omega^{(D_2)}$ the vibrational frequencies of $H_2$ and $D_2$ molecules, and $\varepsilon_0$ is the binding energy of a molecule in gas. Substituting the chemical potentials (2) and (3) into the condition of thermodynamic equilibrium

$$\mu_{\text{surf}}^{(H)} = \mu_{\text{gas}}^{(H)}, \qquad (4)$$

$$\mu_{\text{surf}}^{(D)} = \mu_{\text{gas}}^{(D)}, \qquad (5)$$

gives an equation for the surface densitites $n^{(H)}$ and $n^{(D)}$. Since the local vibrational frequencies $\omega_{loc}^{(H)}$ and $\omega_{loc}^{(D)}$ satisfy $\hbar\omega_{loc}^{(H)}$, $\hbar\omega_{loc}^{(D)} > T$ and since $I_{H_2}/I_{D_2} = M_{H_2}/M_{D_2}$, the ratio of the adatom densities is

$$\frac{n^{(D)}}{n^{(H)}} = \left(\frac{M_{H_2}}{M_{D_2}}\right)^{5/4} \exp\left[\frac{3\hbar[\omega_{loc}^{(H)} - \omega_{loc}^{(D)}]}{2T}\right]$$
$$\times \exp\left[\frac{\hbar\omega^{(D_2)} - \hbar\omega^{(H_2)}}{4T}\right]. \quad (6)$$

The first exponential factor, which contains the difference in the local vibrational frequencies of H and D, makes the main contribution to this expression. Since $\omega_{loc}^{(H)} > \omega_{loc}^{(D)}$, the deuterium density is much higher than the hydrogen density. Our calculation showed that at temperatures $T = 600 - 700$ K typical of this technology this ratio is of the order of 10. It is entirely likely that the difference in the dynamics of the vibrations of H and D isotopes is one of the factors that explains the large isotopic effect in MOS transistors. This effect has been observed in Ref. 6, where hydrogen passivation of a silicon surface was replaced by deuterium passivation.

[1] S. L. Cunnigham, L. Dobrzynski, and A. A. Maradudin, Phys. Rev. B **7**, 4643 (1997).
[2] E. W. Montroll, A. A. Maradudin, G. H. Weiss, and I. P. Ipatova, *Theory of Lattice Dynamics* (Academic Press, New York, 1971).
[3] L. Miglio, P. Ruggerone, and G. Benedek, Phys. Scr. **37**, 768 (1988).
[4] V. A. Burrows, Y. J. Chabal, G. S. Higashi, K. Raghavachary, and S. B. Christman, Appl. Phys. Lett. **53**, 998 (1988).
[5] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, New York; Nauka, Moscow, 1976, Chap. 9).
[6] I. C. Kizilyalli, J. W. Lyding, and K. Hess, IEEE Electron Device Lett. **18**, 81 (1997).

Translated by M. E. Alferieff

# Irradiation as a possible method for producing SiC heterostructures

A. A. Lebedev*⁾

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

Certain aspects of the physics of heteropolytypic junctions based on silicon carbide are examined. It is known that the introduction of certain impurities into the growth zone during epitaxy of silicon carbide results in the growth of films whose polytype is different from that of the initial substrate. It is also known that these impurities lead to the formation of certain deep centers in the band gap of the conductor. Analysis of published data performed in this paper shows that irradiation of SiC with various charged particles also leads to the formation of these deep centers. It is assumed that under certain experimental conditions transformation of the polytype of the already grown epitaxial SiC structure is possible under the action of irradiation and subsequent annealing. © *1999 American Institute of Physics.*
[S1063-7826(99)02009-8]

The term ''silicon carbide'' denotes essentially an entire class of semiconductor compounds, since SiC can crystallize in various modifications — polytypes. The polytypes of SiC with the same chemical composition can differ substantially with respect to their electrical properties. For example, the band gaps range from 2.4 (3*C*-SiC) to 3.3 eV (2*H*-SiC). This makes silicon carbide a promising material from the standpoint of producing various types of heterostructures.

It was found in Refs. 1–3 that when certain impurities are introduced into the growth zone of SiC layers, epitaxial films with a polytype different from that of the substrate employed can be obtained. The introduction of the rare-earth elements Sc and Tb, as well as Al and B, for example, resulted in growth of epitaxial layers of 4*H*-SiC on 6*H*-SiC substrates. The most efficient transformation of the polytype of the growing layer 6*H* ⇒ 4*H* was observed with the introduction of group-IV impurities Sn, Pb, and Ge. Group-V impurities (nitrogen and phosphorus) gave rise to growth of the 3*C*-SiC polytype.

It was also found that the Si:C ratio in the growth zone strongly affected heteropolytypic epitaxy. Thus, increasing the Si concentration increased the likelihood that 3*C*-SiC or other polytypes with a low percentage of hexagonality will form. At the same time, the introduction of excess carbon made it possible to grow on 6*H*-SiC substrates 4*H*-SiC epitaxial layers also from Gd and Dy melts.[3] In Ref. 1 it was also noted that a transformation of the substrate polytype occurs especially easily for growth in the (0001)C direction. The temperature and rate of growth had virtually no effect on the heteropolytypic epitaxy. The formation of 3*C*-SiC layers on 6*H*-SiC substrates is also possible during sublimation growth in excess silicon with no additional doping.[4]

Since the nature of polytypism is still unclear, it is quite difficult to understand the nature of heteropolytypic epitaxy. Besides the impurity composition of the growth zone, it seems that various factors can influence the likelihood of heteropolytypic epitaxy — from thermodynamic (pressure, temperature) to crystallographic (orientation and defect den-

sity of the substrate). It was found in Ref. 5 that the use of 6*H*-SiC Lely substrates with high dislocation density (of the order of $10^5$ cm$^{-2}$) in the standard (for growth of 6*H*-SiC layers) sublimation epitaxy process results in the growth of 3*C*-SiC epitaxial layers.

In Refs. 1 and 6 heteropolytypic epitaxy processes were attributed to the stoichiometric composition of various polytypes of SiC. Previously, it was observed that the concentration ratio [Si]/[C] varies in different polytypes of SiC and decreases with increasing percentage of hexagonality. It was shown that the ratio [Si]/[C] was 1.046, 1.022, and 1.001 for the silicon carbide polytypes 3*C*, 6*H*, and 4*H*, respectively.[7] The data on diffusion and solubility of impurities in various SiC polytypes examined in Ref. 1 also attest to a different concentrations of vacancies $V_C$. In Ref. 1 these data were explained by the fact that as the stresses in the lattice increase with increasing concentration of carbon vacancies, bonds between atoms at cubic sites become energetically more favorable. This is what leads to rearrangement of the crystal and transformation of the polytype. The influence of doping on heteropolytypic epitaxy can then be explained in terms of the introduction of impurities that occupy the carbon lattice sites, which accounts for the decrease in concentration of $V_C$ and for the stimulation of growth of epitaxial films with a higher percentage of hexagonality than the substrate employed. When impurities occupying silicon sites are introduced, the concentration of vacancies $V_C$ increases and, correspondingly, films with a lower percentage of hexagonality than the substrate used by us grow.

Our investigations made it possible to observe at least three types of deep centers, which have similar parameters and which participate in similar radiative-recombination processes in various polytypes of SiC. Comparing the energy position of the levels in the bottom half of the band gap in 4*H*-SiC and 6*H*-SiC, it is evident that most levels lie inside the two rather narrow bands in the band gap: $E_v + (0.5 - 0.6)$ eV (*D* band) and $E_v + (0.2 - 0.3)$ eV (*L* band) (Fig. 1). A *D* center, an *I* center, and a scandium level lie in the
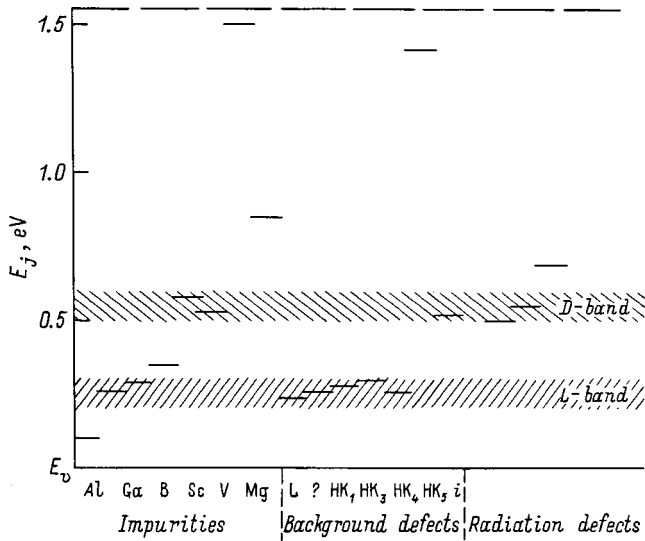
FIG. 1. Arrangement of known deep centers in the bottom half of the band gap in 6H-SiC.



FIG. 2. Density of carbon ($N_v^C$) and silicon ($N_v^{Si}$) vacancies and the average density of deep acceptor centers ($N_{dc}$) in epitaxial layers prepared by sublimation epitaxy as a function of the degree of hexagonality ($D$) of the SiC polytype.[7] The numbers on the curves correspond to vacancy concentrations: $1$ — $N_v^{Si}$, $2$ — $N_v^C$ and the concentration ($N_{dc}$) of centers of the type $i$ (3), $D$ (4), and $L$ (5).

first band; a deep aluminum level, a gallium level, and an $L$ center lie in the second band. A boron level is also present in (4H-SiC) or near (6H-SiC) the edge of the $L$ band. In the top half of the band gap the 4H-SiC and 6H-SiC levels are distributed uniformly, and it is impossible to distinguish any bands that are common to both polytypes.[8] Since the parameters of centers of one type, formed in the bottom half of the band gap of different polytypes of SiC, are approximately the same, it can be inferred that such bands are also present in other polytypes of silicon carbide. This is supported by the fact that the characteristics are similar and the shift of the maxima of the main electroluminescence bands in different polytypes of SiC is proportional to the band gap.

The formation of such bands in the band gaps of different polytypes of SiC attests to the presence of centers characteristic of SiC as a whole and related to the valence band, whose structure in various polytypes is similar. It can also be inferred that each band is related to a ''basic'' center that consists of intrinsic defects (an $i$ center in the first case and an $L$ center in the second case), which can interact with the impurity atoms introduced and form other centers with similar parameters.

Figure 2 shows the average density of three intrinsic defects ($i$, $D$, and $L$ centers) in epitaxial 6H and 4H-SiC layers obtained by sublimation epitaxy.[9] As is evident from Fig. 1, there is a rather good correlation between the decrease in the density of carbon vacancies ($V_C$) and the increase in the percentage of hexagonality of the polytype ($D$) and the decrease in the density of intrinsic defects.

Data obtained by various authors, who assumed that $V_C$ vacancies participate in the formation of different centers and that individual vacancies merge into stable clusters,[10,11] also support the idea that carbon vacancies are the main complex-forming defect in SiC. We note that a number of impurities (Sc, B, Al) employed in SiC heteroepitaxy[1] form deep centers in the above-noted $L$ and $D$ bands. The model of heteropolytypic epitaxy proposed in Ref. 1 can also be interpreted as follows. Each polytype of SiC corresponds to an
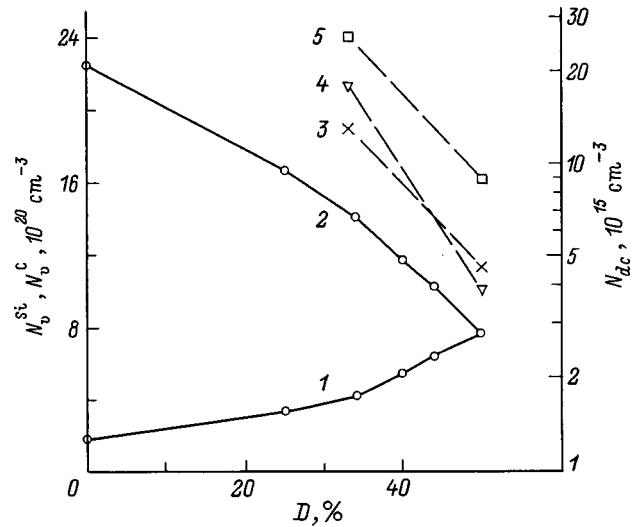
equilibrium concentration of defects (carbon vacancies), whose variation due to the addition of impurities that generate excess defects or, conversely, bind vacancies into complexes, accounts for the heteroepitaxy of films of polytypes whose degree of hexagonality is lower or higher, respectively, than that of the substrate. Thus, the introduction of Sc during heteroepitaxy of 4H-SiC films on 6H-SiC substrates, for example, leads to the formation of a deep center, which seems to bind vacancies into complexes and decreases their concentration to values characteristic of the 4H-SiC polytype.

It is well known that irradiation of semiconductors leads to the formation of a large number of defects, vacancies, and interstices in them. During subsequent annealing they migrate, recombine, and merge into complexes. For SiC, an increase in the concentrations of $L$, $D$, and $i$ centers, which, as noted above, are associated with carbon vacancies, was observed after irradiation and implantation.[12–14]

Let us assume that a double-layer SiC structure, one layer of which was doped with an impurity capable of binding a carbon vacancy, has been subjected to irradiation and subsequent annealing. The concentration of deep acceptors then will increase in this layer without a change in the crystal structure of the material. In the undoped layer, however, the concentration of vacancies formed after irradiation can be higher than the equilibrium concentration for the given polytype, and the stresses arising in the crystal lattice will lead to a transformation of the polytype.

If the mechanism for producing heteropolytypic structures, which has been examined above, is realizable in practice, then irradiation could become a method for producing SiC-based heterojunctions. The results obtained in Ref. 15, where it was found that inclusions of the polytype 3C-SiC appeared in photodiodes which are based on 6H-SiC and which are degraded as a result of application of a direct

current, show that transformation of the polytype of as-grown epitaxial structures is possible.

*⁾E-Mail: shura.lebe@pop.ioffe.rssi.ru

———————

[1] Yu. A. Vodakov, G. A. Lomakina, and E. N. Mokhov, Fiz. Tverd. Tela (Leningrad) **24**, 1377 (1982) [Sov. Phys. Solid State **24**, 780 (1982)].

[2] Yu. Vakhner and Yu. M. Tairov, Fiz. Tverd. Tela (Leningrad) **12**, 1543 (1970) [Sov. Phys. Solid State **12**, 1213 (1970)].

[3] Yu. A. Vodakov, E. N. Mokhov, A. D. Roenkov, and M. M. Anikin, Pis'ma Zh. Tekh. Fiz. **5**, 367 (1979) [Sov. Tech. Phys. Lett. **5**, 147 (1979)].

[4] Yu. A. Vodakov, G. A. Lomakina, E. N. Mokhov, and V. G. Oding, Fiz. Tverd. Tela (Leningrad) **19**, 1812 (1977) [Sov. Phys. Solid State **19**, 1695 (1977)].

[5] A. A. Lebedev, in *Abstracts of Reports at the 19th International Conference on Defects in Semiconductors* (Averio, Portugal, 1997), p. 239.

[6] Yu. A. Vodakov and E. N. Mokhov, Inst. Phys. Conf. Ser., No. 137, 197 (1994).

[7] N. D. Sorokin, Yu. M. Tairov, V. F. Tsvetkov, and M. A. Chernov, Kristallografiya **28**, 910 (1983) [Sov. Phys. Crystallogr. **28**, 539 (1983)].

[8] A. A. Lebedev, Fiz. Tekh. Poluprovodn. **33**, 129 (1999) [Semiconductors **33**, 107 (1999)].

[9] A. A. Lebedev, in *Abstracts of Reports at the International Conference on SiC, III Nitrides and Related Materials* (Stockholm, Sweden, 1997), p. 67.

[10] R. N. Kyutt, A. A. Lepeneva, G. A. Lomakina, E. N. Mokhov, A. S. Tregubova, and G. F. Yuldashev, Fiz. Tverd. Tela (Leningrad) **30**, 2606 (1988) [Sov. Phys. Solid State **30**, 1500 (1988)].

[11] A. I. Girka, V. A. Kuleshin, A. D. Mokrushin, E. N. Mokhov, S. V. Svirida, and A. V. Shishkin, Fiz. Tekh. Poluprovodn. **23**, 1270 (1989) [Sov. Phys. Semicond. **23**, 790 (1989)].

[12] M. M. Anikin, A. A. Lebedev, A. L. Syrkin, and A. V. Suvorov, Fiz. Tekh. Poluprovodn. **19**, 114 (1985) [Sov. Phys. Semicond. **19**, 69 (1985)].

[13] M. M. Anikin, A. A. Lebedev, A. L. Syrkin, and A. V. Suvorov, Fiz. Tekh. Poluprovodn. **20**, 2169 (1986) [Sov. Phys. Semicond. **20**, 1357 (1986)].

[14] G. C. Rubicki, J. Appl. Phys. **78**, 2996 (1995).

[15] G. Zeinther and D. Theeis, IEEE Trans. Electron. **ED-28**, 425 (1981).

Translated by M. E. Alferieff

# InGaAs/InP heterostructures with strained quantum wells and quantum dots ($\lambda = 1.5 - 1.9$ $\mu$m)

Z. N. Sokolova,*) D. A. Vinokurov, I. S. Tarasov, N. A. Gun'ko, and G. G. Zegrya

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

Strongly strained $In_xGa_{1-x}As/In_{0.53}Ga_{0.47}As/InP$ heterostructures with indium content $x = 0.69 - 1.0$ in the active region were investigated experimentally and theoretically. Two types of structures were obtained by vapor-phase epitaxy from metalorganic compounds: 1) with isolated compression-strained quantum wells and 2) with self-organized nanosize InAs clusters (quantum dots). The temperature dependence of the quantum radiation efficiency of samples with quantum wells in the temperature range 77–265 K is characterized by $T_0 = 43$ K. One reason for the low value of $T_0$ is electron delocalization in the active region. The maximum radiation wavelength obtained in structures with quantum dots is 1.9 $\mu$m at 77 K. © *1999 American Institute of Physics.* [S1063-7826(99)02109-2]

## 1. INTRODUCTION

A great deal of attention has recently been focused on producing perfect mid-IR-range radiation sources (1.5–5.0 $\mu$m), which find application in fiber optic communication lines, lidars, and gas analyzers. Solid solutions of III–V semiconductors cover the 1.5 to 2-$\mu$m range. Compounds containing antimony are now being studied extensively.[1,2] $In_xGa_{1-x}As/InP$ heterostructures with strained quantum wells (QWs)[3,4] or quantum dots (QDs)[5-7] also emit in the indicated range.

In this article we report the results of an experimental and theoretical study of strongly strained $In_xGa_{1-x}As/In_{0.53}Ga_{0.47}As/InP$ heterostructures with composition $x = 0.69 - 1.0$ in the active region.

## 2. PREPARATION AND PHOTOLUMINESCENCE STUDIES OF HETEROSTRUCTURES

Heterostructures with an isolated compression-strained QW with composition $x = 0.69 - 0.81$ (Ref. 4) and self-organized nanosize InAs clusters (QDs) (Ref. 5) were grown on InP (100) substrates by vapor-phase epitaxy from metal-organic compounds at low pressure.

The samples that we grew were heterostructures which we did not dope intentionally. These samples consisted of a 0.2-$\mu$m-thick InP buffer layer, a 0.2-$\mu$m-thick $In_{0.53}Ga_{0.47}As$ barrier layer (0.3 $\mu$m for QD) at the bottom, lattice-matched with the substrate, and then InAs quantum dots or an $In_xGa_{1-x}As$ quantum well with indium content in the range 0.69–0.81 and thickness from 25 to 100 Å , as well as a 0.2-$\mu$m-thick (0.04 $\mu$m for QD) (Refs. 4 and 5) $In_{0.53}Ga_{0.47}As$ layer on top. The compression strains in the quantum well, which are due to the mismatch of the lattice parameters, range from 1% to 1.8%, depending on the indium content in the well. For a structure with QDs the growth rate was 8 Å /s for barriers and 2 Å /s for QDs; the growth temperature was

500–600 °C. The amount of InAs deposited varied from 1.5 to 7 monolayers (ML).

Transmission microscopy investigations of the samples showed that two types of QDs can be identified: with average transverse size 50–60 and 10–20 nm.

The band diagram of the heterostructures is shown in Fig. 1. It thus follows that for QW a type-II heterojunction is investigated; in the QW region there is no potential well for electrons. Depending on the composition of the solid solution ($x = 0.69 - 0.81$) $\Delta E_c = (-3) - (-16)$ meV and $E_g = 664 - 607$ meV, the stress-induced splitting of the light- and heavy-hole subbands is $\delta = 64 - 115$ meV, and the QW depth for heavy holes is 105–175 meV.

The photoluminescence (PL) was investigated in the temperature range 77–265 K for samples with QWs and at 77 K for the samples with QDs. An argon laser ($\lambda = 0.514$ $\mu$m) with excitation density 1 kW/cm² served as the source of excitation radiation. Curves of the wavelength of the PL of a QW at the maximum intensity versus the QW thickness at 77 K were obtained.[4] The longest radiation wavelength is 1.66 $\mu$m. The energies of the electron transitions from the conduction-band bottom to the first heavy-hole quantum-well level were calculated. It was found that the transition energies computed using a rough estimate are 30–70 meV less than the measured values.

The temperature dependence of the half-width of the radiation spectrum for a 50-Å -thick structure with a QW and $x = 0.73$ in the temperature range 77–265 K was determined in the same manner. At 77 K the half-width is 20 meV and increases rapidly with temperature.

The temperature dependence of the quantum emission efficiency $\eta$ was measured in the temperature range 77–265 K for two structures with QWs: $1 - d = 50$ Å , $x = 0.73$ and $2 - d = 30$ Å , $x = 0.81$ (Fig. 2). The parameter $T_0$ was determined from this dependence using the formula

$$\eta = \eta_0 \exp(-T/T_0), \qquad (1)$$

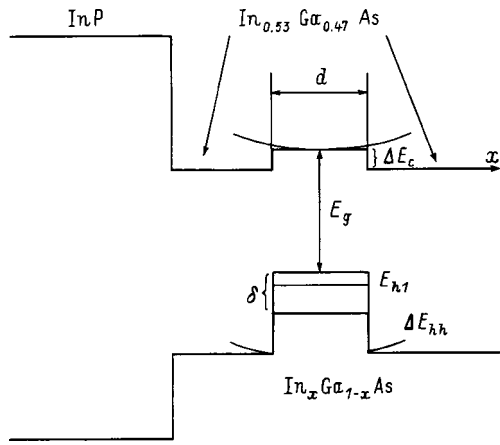FIG. 1. Band diagram of an $In_xGa_{1-x}As/In_{0.53}Ga_{0.47}As/InP$ heterostructure.



FIG. 3. Position of two energy maxima in the 77-K quantum-dot PL spectra versus the nominal thickness of InAs deposited at $T = 600\,°C$.

where $T_0 = 40$ K for case 1 and 43 K for case *2*.

Two intense peaks with $\lambda = 1.68$ and 1.88 $\mu$m appear in the PL spectra measured at 77 K with the formation of InAs QDs grown at $T = 600\,°C$. The radiative recombination efficiency from QDs is approximately an order of magnitude higher than luminescence from the $In_{0.53}Ga_{0.47}As$ matrix. Figure 3 shows the energy of these two PL peaks versus the amount of InAs deposited. It is evident that the position of the peaks remains virtually unchanged as the nominal thickness of the InAs deposited layer varies from 2.5 to 7 ML. This is different from the QW case, where the energy of the PL peak decreases with increasing layer thickness. Hence it can be concluded that according to the Stranskiĭ–Krastanov model, three-dimensional clusters develop up to their critical sizes determined by the ''equilibrium'' of the stresses between the matrix and QD materials. In our opinion, the data obtained indicate the existence of two sizes of QDs in the structure investigated.
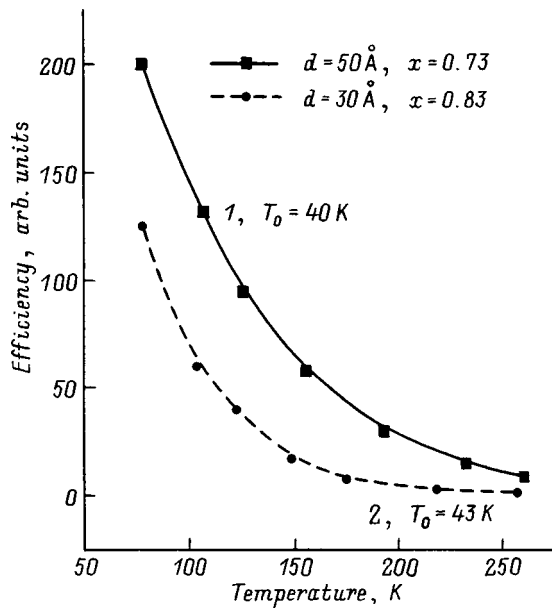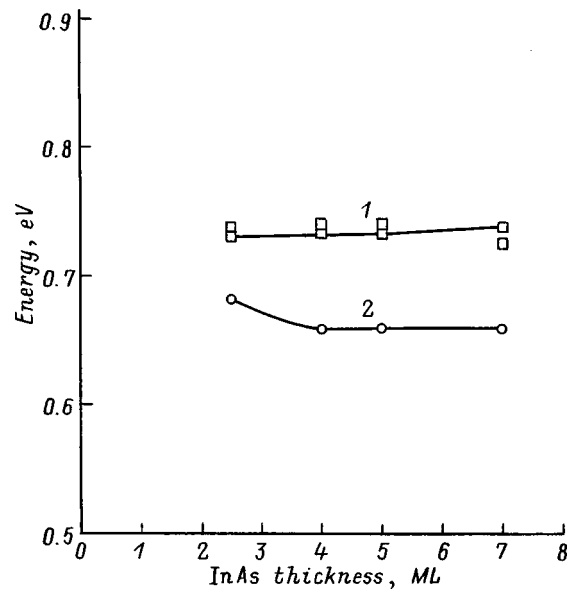
Figure 4 shows the dependence of the PL intensity at 77 K on the amount of InAs deposited for QDs grown at 600 °C. With 4–5 ML InAs the maximum intensity is observed for the peak with $\lambda = 1.68$ $\mu$m; for the $\lambda = 1.88$-$\mu$m peak the intensity remains nearly constant for 2.5–5 ML and much lower. The sharp degradation of the emission characteristics of QDs with thickness greater than 5 ML can be explained by an increase in the number of dislocations.

The spectral dependences of the PL intensity for InAs QDs grown at $T = 500$, 550, and 600 °C were obtained. At $T = 500\,°C$ a structure with QDs emitting at $\lambda = 1.91$ $\mu$m at 77 K was obtained.



FIG. 2. Temperature dependence of the external quantum efficiency $\eta$ for $In_xGa_{1-x}As$ quantum wells: *1* — $d = 50$ Å, $x = 0.73$, $T_0 = 40$ K; *2* — $d = 30$ Å, $x = 0.81$, $T_0 = 43$ K.
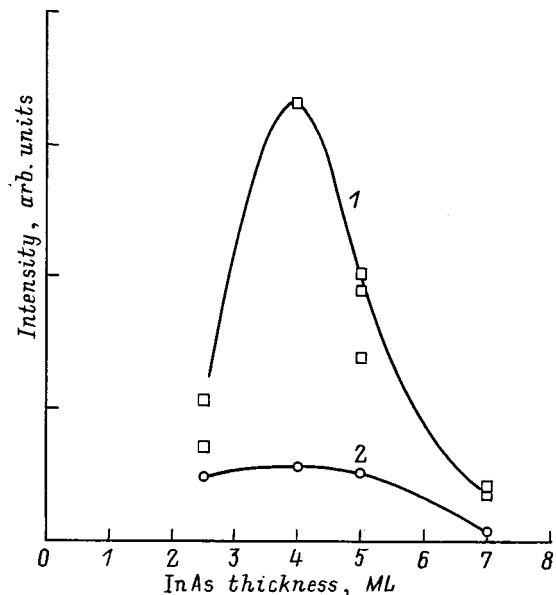


FIG. 4. Intensity of two 77-K quantum-dot PL peaks versus the nominal thickness of InAs deposited at $T = 600\,°C$. *1* — $\lambda = 1.68$ $\mu$m, *2* — $\lambda = 1.88$ $\mu$m.

## 3. COULOMB POTENTIAL WELL FOR ELECTRONS

To avoid discrepancies between the measured values of the radiation wavelengths of QW and the wavelength estimated in a very simple model[4] it is necessary to take into account that the electrons in the active region are localized in a QW which occurs as a result of the Coulomb attraction to holes. The shape of the Coulomb potential well is determined by solving Poisson's equation.[8]

The widths of the region occupied by electrons were determined by solving this equation. We estimated the electron density localized in the active region of a strained heterostructure in the temperature range 77–250 K to be 3–10 times lower than the hole density localized here. In the experimental structures with QWs, the rate of the zero-threshold Auger-recombination process[9] is also weakened, because the potential barrier for electrons is not vertical but rather varies continuously as a function of the transverse coordinate.[10]

## 4. DISCUSSION OF THE RESULTS AND CONCLUSIONS

Our experimental and theoretical studies have shown that 1.5 to 1.9-$\mu$m radiation can be obtained by using heterostructures in a system with strained quantum wells and quantum dots $In_xGa_{1-x}As/In_{0.53}Ga_{0.47}As/InP$ ($x = 0.69 - 1.0$).

For heterostructures with strained quantum wells, despite the weakening of Auger processes, as the temperature is raised, the half-width of the emission spectrum increases appreciably and the external quantum efficiency decreases. This behavior is attributable to delocalization of electrons in the active region due to the formation of a wide Coulomb potential well. As a result, the overlap integral of electrons and holes in the region of the quantum well decreases and the fraction of indirect transitions also increases. In addition, in the heterostructures which we studied there is a strong absorption of light by holes with a transition of holes into a spin-orbit splitoff band.[11]

Our study shows that efficient radiation sources cannot be produced on the basis of these heterostructures with a small offset ($\Delta E_g$) at the boundaries of the quantum well and barrier. Switching to InAs quantum dots makes it possible to advance to longer wavelengths.

*)E-Mail: zina.sokolova@pop.ioffe.rssi.ru

———————

[1] J. Dias, H. Yu, A. Rybaltowski, B. Lane, G. Lukas, D. Wu, S. Kim, M. Erdtmann, E. Kaas, and M. Razeghi, Appl. Phys. Lett. **70**, 40 (1997).
[2] D. Z. Garbuzov, R. U. Martinelli, H. Lee, P. K. York, R. J. Menna, J. C. Connoly, and S. Y. Narayan, Appl. Phys. Lett. **69**, 2006 (1996).
[3] L. Zheng, C. H. Lin, K. E. Singer, and M. Missous, IEE Proc.-J: Optoelectron. **144**, 360 (1997).
[4] A. D. Bondarev, D. A. Vinokurov, V. A. Kapitonov, O. V. Kovalenko, Z. N. Sokolova, and I. S. Tarasov, Pis'ma Zh. Tekh. Fiz. **24**, 46 (1998) [Tech. Phys. Lett. **24**, 886 (1998)].
[5] O. V. Kovalenkov, I. S. Tarasov, D. A. Vinokurov, and V. A. Kapitonov, in *Abstracts of Reports at the 6th International Symposium on ''Nanostructures: Physics and Technology,''* St. Petersburg, Russia, June 22–26 (1998), p. 268.
[6] V. M. Ustinov, A. E. Zhukov, A. F. Tsatsul'nikov, A. Yu. Egorov, A. R. Kovsh, M. V. Maksimov, A. A. Suvorova, N. A. Bert, and P. S. Kopaev, Fiz. Tekh. Poluprovodn. **31**, 1256 (1997) [Semiconductors **31**, 1080 (1997)].
[7] V. M. Ustinov, A. R. Kovsh, A. E. Zhukov, A. Yu. Egorov, N. N. Ledentsov, A. V. Lunev, Yu. M. Shernyakov, M. V. Maksimov, A. F. Tsatsul'nikov, B. V. Volovik, P. S. Kopaev, and Zh. I. Alferov, Pis'ma Zh. Tekh. Fiz. **24**, 49 (1998) [Tech. Phys. Lett. **24**, 22 (1998)].
[8] M. Silver and E. P. O'Reilly, IEEE J. Quantum Electron. **30**, 547 (1994).
[9] G. G. Zegrya and A. S. Polkovnikov, Zh. Éksp. Teor. Fiz. **113**, 1491 (1998) [JETP **86**, 815 (1998)].
[10] G. G. Zegrya and V. A. Kharchenko, Zh. Éksp. Teor. Fiz. **101**, 327 (1992) [Sov. Phys. JETP **74**, 173 (1992)].
[11] N. A. Gun'ko, V. B. Khalfin, Z. N. Sokolova, and G. G. Zegrya, J. Appl. Phys. **84**, 547 (1998).

Translated by M. E. Alferieff

# Spontaneously assembling periodic composition-modulated InGaAsP structures

L. S. Vavilova, V. A. Kapitonov, A. V. Murashova, N. A. Pikhtin, I. S. Tarasov,*
I. P. Ipatova, V. A. Shchukin, N. A. Bert, and A. A. Sitnikova

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1108–1110 (September 1999)

It is established theoretically and experimentally that in certain temperature and composition
ranges the solid solutions InGaAsP comprise a system of strained, alternating (in mutually
perpendicular directions [100] and [010]) domains of a solid solution with two different
compositions and different lattice constants. The domains are clearly seen at the surface of an
epitaxial film and wash out into its depth in the direction of the substrate. The data
obtained most likely show spinodal decomposition of InGaAsP solid solutions in the experimental
samples. © *1999 American Institute of Physics.* [S1063-7826(99)02209-7]

## 1. INTRODUCTION

In this paper we make a detailed comparison of the results of theoretical and experimental studies of the quaternary solid solutions InGaAsP in the instability region.

## 2. THEORY

Quaternary solid solutions of the type $A_{1-x}B_xC_{1-y}D_y$ with positive formation enthalpy become unstable with respect to composition fluctuations $\delta x(\mathbf{r})$ and $\delta y(\mathbf{r})$ at a critical temperature $T_c$.[1–4] The instability means that the free energy of a solid solution with modulation of the composition $x(\mathbf{r}) = \bar{x} + \delta x(\mathbf{r})$ and $y(\mathbf{r}) = \bar{y} + \delta y(\mathbf{r})$ is lower than that of uniform solid solution with the composition $\bar{x}$ and $\bar{y}$. As a result, a uniform solid solution decays into a composition-modulated periodic structure. Such a thermodynamic phase transition is called a spinodal decomposition of solid solutions.

Figure 1 shows regions of absolutely unstable (1), metastable (2), and stable (3) solid solutions, calculated at the temperature $T = 900$ K. The solid solutions in region 1 are unstable even with respect to infinitesimal composition fluctuations and the solutions in region 2 are unstable relative to composition fluctuations of finite amplitude. The computed value of $T_c$ for InGaAsP is 1275 K (Ref. 4).

The stability of an epitaxial film of the solid solution $A_{1-x}B_xC$, coherently coupled with a (001) substrate consisting of a cubic crystal, has been investigated in Refs. 5, 6, and 7. It was shown that the final state of spinodal decomposition of a solid solution is a system of strained, i.e., compressed and stretched, domains with alternating composition. Domain alternation occurs in one of the directions of smallest compression of the cubic crystal: [100], [010], and [001]. Composition modulation is maximum on a growing surface and decreases exponentially into the interior of the film ($z$ direction), as shown in Fig. 2.

## 3. EXPERIMENTAL RESULTS

**I**. The InGaAsP epitaxial layers grown on InP (001) and GaAs (001) substrates by liquid-phase epitaxy at temperatures below the critical temperature ($T_{growth} < T_c$) possess in the region of spinodal decomposition a characteristic two-band photoluminescence (PL) spectrum. As an example, such a spectrum is shown in Fig. 3 for an InGaAsP sample grown on a GaAs (001) substrate at the temperature $T_{growth} = 1023$ K.

The existence of two bands in the PL spectra of such samples suggests that the epitaxial layer is nonuniform and consists of two solid solutions with different compositions.

**II**. A periodic diffraction pattern (for example, Fig. 4a) was obtained on the same samples, using transmission electron microscopy (TEM), in the (200) reflection, which is sensitive to the composition of the solid solution.

In samples with a planar geometry, intensity (composition) modulation in the TEM image is observed in mutually perpendicular directions [100] and [010]. This is an indica-
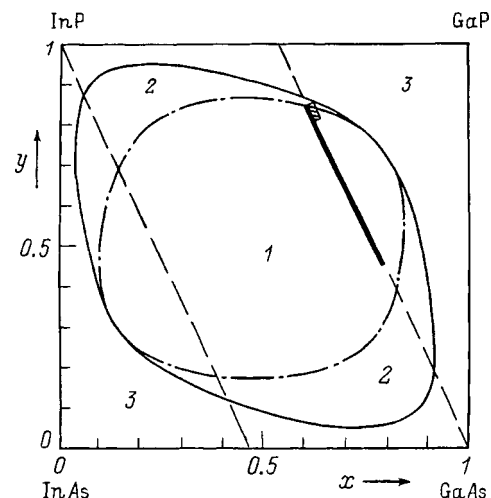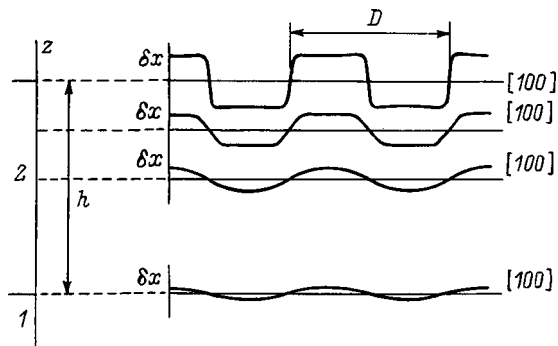


FIG. 1. Composition diagram for InGaAsP.[4]

FIG. 2. Theoretically computed equilibrium composition profile of the solid solution $A_{1-x}B_xC$ in an epitaxial film (*2*) on a substrate (*1*).[7]



FIG. 3. Photoluminescence spectrum ($T=300$ K) of a InGaAsP epitaxial layer grown in the region of instability of quaternary solid solutions.

tion of the possibility of spinodal decomposition of the solid solution.

InGaAsP epitaxial layers grown in the region of absolutely stable solid solutions have a single peak in the PL spectra. A periodic diffraction pattern is not observed in the corresponding TEM images of these samples. These two facts show that such epitaxial InGaAsP layers are uniform solid solutions.

**III.** Figure 1 shows two experimentally determined regions of spinodal decomposition of InGaAsP solid solutions which are isoperiodic with GaAs (001). The solid solutions were grown at temperatures ($T_{growth}$) 1023 K (continuous segment) and 1063 K (Ref. 8) (hatched segment). As the growth temperature is raised, the region of compositions of the InGaAsP solid solutions, for which there are, in the first place, two bands in the PL spectra (Fig. 3) and, secondly, a periodic TEM composition modulation pattern (Fig. 4a), decrease appreciably. This indicates that the growth temperature is approaching the critical temperature for spinodal decomposition.

**IV**. TEM investigations of InGaAsP/InP (001) epitaxial layers grown in the region of instability established the existence of elastic deformations at the boundaries of regions with different compositions. In samples with planar geometry, TEM patterns in the composition-sensitive (200) reflection (Fig. 4a) and in the stress-sensitive (220) reflection (Fig. 4b) were investigated. The repetition period of the alternating regions in the (200) reflection is two times smaller than in the composition-sensitive (200) reflection. It thus follows that elastic deformations exist at the boundaries of alternating regions with different compositions. The strained periodic structure possesses a nearly 100% PL quantum efficiency, which rules out residual deformation (dislocations due to lattice mismatch of the two solid solutions).

**V**. TEM investigations of the transverse section of InGaAsP/InP (001) epitaxial layers revealed the existence of a periodic structure predominantly at the surface of the film. The periodic structure washes out in a direction into the layer, and it is completely absent at the interface with the substrate (Fig. 5), which qualitatively agrees with the computed dependence shown in Fig. 2.
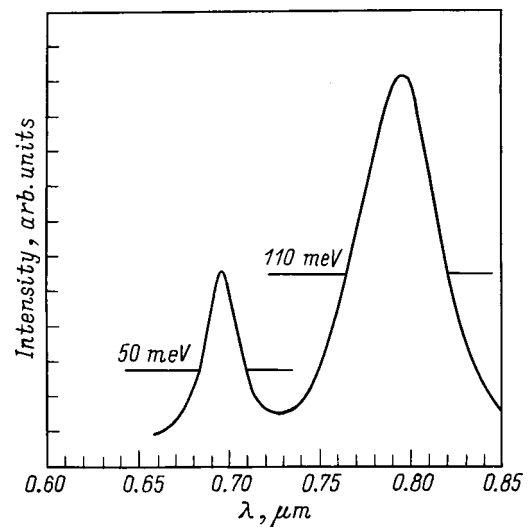


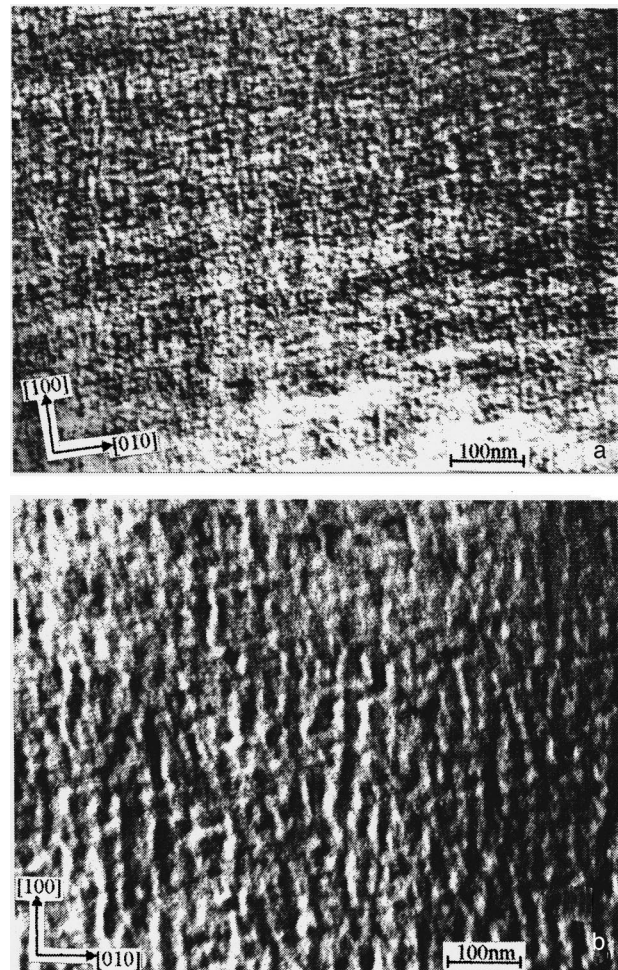FIG. 4. Photographs of an InGaAsP sample in a planar section. The photographs were obtained in an EM-420 transmission electron microscope in the reflections: a — (200) (sensitive to the composition of the solid solution) and b — (220) (sensitive to stresses in the sample). The InGaAsP epitaxial layer was grown at $T_{growth}=867$ K on an InP(001) substrate in the region of instability of solid solutions.
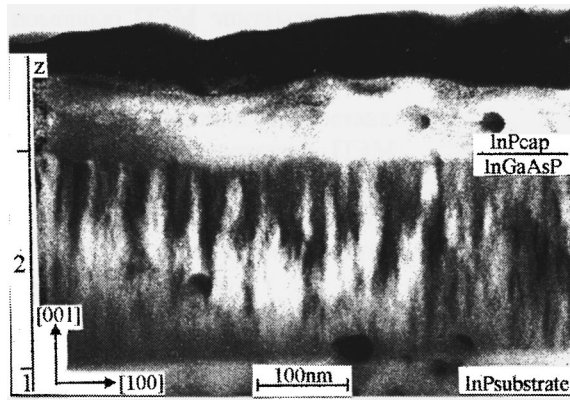
FIG. 5. Photograph of an InGaAsP sample in a transverse section. The photograph was obtained in an EM-420 transmission electron microscope. The InGaAsP epitaxial layer investigated was grown at $T_{growth} = 870$ K on an InP (001) substrate in the region of instability of solid solutions.

## 4. CONCLUSIONS

In summary, comparing the results of theoretical investigations of spinodal decomposition of solid solutions of III–V semiconductors with the study of InGaAsP samples grown by liquid epitaxy in the region of instability enabled us to conclude that in our experimental samples we most likely observed spinodal decomposition of InGaAsP solid solutions.

Our results allow us to infer that the spinodal decomposition of solid solutions can be used to obtain spontaneously assembling superlattices.

This work was supported by the Russian Fund for Fundamental Research (Project 97-02-18105).

[1] G. B. Stringfellow, J. Cryst. Growth **65**, 454 (1983).
[2] I. P. Ipatova, V. A. Shchukin, V. G. Malyshkin, A. Yu. Maslov, and E. Anastassakis, Solid State Commun. **78**, 19 (1991).
[3] B. de Cremoux, J. Physique **43**, C5–19 (1982).
[4] I. P. Ipatov, V. G. Malyshkin, A. Yu. Maslov, and V. A. Shchukin, Fiz. Tekh. Poluprovodn. **27**(2), 259 (1993) [Semiconductors **27**, 158 (1993)].
[5] I. P. Ipatova, V. G. Malyshkin, and V. A. Shchukin, J. Appl. Phys. **74**, 7198 (1993).
[6] I. P. Ipatova, V. G. Malyshkin, and V. A. Shchukin, Philos. Mag. **70**, 557 (1994).
[7] D. Bimberg, I. P. Ipatova, P. S. Kop'ev, N. N. Ledentsov, V. G. Malyshkin, and V. A. Shchukin, Usp. Fiz. Nauk **167**(3), 552 (1997).
[8] Seiji Mukai, J. Appl. Phys. **54**(5), 2635 (1983).

Translated by M. E. Alferieff

# Gain characteristics of quantum-dot injection lasers

A. E. Zhukov,*) A. R. Kovsh, V. M. Ustinov, A. Yu. Egorov, N. N. Ledentsov,
A. F. Tsatsul'nikov, M. V. Maksimov, S. V. Zaĭtsev, Yu. M. Shernyakov, A. V. Lunev,
P. S. Kop'ev, and Zh. I. Alferov

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

D. Bimberg

*Institut fur Festkorperphysik, Technische Universitat Berlin, D-10623 Berlin, Germany*

The current dependence of the optical gain in lasers based on self-organized InGaAs quantum
dots in a AlGaAs/GaAs matrix is investigated experimentally. A transition from lasing
via the ground state of quantum dots to lasing via an excited state is observed. The saturated
gain in the latter case is approximately four times greater than for the ground state.
This result is attributable to the fourfold degeneracy of the excited level of quantum dots. The
effect of the density of the quantum-dot array on the threshold characteristics is
investigated. A lower-density array of dots is characterized by a lower threshold current density
in the low-loss regime, because the transmission current is lower, while dense quantum-
dot arrays characterized by a high saturated gain are preferable at high threshold gains. © *1999
American Institute of Physics.* [S1063-7826(99)02309-1]

As predicted theoretically,[1,2] injection lasers with an active region based on quantum dots (QDs) make it possible to decrease substantially the threshold current density $J_{th}$. A finite surface density $N_{QD}$ of a QD array actually signifies a low transmission current because of the small number of states on which a population inversion must be maintained. On the other hand, a finite value of $N_{QD}$ combined with inhomogeneous broadening of the density of states could lead to gain saturation. Progress in producing QD arrays by self-organization of strongly strained thin layers[3] has made it possible to obtain in long laser diodes based on them $J_{th}$ on the order of $60-100$ A/cm$^2$ at 300 K (Ref. 4). However, higher losses in the laser diode led to superlinear growth of $J_{th}$,[5,6] indicating that gain saturation is indeed a serious problem in lasers based on self-organized QDs.

In the present paper we report the results of an experimental study of the gain characteristics of injection lasers based on InGaAs QDs in an AlGaAs/GaAs host. The laser structures were grown by molecular-beam epitaxy on $n^+$-GaAS (100) substrates in a laser geometry with separate confinement (SC DHS). The active region is formed by three rows of InGaAs QDs separated by 5-nm-thick Al$_{0.15}$Ga$_{0.85}$As spacer layers. The laser diodes were fabricated in a fine-mesa geometry, and the device characteristics were investigated in a pulsed regime at 300 K. The growth regime and the measurements are described in greater detail in Ref. 7.

The dependence of the gain on the current density is one of the most important characteristics of an injection laser, and it is closely related to the electronic structure of the active medium. The mode gain ($g$) as a function of the pump current was calculated from the experimental dependence of $J_{th}$ on the cavity length, using the fact that $g$ is the sum of the losses (internal and on the mirrors) at the lasing thresh-

old. The internal losses were estimated as an adjustable parameter from the experimental dependence of the external differential efficiency on the cavity length.

Figure 1 (dots) shows the experimental dependence of $g$ and the lasing wavelength $\lambda_{las}$ on the pump current density. It is evident that both curves have a kink at a current density of about 700 A/cm$^2$ that separates regions of different gain behavior accompanied by a sharp hop of $\lambda_{las}$. In each section, an increase in the pumping results in rapid gain saturation at a certain level $g^{sat}$, while $\lambda_{las}$ gradually shifts in the short-wavelength direction.
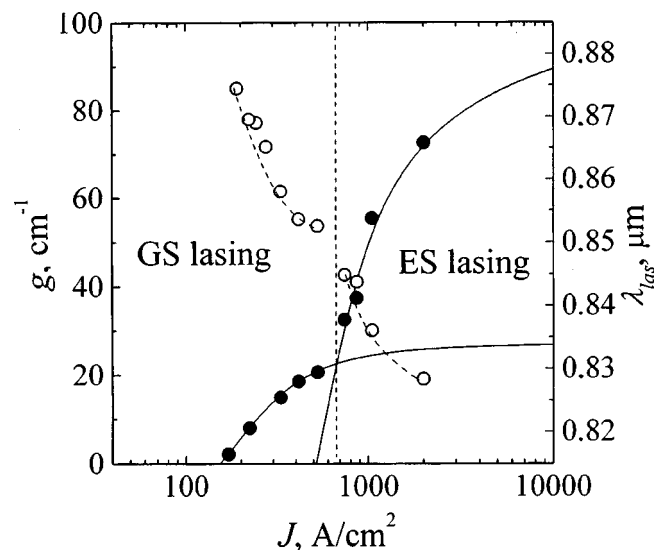


FIG. 1. Experimental curves of the mode gain $g$ (filled dots) and lasing wavelength $\lambda_{las}$ (open dots) versus the current density. Solid lines — calculation of $g$ using Eq. (1) and the gain parameters presented in Table I.

We observed two peaks, at 855 and 960 nm, in the photoluminescence (PL) spectrum of this structure. The long-wavelength peak is due to radiative recombination of excitons localized in the ground state (GS) of a QD, while the short-wavelength peak corresponds to recombination through an excited state (ES).[8] Comparing the pump dependence of $\lambda_{las}$ with the spectral position of these lines shows that the observed current dependence of the gain is due to a transition from lasing via the GS to lasing via the ES of the QD array.

We have shown previously[9] that in a simplified form the relation between the gain and the current density in a QD laser can be described by the expression

$$J = J_0\left(1 + \frac{g}{g^{sat}}\right) + J_1\left(\frac{g^{sat}+g}{g^{sat}-g}\right) + J_2\left(\frac{g^{sat}+g}{g^{sat}-g}\right)^2. \quad (1)$$

For lasing via the GS the first two terms in Eq. (1) describe the contributions of recombination via the GS and the ES, respectively, while the last term corresponds to recombination via a state of the wetting layer (WL) and the matrix. If lasing occurs via the ES, then the second term in Eq. (1) is absent, and the first term becomes the combined contribution of both the GS and the ES.

According to Eq. (1), the transmission current density $J_{tr}$ corresponding to zero gain is given by

$$J_{tr} = J_0 + J_1 + J_2, \quad (2)$$

and for a small gain ($g \ll g^{sat}$) the dependence of the gain on the pump current is approximately linear

$$g_{QD} = g^{sat}\frac{J - J_{tr}}{J_{tr}}\gamma, \quad \text{where} \quad \gamma = \frac{J_{tr}}{J_{tr}+J_2}. \quad (3)$$

It is shown in Fig. 1 that for an appropriate choice of the parameters $J_0$, $J_1$, $J_2$, and $g^{sat}$, presented in Table I, the expression (1) makes it possible to describe to a high degree of accuracy the current dependence of the gain for lasing via the GS and the ES.

According to the simple hydrogen-like model of the electronic structure of a QD,[10] the ES is fourfold degenerate (in addition to the spin degeneracy). Four times as many carriers therefore can be localized in the ES; consequently $g^{sat}$ for lasing via the ES should be four times greater than $g^{sat}$ for the GS provided that the GS and ES have the same oscillator strength. In general, the degeneracy of the excited level is lifted because the potential well of a QD is not described by a Coulomb potential and real QDs are not spherical. However, the experimentally observed ratio of $g^{sat}$ is indeed close to 4, indicating that the splitting of the ES is much smaller than the inhomogeneous broadening of the density of states of the QD as a result of size variance.
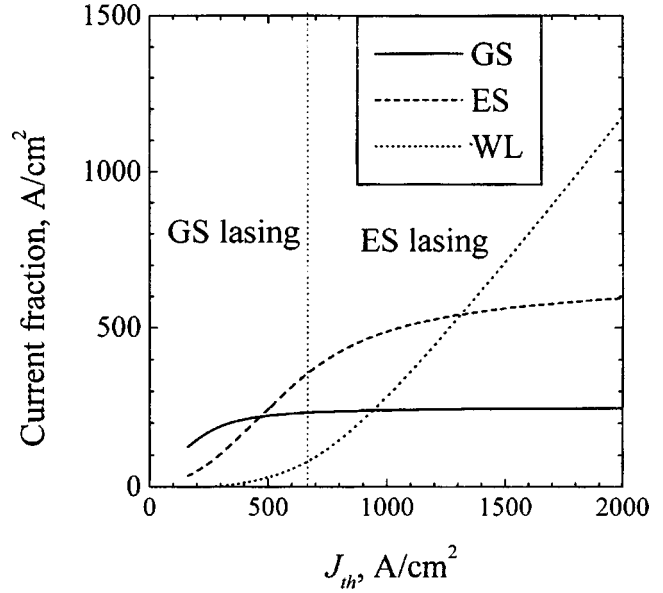


FIG. 2. Contributions due to recombination via the ground state (GS) and excited state (ES) of QDs and states of the wetting layer (WL) to the threshold current density $J_{th}$.

According to the data presented in Table I, the fraction of the transmission current going to recombination via the WL ($J_2$) in the case of lasing via the GS is negligible compared with lasing via an ES. Correspondingly, according to Eqs. (2) and (3), $\gamma$ in the latter case decreases. This effect can be explained by the fact that to achieve lasing via the ES the Fermi level must lie much higher and therefore more carriers will be located in the WL.

Figure 2 shows the contributions to $J_{th}$ from recombination via various states as a function of the threshold current density. For small $J_{th}$ the current density is mainly determined by recombination via the GS. As $J_{th}$ increases, the contribution of the GS saturates, and the current due to recombination via an ES starts to dominate. At the same time, the fraction of the current associated with the WL increases. As $J_{th}$ increases further, the fraction of the current going into recombination in the WL becomes dominant, and the contribution of the ES saturates, but at a higher level compared with the GS, since a larger number of carriers can be localized in the ES. However, even with strong pumping lasing again occurs via the states of the QD, and a transition to lasing via the WL is not observed experimentally.

The expression (1) can be used to estimate the effect of the surface density $N_{QD}$ of QDs on the threshold characteristics of QD lasers. Assuming that the electronic and radiative properties of the QD array do not change, a change in $N_{QD}$ by a factor of $N$ results in an $N$-fold scaling of $g^{sat}$ and also of the contributions to the transmission current due to the GS ($J_0$) and the ES ($J_1$). If $N_{QD}$ changes as a result of a change in the number of deposited QD layers,[11] then the parameter $J_2$ also scales by a factor of $N$ since each QD row lies on the WL itself. If special growth regimes[7] are used to change directly the surface density of QDs while keeping the number of deposited planes constant, then $J_2$ remains the same:

TABLE I. Gain parameters of a laser based on self-organized QDs.

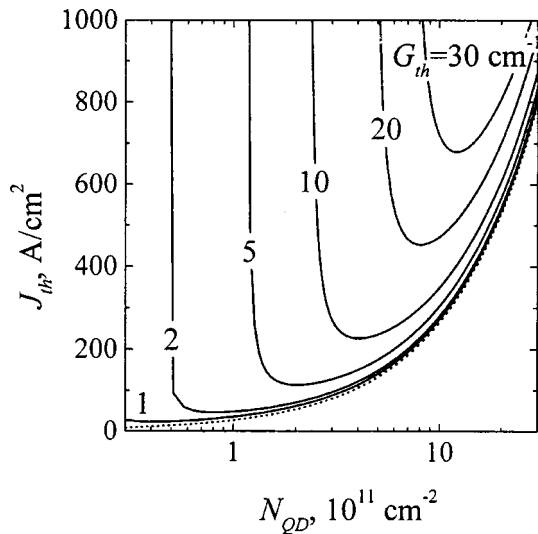| State | $g^{sat}$, cm$^{-1}$ | $J_0$, A/cm$^2$ | $J_1$, A/cm$^2$ | $J_2$, A/cm$^2$ | $J_{tr}$, A/cm$^2$ | $\gamma$ |
|---|---|---|---|---|---|---|
| GS | 27 | 119 | 38 | 0.6 | 157.6 | 0.997 |
| ES | 102 | 480 | – | 33 | 513 | 0.93 |

FIG. 3. Threshold current density $J_{\text{th}}$ versus the surface density $N_{QD}$ of a QD array for various threshold gains $G_{\text{th}}$. (The dotted line corresponds to the transmission current density.)

$$J(N) = NJ_0 + \left(1 + \frac{g}{Ng^{\text{sat}}}\right) + NJ_1\left(\frac{Ng^{\text{sat}} + g}{Ng^{\text{sat}} - g}\right)$$
$$+ NJ_2\left(\frac{Ng^{\text{sat}} + g}{Ng^{\text{sat}} - g}\right)^2. \tag{4}$$

The $N_{QD}$ dependence of $J_{\text{th}}$, calculated according to Eq. (4) using the gain parameters presented in Table I, is shown in Fig. 3 for various values of the total losses ($G_{\text{th}}$). It follows from Fig. 3 that in the low-gain regime a lower-density QD array is characterized by a lower value of $J_{\text{th}}$, which is attributable to the lower transmission current. On the other hand, when high gains are required, a higher-density QD array makes it possible to decrease $J_{\text{th}}$ due to an increase in $g^{\text{sat}}$. If the required gain is greater than $g^{\text{sat}}$ for a certain QD density, then lasing via the QD states cannot be attained at any current density.

It is evident that there exists an optimal density $N_{\text{opt}}$ corresponding to a minimum value of $J_{\text{th}}$ for a given level of losses. For most device applications, $G_{\text{th}}$ is relatively large, since short diodes are required in order to obtain high differential efficiency. For this reason, higher-density QD arrays, which make it possible to overcome the gain-saturation to a certain degree, are more desirable for practical applications.

In summary, the threshold characteristics of lasers based on InGaAs QDs in an AlGaAs/GaAs host were investigated.

Two characteristic sections are observed on the curves of the optical gain and lasing wavelength versus the current density. These sections are attributable to lasing via the ground state and the excited state of the QDs. As the current is raised, the gain attained on the GS saturates, and a transition to lasing via an ES occurs. As a result of the degeneracy of the ES, the saturated gain in the latter case is approximately four times greater than the corresponding value for the GS. As the QD density increases, the transmission current and the saturated gain increase. As a result, a lower-density QD array gives a lower threshold current density in the low-loss regime, while for large threshold gains dense QD arrays are preferable.

*)E-Mail: Zhukov@beam.ioffe.rssi.ru

[1] Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).

[2] M. Asada, Y. Miyamoto, and Y. Suematsu, J. Quant. Electron. **QD-22**, 1915 (1986).

[3] L. Goldstein, F. Glas, J. Y. Marzin, M. N. Charasse, and G. Le Roux, Appl. Phys. Lett. **47**, 1099 (1985).

[4] V. M. Ustinov, A. Yu. Egorov, A. R. Kovsh, A. E. 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. Bohrer, D. Bimberg, A. O. Kosogov, P. Werner, and U. Gosele, J. Cryst. Growth **175/176**, 689 (1997).

[5] S. V. Zaitsev, N. Yu. Egorov, Yu. M. Sherniakov, V. M. Ustinov, A. E. Zhukov, A. Yu. Egorov, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, N. N. Ledentsov, N. Kirstaedter, and D. Bimberg, Superlattices Microstruct. **21**, 559 (1997).

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

[7] A. R. Kovsh, A. E. Zhukov, A. Yu. Egorov, V. M. Ustinov, Yu. M. Shernyakov, M. V. Maksimov, A. F. Tsatsul'nikov, B. V. Volovik, A. V. Lunev, N. N. Ledentsov, P. S. Kop'ev, Zh. I. Alferov, and D. Bimberg, Fiz. Tekh. Poluprovodn. **32**, 1114 (1998) [Semiconductors **32**, 997 (1998)].

[8] A. E. Zhukov, A. R. Kovsh, V. M. Ustinov, A. Yu. Egorov, N. N. Ledentsov, A. F. Tsatsul'nikov, M. V. Maximov, Yu. M. Shernyakov, V. I. Kopchatov, A. V. Lunev, P. S. Kop'ev, D. Bimberg, and Zh. I. Alferov, Semicond. Sci. Technol. **14**, 118 (1999).

[9] A. R. Kovsh, A. E. Zhukov, A. Yu. Egorov, V. M. Ustinov, N. N. Ledentsov, M. V. Maksimov, A. F. Tsatsul'nikov, and P. S. Kop'ev, Fiz. Tekh. Poluprovodn. **33**, 215 (1999) [Semiconductors **33**, 184 (1999)].

[10] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: The Nonrelativistic Theory* [Pergamon Press, New York; Nauka, Moscow, 1989, 300 pp.].

[11] N. N. Ledentsov, V. A. Schukin, M. Grudmann, N. Kirstaedter, J. Bogrer, O. Schmidt, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, S. V. Zaitsev, Zh. I. Alferov, A. I. Borovkov, A. O. Kosogov, S. S. Ruvimov, P. Werner, U. Gosele, and J. Heydenreich, Phys. Rev. B **54**, 8743 (1996).

Translated by M. E. Alferieff

# Blue-green ZnSe lasers with a new type of active region

S. V. Ivanov,*) A. A. Toropov, S. V. Sorokin, T. V. Shubina, I. V. Sedova, P. S. Kop'ev, and Zh. I. Alferov

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

A. Waag, H. J. Lugauer, G. Reuscher, M. Keim, F. F. Fischer, and G. Landwehr

*Institute of Physics, University of Würzburg, Würzburg, Germany*

We report the results of an experimental study of molecular-beam epitaxy of ZnSe-based laser heterostructures with a new structure of the active region, which contains a fractional-monolayer CdSe recombination region in an expanded ZnSe quantum well and a waveguide based on a variably-strained, short-period superlattice are reported. Growth of a fractional-monolayer CdSe region with a nominal thickness of 2–3 ML, i.e., less than the critical thickness, on a ZnSe surface ($\Delta a/a \sim 7\%$) leads to the formation of self-organized, pseudomorphic, CdSe-enriched islands with lateral dimensions $\sim 10-30$ nm and density $\sim 2 \times 10^{10} \, \text{cm}^{-2}$, which serve as efficient centers of carrier localization, giving rise to effective spatial separation of defective regions and regions of radiative recombination and, as a result, a higher quantum efficiency. Laser structures for optical pumping in the (Zn, Mg) (S, Se) system with a record-low threshold power density (less than 4 kW/cm$^2$ at 300 K) and continuous-wave laser diodes in the system (Be, Mg, Zn) Se with a 2.5 to 2.8-ML-thick, fractional-monolayer CdSe active region have been obtained. The laser structures and diodes have an improved degradation resistance. © *1999 American Institute of Physics.*
[S1063-7826(99)02409-6]

## 1. INTRODUCTION

Despite the great efforts made recently to optimize laser diodes with ZnSe quantum wells, increasing the lifetime of blue–green II–VI lasers to a commercially useful level remains an unsolved problem.[1] Just as in GaAs-based laser diode structures, degradation is the main reason for the propagation and multiplication of defects in the active region, even if the initial (immediately after growth) defect density does not exceed $10^4 \, \text{cm}^{-2}$. One way to overcome the degration problem is to develop a new structure of the active region which could give: 1) better optical and electronic (hole) confinement and 2) higher activation energy for formation and propagation of extended defects in the active region. It has been shown that a variably-strained, short-period superlattice (SL) used as a waveguide is a very effective method for solving this problem.[2–4] Another possibility of improving the characteristics of ZnSe laser structures is to decrease the dimension of the active medium from 2D [quantum wells (QWs)] to 0D [quantum dots (QDs)]. Aside from the obvious advantages of a delta-function density of states, which can be obtained, nonetheless, by producing a uniform size distribution of extremely small (the Bohr radius of an exciton in ZnSe is $a_B \sim 3$ nm) 0D objects, the introduction of centers of three-dimensional localization can lead to spatial separation of carrier localization locations and defective locations in the active region of the laser structure.[5] It can then be expected that the internal quantum yield will increase and the rate of development of defects during lasing will de-

crease, since nonradiative recombination on defects, as shown in Ref. 6, is the main reason for defect multiplication and migration. Finally, great hopes for improving the degradation characteristics are based on the use of Be-containing compounds, which among wide-gap II–VI compounds have the hardest lattice.[7]

To obtain lasing in the blue–green region of the spectrum (460–550 nm), corresponding to the maximum sensitivity of the human eye, a great deal of attention has recently been devoted to the study of the heteropair CdSe/ZnSe. This pair is similar in many respects to the well-investigated heteropair InAs/GaAs, where a large lattice mismatch ($\sim 7\%$) is the driving force for the formation of self-organized (In, Ga)As QDs. Until recently, predominantly two technological approaches have been used to produce such nanoobjects based on (Cd, Zn)Se: 1) the Stranskiĭ–Krastanov (SK) method of growth with the formation of relatively large 3D islands with CdSe layer thickness greater than the critical value [$\sim 3$ monolayers (ML)][8–10] and 2) the method where incomplete CdSe monolayers are deposited on a ZnSe surface.[11] The formation of CdSe nanoislands in a ZnSe host by the SK method is ordinarily impeded because strain relaxation via defect formation, which is accompanied by a sharp drop in the photoluminescence (PL) intensity, predominates. The second approach is based on growing ultrathin CdSe QWs, which to a first approximation can be viewed as layers of the solid solution ZnCdSe formed on a CdSe/ZnSe interface.[12] Recent optical investigations have demonstrated that such layers have a more complicated in-

ternal structure, including flat CdSe islands of uniform thickness (~1 ML).[11,13,14] High-resolution transmission electron microscopy (TEM) investigations of CdSe/ZnSe SLs in the cross-sectional geometry have also confirmed the formation of nanosize CdSe islands.[15] It should also be noted that previous attempts to use CdSe/ZnSe SLs in the active region of laser structures did not lead to appreciable improvement of the laser characteristics.[16–18]

We report here the results of investigations on the development of room-temperature blue-green lasers based on II–VI compounds in the systems (Zn, Mg)(S, Se) and (Be, Mg, Zn)Se. A new laser structure is proposed and implemented, including a waveguide in the form of variably-strained, short-period SLs (of the type ZnSSe/ZnCdSe or BeZnSe/ZnCdSe) combined with an active region based on fractional-monolayer (FML) CdSe inserts (2–3 ML thick) placed in a 10-nm ZnSe QW. Isolated FML CdSe/ZnSe structures were investigated, and special attention was devoted to the region of subcritical nominal CdSe thicknesses $w$ ($w = 1 - 3$ ML). Island formation in FML structures grown by conventional molecular-beam epitaxy (MBE) and epitaxy with enhanced surface migration (ESM) was observed from the surface by TEM. It is shown that isolated FML CdSe/ZnSe structures can be used as an active region for ZnSe-based green lasers, including the lasers containing Be chalcogenides.

## 2. EXPERIMENT

The structures of (Zn, Mg) (S, Se) lasers for optical pumping and (Be, Mg, Zn)Se laser diodes with an active region based on an FML CdSe/ZnSe nanostructure, surrounded by an SL waveguide, were grown by MBE pseudomorphically on GaAs(001) substrates at growth temperatures 270–280 °C. The MBE growth parameters and methods for monitoring the composition of S- and Be-containing compounds have been published previously.[19,20] The structures of lasers for optical pumping contained, respectively, 0.5 and 0.1-$\mu$m-thick $Zn_{0.9}Mg_{0.1}S_{0.15}Se_{0.85}$ confining layers at the bottom and top, a symmetric waveguide in the form a 3-nm $ZnS_{0.14}Se_{0.86}$/5-nm ZnSe SL with a total thickness of 0.2 $\mu$m, and an FML CdSe/ZnSe active region with CdSe thickness $w = 2.8$ ML at the center of a 10-nm ZnSe QW. The active region of the BeMgZnSe/ZnCdSe laser diode consisted of a (1-nm $Be_{0.05}Zn_{0.95}$Se/1.5-nmZnSe)$_{82}$ SL waveguide with 10-nm ZnSe QW at the center, containing 2.6 ML CdSe. The laser-diode structures also included wide-gap, 1-$\mu$m-thick, $n$- and $p$-$Be_{0.05}Mg_{0.06}Zn_{0.91}$Se confining layers, doped with iodine and nitrogen, respectively, and a ZnSe/BeTe doping-modulated short-period SL with a variable step and strongly $p$-doped 10-nm BeTe:N/30-nm ZnTe:N layers on top. The characteristics of the waveguides based on variably-strained, short-period SLs have been investigated previously.[2–4]

The characteristic features of the growth of an FML CdSe active region by MBE and ESM are described in Ref. 11. For structural and optical characterization, a series of isolated FML CdSe/ZnSe structures with $w = 0.25 - 3.2$ ML were grown under the same MBE conditions as the active regions of the laser structures. These structures contained a

buffer layer (50 nm) and (10 nm) ZnSe layer on top. In addition, for comparison laser structures with conventional QWs at the center of the SL waveguide were grown for optical (7-nm $Zn_{0.75}Cd_{0.25}$Se QW) and injection (4-nm $Zn_{0.63}Cd_{0.37}$Se QW) pumping.

Surface images were obtained by TEM using a Philips EM420 electron microscope. The PL spectra of the structures
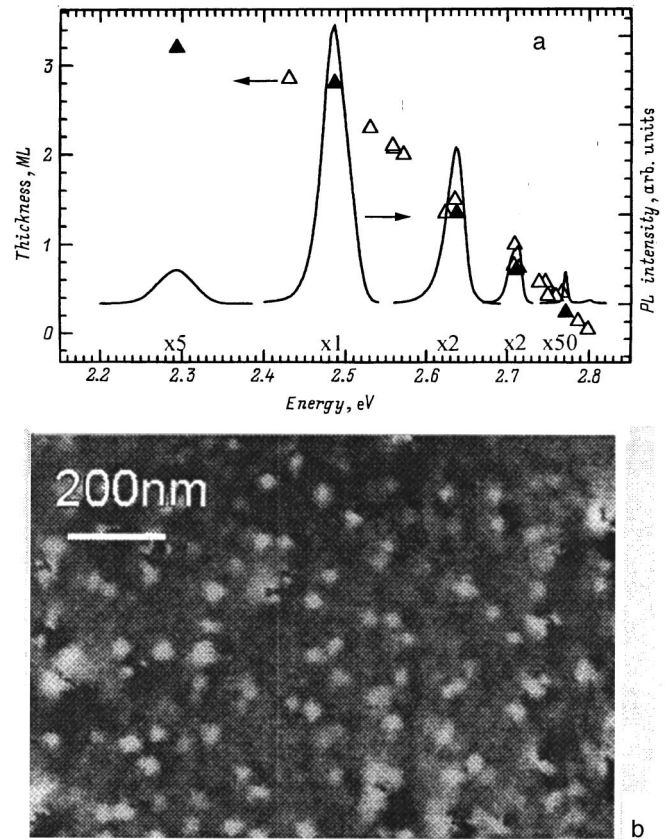


FIG. 1. Relation between the nominal CdSe thickness and the energy position of the PL peak. a) PL spectra for samples marked by the filled triangles. b) TEM image from a surface for a structure with a single CdSe FML insert with nominal thickness 2.8 ML.
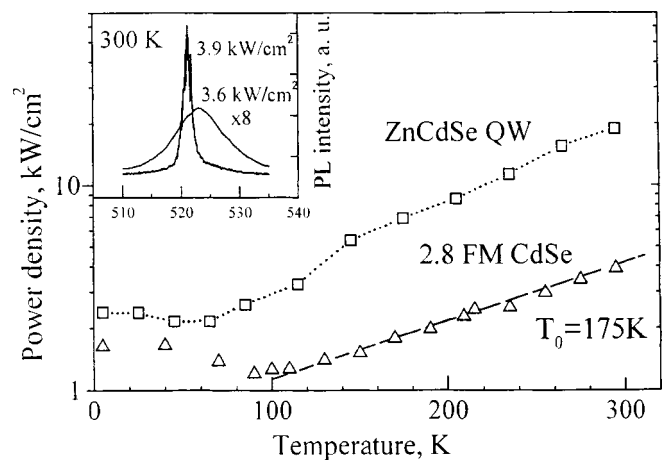


FIG. 2. Temperature dependence of the threshold power density for laser structures with an FML active region (triangles) and a QW (squares). Inset: Lasing spectra of an FML structure at 300 K.
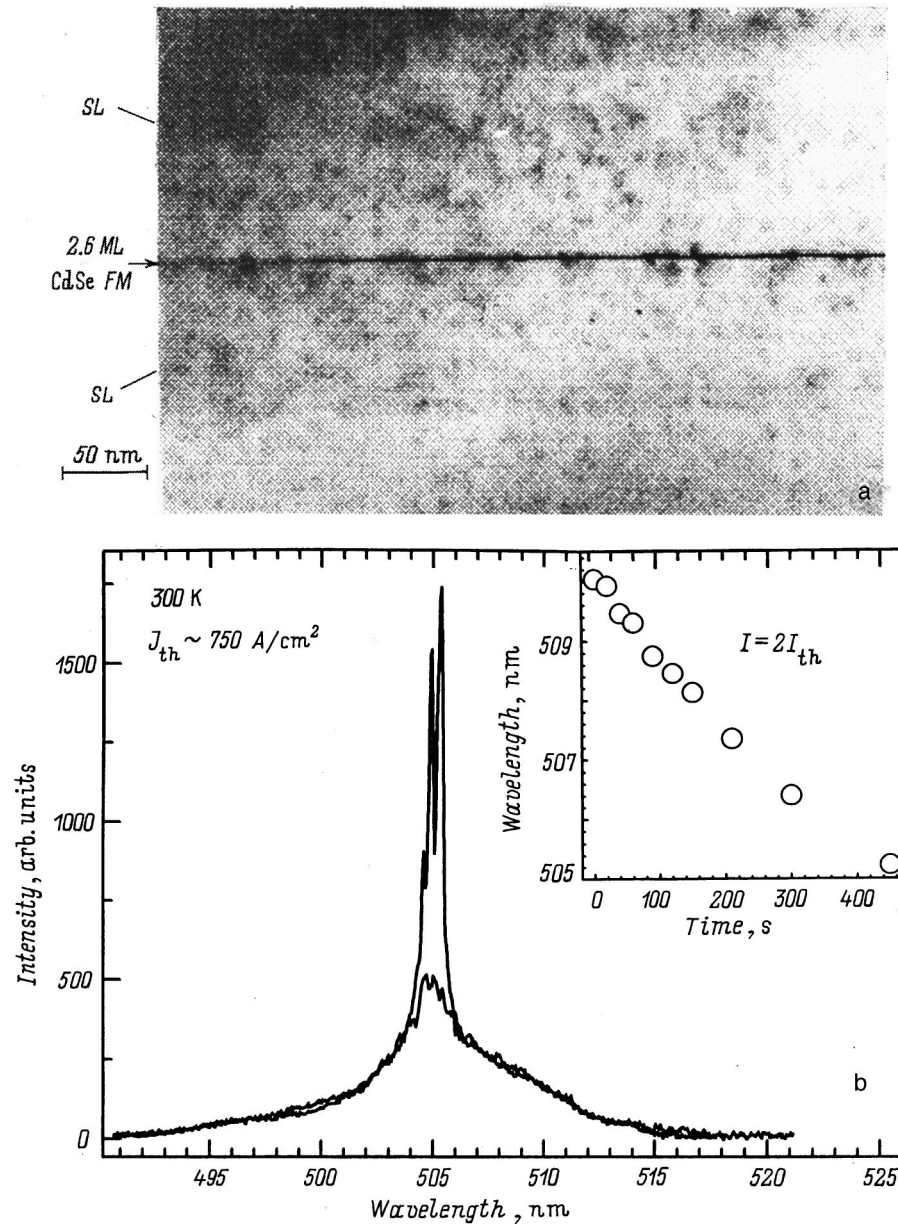
FIG. 3. a) Image of a BeMgZnSe laser structure with an FML active region. The image was obtained by TEM in the cross-sectional geometry. b) Electroluminescence spectra (300 K) of this structure with continuous excitation above and below threshold. Inset: Shift of the lasing wavelength in time under pulsed excitation.

were measured at 5 K using the 351-nm line of a continuous-wave $Ar^+$ laser. The measurements of the PL excitation (PLE) spectra were performed with a halogen lamp. The laser characteristics with optical pumping were investigated using excitation with a $N_2$ laser with 8-ns pulse duration. The structures of injection laser diodes with flat contacts were obtained by conventional photolithography and measured by the methods described in Ref. 20.

## 3. RESULTS AND DISCUSSION

The relationship between the nominal thickness of the CdSe layer and the energy position of the maximum in the PL spectrum for isolated FML CdSe/ZnSe structures is shown in Fig. 1a. Increasing $w$ to values not exceeding the critical thickness results in a rapid increase of the PL efficiency, accompanied by a substantial red shift of the PL peak. The PL intensity with the same pumping increases by more than a factor of 350 with a tenfold increase in the

amount of CdSe deposited. Above the critical thickness a sharp (by more than a factor of 30) drop in the PL intensity is observed.

The results of optical measurements are in good agreement with the TEM data from the surface (Fig. 1b) for an isolated FML structure with the same $w = 2.8$ ML. Extended CdSe-enriched islands appear in the TEM image as light spots with lateral dimensions 10–40 nm and density $2 \times 10^{10}$ cm$^{-2}$. Although large (>30 nm) islands relaxed with formation of clearly distinguishable, identically oriented defects, their density is only ~15% of the total island density, while the remaining 85% (with lateral dimensions <30 nm) seem to be pseudomorphic. We note that the appearance of CdSe islands with lateral dimensions of 25–40 nm is typical for thicknesses $w$ greater than 0.7 ML, and the fraction of relaxed islands is virtually independent of $w$,[21] indicating that they are related to the initial nonuniformities of the growth surface of ZnSe (defects, growth steps, impurities),

giving rise to a predominant accumulation of Cd. Some islands in Fig. 1b demonstrate a square faceting, but the strong contrast, due to deformations, makes it impossible to determine their shape and real sizes accurately. As a result of the distortions of the image by the ZnSe layer at the top, the real dimensions of the islands can be less than those observed but definitely much greater than $a_B$. We believe that these islands are not electronic quantum dots, but rather they are extended sections of ultrathin QWs with a specific carrier-localization potential arising as a result of local thickness or Cd content fluctuations.

The use of such a CdSe/ZnSe FML active region in the ZnMgSSe laser structure led to a large decrease of the threshold pump power density at 300 K to a level below $P_{th} = 4$ kW/cm$^2$ ($\lambda \sim 523$ nm), compared with 19 kW/cm$^2$ for a laser structure with the conventional ZnCdSe QW, emitting at the same wavelength.[21] The temperature dependence of the threshold power density for both types of structures and the room-temperature lasing spectra of structures with an FML CdSe insert are shown in Fig. 2. For both structures the temperature dependence is nonmonotonic with a minimum in the range 80–100 K, which, as shown in Ref. 22, is due to a thermally activated mechanism of vertical hole transport in the SL waveguide. The structure with FML CdSe demonstrates a 2.5-fold increase in threshold on passing from liquid-helium (5 K) to room temperature (295 K) with characteristic temperature $T_0 \sim 175$ K at 300 K, while for the control structure with a QW the threshold increases by a factor of 4.75. In addition, the FML laser structure did not show appreciable degradation for 24 h with pump density 25 times greater than the threshold value ($\sim 100$ kW/cm$^2$), while the structure with a QW and the same post-growth density of extended defects under the same conditions degraded completely in one hour.

Figure 3 shows the characteristic electroluminescence spectra of a Be-containing continuous-wave laser diode with 2.6-ML CdSe/ZnSe FML active region, operating at room temperature, and a TEM image of the active region in cross-sectional geometry. The thickness modulations of the FML region are probably due to fluctuations of the strain fields caused by CdSe islands randomly distributed in the plane, as observed previously by TEM from the surface (see Fig. 1b). An appreciable shift of the lasing wavelengths with time, accompanied by approximately a twofold increase in the threshold current density, was observed. It can be inferred that this phenomenon, which is not observed in the control structure with a ZnCdSe QW, is due to a gradual degradation of large islands of nearly critical size with deeper-lying energy levels. Preliminary degradation tests also revealed that the FML laser diodes are more stable than QW samples with the same density of post-growth extended defects and optical confinement.

## 4. CONCLUSIONS

Pseudomorphic CdSe-enriched islands with lateral dimensions 10–30 nm and density $> 2 \times 10^{10}$ cm$^{-2}$, demonstrating a high internal quantum yield of radiative recombination, were observed to form in MBE-grown fractional-monolayer CdSe/ZnSe structures with subcritical CdSe-layer nominal thickness ($\lesssim 3.0$ ML). These results were used to produce laser structures for optical and injection pumping. We believe that the record-low values obtained for the threshold power density of optically pumped lasers and their higher degradation resistance are due to effective carrier localization in these nanoislands, which prevents carrier diffusion to larger relaxed islands and other defective regions, where they can recombine nonradiatively and give rise to accelerated multiplication of defects in the active region. Continuous-wave BeMgZnSe injection lasers (at 300 K) with an FML CdSe/ZnSe active region and an SL waveguide were also obtained.

*)E-Mail: ivan@beam.ioffe.rssi.ru

[1] E. Kato, H. Noguchi, M. Nagai, H. Okuyama, S. Kijima, and A. Ishibashi, Electron. Lett. **34**, 282 (1998).

[2] T. V. Shubina, S. V. Ivanov, A. A. Toropov, G. N. Aliev, M. G. Tkatchman, S. V. Sorokin, N. D. Il'inskaya, and P. S. Kop'ev, J. Cryst. Growth **184/185**, 596 (1998).

[3] S. Ivanov, A. Toropov, S. Sorokin, T. Shubina, A. Lebedev, P. Ko'pev, Zh. Alferov, H.-J. Lugauer, G. Reuscher, M. Keim, F. Fischer, A. Waag, and G. Landwehr, Appl. Phys. Lett. **73**, 2104 (1998).

[4] S. V. Ivanov, A. A. Toropov, S. V. Sorokin, T. V. Shubina, N. D. Il'inskaya, A. V. Lebedev, I. V. Sedova, P. Ko'pev, Zh. Alferov, H.-J. Lugauer, G. Reuscher, M. Keim, F. Fischer, A. Waag, and G. Landwehr, Fiz. Tverd. Tela (St. Petersburg) **32**, 1272 (1998) [Phys. Solid State **32**, 1137 (1998)].

[5] J. M. Gerald and O. Sermage, Appl. Phys. Lett. **68**, 3123 (1996).

[6] K. Nakano, in *Proceedings of the 2nd International Symposium on Blue Laser and Light Emitting Diodes* (Chiba, Japan, 1998), p. 395.

[7] A. Waag, Th. Litz, F. Fischer, H.-J. Lugauer, T. Baron, K. Schüll, U. Zehnder, T. Gerhart, U. Lunz, and M. Keim, G. Reuscher, and G. Landwehr, J. Cryst. Growth **184/185**, 1 (1998).

[8] S. H. Xin, P. D. Wang, A. Yin, C. Kim, M. Dobrovollska, J. L. Merz, and J. K. Furdyna, Appl. Phys. Lett. **69**, 3884 (1996).

[9] H.-C. Ko, D.-C. Park, Y. Kawakami, and S. Fujita, Appl. Phys. Lett. **70**, 3278 (1997).

[10] K. Leonardi, H. Heinke, K. Ohkawa, D. Hommel, H. Selke, F. Gindele, and U. Woggon, Appl. Phys. Lett. **71**, 1510 (1997).

[11] S. V. Ivanov, A. A. Toropov, T. V. Shubina, S. V. Sorokin, A. V. Lebedev, P. S. Kop'ev, G. R. Pozina, J. P. Bergman, and B. Monemar, J. Appl. Phys. **83**, 3168 (1998).

[12] Z. Zhu, H. Yoshihara, K. Takebayashi, and T. Yao, Appl. Phys. Lett. **63**, 1678 (1993).

[13] A. A. Toropov, S. V. Ivanov, T. V. Shubina, A. V. Lebedev, S. V. Sorookin, P. S. Kop'ev, G. R. Pozina, J. P. Bergman, and B. Monemar, J. Cryst. Growth **184/185**, 293 (1998).

[14] F. Gindele, C. Varkle, U. Woggon, W. Langbein, J. M. Hvam, K. Leonarke, K. Ohkawa, and D. Hommel, J. Cryst. Growth **184/185**, 306 (1998).

[15] M. Strassburg, V. Kutzer, U. W. Pohl, A. Hoffmann, I. Broser, N. N. Ledentsov, D. Bimberg, A. Rosenauer, U. Fischer, D. Gerthsen, I. L. Krestnikov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, Appl. Phys. Lett. **72**, 942 (1998).

[16] M. A. Haase, P. F. Baude, M. S. Hagedorn, J. Qiu, J. M. DePuydt,

H. Cheng, S. Guba, G. E. Höfler, and B. J. Wu, Appl. Phys. Lett. **63**, 2315 (1993).

[17] K. P. O'Donnell, D. M. Bagnall, O. J. Wright, and B. Cockayne, Phys. Status Solidi B **187**, 451 (1995).

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

[19] S. V. Ivanov, S. V. Sorokin, P. S. Kop'ev, J. R. Kim, N. D. Jung, and H. S. Park, J. Cryst. Growth **159**, 16 (1996).

[20] A. Waag, F. Fischer, K. Schüll, T. Baron, H.-J. Lugauer, Th. Litz, U. Zehnder, W. Ossau, T. Gerhardt, M. Keim, G. Reuscher, and G. Landwehr, Appl. Phys. Lett. **70**, 280 (1997).

[21] A. Toropov, S. Ivanov, T. Shubina, S. Sorokin, A. Lebedev, A. Sitnikova, P. Kop'ev, M. Willander, G. Pozina, J. Bergman, and B. Monemar, Jpn. J. Appl. Phys. **38**, Pt 1, No. 113 (1999) (in print).

[22] A. Lebedev, S. Sorokin, A. Toropov, T. Shubina, N. Il'inskaya, O. Nekrutkina, S. Ivanov, G. Pozina, P. Bergman, and B. Monemar, Acta Phys. Pol. A **94**, 421 (1998).

Translated by M. E. Alferieff

# Optical spectra of wide band gap $Be_xZn_{1-x}Se$ alloys

A. M. Mintairov,* S. Raymond, and J. L. Merz

*Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA*

F. C. Peiris, S. Lee, U. Bindley, and J. K. Furdyna

*Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA*

A. M. Mintairov, V. G. Melehin, and K. Sadchikov

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

We report optical measurements (photoluminescence, Raman scattering, and infrared reflectance) of direct band gap and of optical phonon energies of $Be_xZn_{1-x}Se$ alloys grown by MBE on (001) GaAs substrates for a wide range of Be concentrations. The high band gap of BeSe (5.15 eV) suggests the possibility of using isoternary alloys for ultraviolet optoelectronic applications. $Be_xZn_{1-x}Se$ has the unique advantage that it can be lattice matched to Si at about $x=0.5$. We observed a strong linear shift of the $Be_xZn_{1-x}Se$ direct band gap to higher energies with increasing Be content (to 3.63 eV for $x=0.34$). Furthermore, optical phonon parameters for the entire range of BeSe content have been obtained. Finally, polarized infrared and Raman spectra revealed local atomic ordering (anti-clustering) effects in the group-II sublattice. © *1999 American Institute of Physics.* [S1063-7826(99)02509-0]

## INTRODUCTION

Be-chalcogenide semiconductor alloys have been recently proposed for improving the performance of ZnSe-based blue lasers.[1] It is expected that the incorporation of beryllium will lead to bond strengthening within the II–VI lattice, and also will increase the energy of stacking-fault formation, thus reducing defect propagation that has been seriously limiting the lifetimes of ZnSe-based devices. There are also other important opportunities for Be-chalcogenides. The high band gap of BeSe (5.15 eV) suggests the possibility of using these materials for ultraviolet (UV) optoelectronic applications. In addition, $Be_xZn_{1-x}Se$ has the unique advantage that it can be lattice matched to Si at about $x=0.5$. In the present paper we report optical measurements [photoluminescence, Raman scattering, and infrared (IR) reflectance] of the direct band gap and of the optical photon energies of $Be_xZn_{1-x}Se$ alloys for a wide range of Be concentrations.

## 1. EXPERIMENT

The $Be_xZn_{1-x}Se$ ($0<x\leqslant1$) samples used in this study were grown by molecular-beam epitaxy (MBE) on (001) GaAs substrates at 310 °C. The thicknesses of the samples were between 0.3–1 $\mu$m. During the growth process the surface reconstruction was monitored by reflection high-energy electron diffraction (RHEED), which indicated a ''streaky'' pattern for all Be concentrations. The Be content of the alloys was determined from lattice parameter measurements by x-ray diffraction.

Photoluminescence (PL) spectra were excited by the 325-nm line (10-mW power) of a HeCd laser, and measured

using 270-mm focal length monochromator equipped with a CCD detector.

IR reflectance spectra were taken at oblique incidence (10 to 50°) with a Bruker IFS-66V spectrometer. The photon
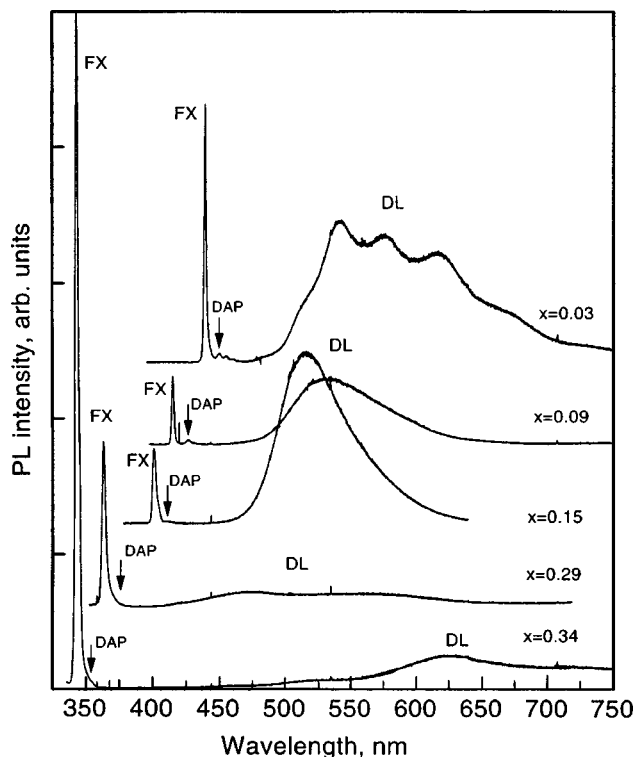


FIG. 1. 10-K PL spectra of $Be_xZn_{1-x}Se$ ($x=0.03$, 0.09, 0.15, 0.29, and 0.34) grown on GaAs.
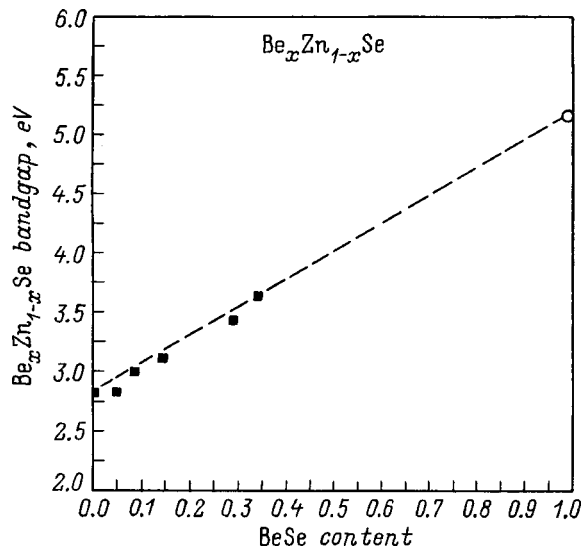
FIG. 2. $Be_xZn_{1-x}Se$ band gap: solid squares—10-K PL spectra; open circle—300 K reflectivity data from Ref. 13; dotted line—linear extrapolation from ZnSe to BeSe.

frequencies and damping were extracted using a standard modeling procedure of the reflectivity spectra via multimode representation of the alloy dielectric function.

The Raman spectra were obtained with the argon ion laser line excitation using 514.5-nm line, and measured at 300 K using a DFS-52 double monochromator. The *xy* and *xx* components of the Raman spectra (where the $x\|[100]$ and $y\|[010]$) excited from the (001) growth plane were analyzed.
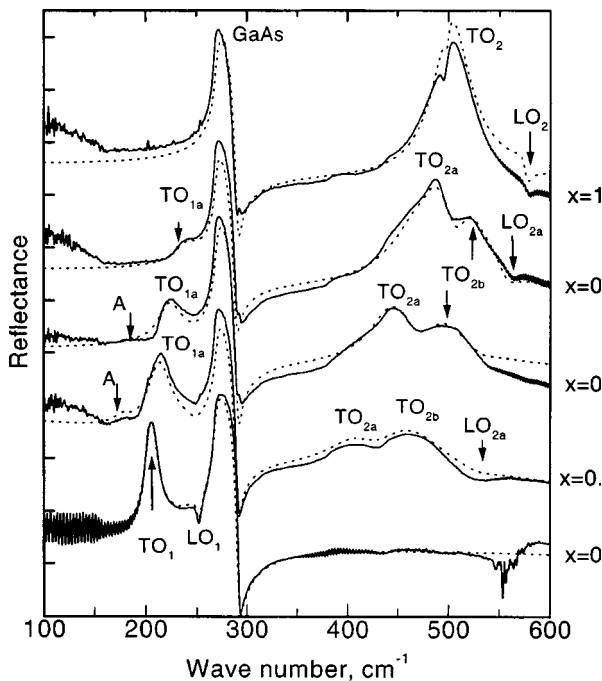


FIG. 4. Raman spectra of $Be_{0.34}Zn_{0.66}Se$ measured in *xy* (solid curve) and *xx* (dotted curve) configurations.

## 2. PL SPECTRA

PL spectra were used to measure the direct band gap energy of $Be_xZn_{1-x}Se$ from $x=0$ to 0.34. As in the case of pure ZnSe epitaxial layers,[2] the low-temperature PL spectra of $Be_xZn_{1-x}Se$ were dominated by a free exciton (*FX*) line and a broad deep level (*DL*) band (see Fig. 1). For low *x* (0.03, 0.09, and 0.15), donor-acceptor pair (*DAP*) transition bands having a *LO* phonon replica (for $x=0.03$) are also clearly resolved. Because of the increase in the half-width of the *FX* line with increasing BeSe content (from 12 meV for $x=0.03$ to 40 meV for $x=0.34$), the *DAP* transition for higher *x* (0.29 and 0.34) appears as a weak shoulder at the low-energy side of the *FX* line. We observed a strong suppression of the *DL* emission for $x>0.3$, indicating improvement of the quality of the material.

Increasing the Be content results in a strong shift of the *FX* line toward higher energies. As an example, for $x=0.34$ the *FX* line was observed at an energy as high as 3.63 eV. Figure 2 shows the *FX* line position versus BeSe content, including BeSe direct band energy measured in Ref. 3 by reflectance spectroscopy. Note that all experimental



FIG. 3. IR spectra of $Be_xZn_{1-x}Se$ ($x=0$, 0.23, 0.57, 0.81, and 1) measured at 40° angle of incidence: solid curve—experiment; dotted curve—modeling.



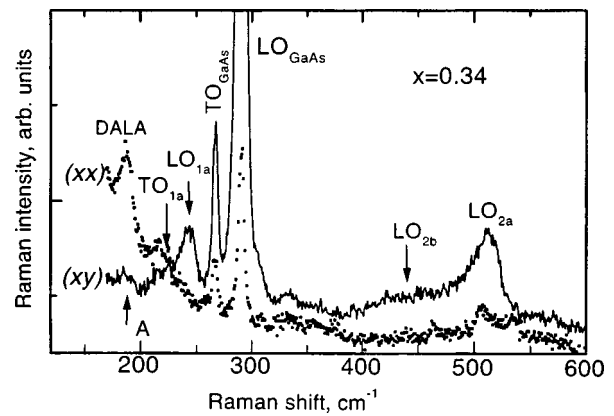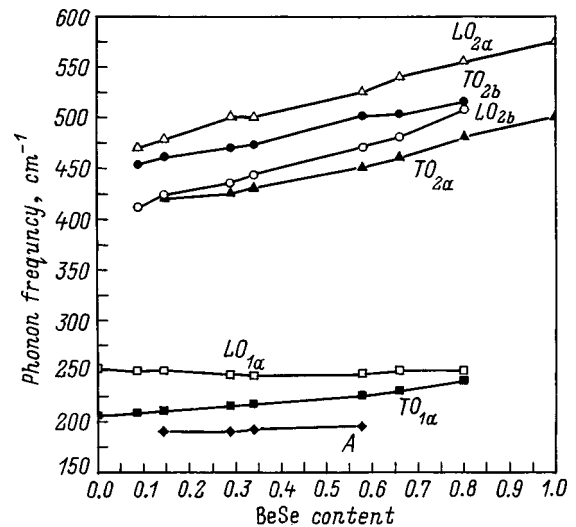FIG. 5. Dependence of the polar optic phonon frequencies on Be content in $Be_xZn_{1-x}Se$.

points are close to the straight line that connects ZnSe and BeSe band gaps, indicating negligible bowing.

## 3. PHONON MODES OF Be$_x$Zn$_{1-x}$Se

IR (Fig. 3) and Raman (Fig. 4) spectra reveal three polar optic phonon modes of the Be$_x$Zn$_{1-x}$Se lattice for $0.01 < x < 0.08$ (Fig. 5). They are denoted in Figs. 3, 4, and 5 as $TO_{1a}$, $TO_{2a}$, and $TO_{2b}$ (transverse components) and $LO_{1a}$, $LO_{2a}$, and $LO_{2b}$ (longitudinal components). We identify two of them (subscripts $2a$ and $2d$) that appear in the frequency range $400-570$ cm$^{-1}$ as BeSe-type modes and one (subscript $1a$) that appears at $200-250$ cm$^{-1}$ as a ZnSe-type mode. For $x < 0.6$ we also identify a weak polar mode $A$ at $\sim 190$ cm$^{-1}$, which can be attributed to ZnSe-type vibration. The ZnSe-type $TO_{1a}$ mode shows an unusual behavior: an increase of frequency with decreasing density of the ZnSe content (see Fig. 5). We suggest that this follows from the positive $TO$ phonon dispersion characteristic of bulk binary ZnSe.[4] As can be seen from Fig. 4, all observed modes obey zinc-blende Raman selection rules, which imply a dominance of the $LO$ components in $xy$ configuration.

Our observation of the two types of BeSe modes indicates that there are two different local arrangements (ordering) of the BeSe bonds in the alloy lattice. Our observation also gives evidence of anticlustering effects in the group-II sublattice.[5] Such effects may be driven by the very large difference in the ZnSe and BeSe bond lengths. We also note that the absence of zinc-blende forbidden components for $LO_{2a}$ and $LO_{2b}$ bands in the Raman spectra[1) ] is indicative of tetragonal symmetry of the ordered atomic arrangement.[6]

---

*E-mail: Alexander.Mintairov.l@nd.edu
1) The weak structure at $510$ cm$^{-1}$ in $xx$ configuration is the second order spectra of GaAs substrate.

---

[1] A. Waag, F. Fischer, H. J. Lugauer, Th. Litz, J. Laubender, U. Lunz, U. Zehneder, W. Ossau, T. Ferhardt, M. Moller, and G. Landwehr, Appl. Phys. Lett. **70**, 280 (1997).

[2] J. Gutowski, N. Presser, and G. Kudlek, Phys. Status Solidi A **120**, 11 (1990).

[3] J. P. Faurie, V. Bousquet, P. Brunet, and E. Tournie, J. Cryst. Growth **184/185**, 11 (1998).

[4] B. Hennion, F. Moussa, G. Pepyand, and K. Kunc, Phys. Lett. A **36**, 377 (1971).

[5] A. M. Mintairov and H. Temkin, Phys. Rev. B **55**, 5177 (1977).

[6] A. M. Mintairov, P. A. Blagnov, V. G. Melehin, N. N. Faleev, J. L. Merz, Y. Qiu, S. A. Nikishin, and H. Temkin, Phys. Rev. B **56**, 15836 (1997).

# Excitonic molecules trapped by quantum dots and isoelectronic impurities in many-valley semiconductors

A. A. Rogachev*)

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1123–1125 (September 1999)

Excitonic molecules with two to six excitons trapped by quantum dots or isoelectronic impurity centers are studied theoretically and experimentally for the first time. Such excitonic molecules can exist only in many-valley semiconductors. A model of an excitonic molecule, valid for a large number of valleys in the conduction and valence bands, is examined. The formation kinetics of excitonic molecules is discussed. A material feature of excitonic molecules bound on quantum dots is the existence of tunneling transitions of excitons between quantum dots. © *1999 American Institute of Physics.* [S1063-7826(99)02609-5]

The properties of excitonic molecules and excitonic liquid in single-valley semiconductors differ markedly from those in many-valley semiconductors. For single-valley semiconductors the terms excitonic molecule and biexciton are synonyms.[1,2] An excitonic liquid in single-valley semiconductors is an insulator.[3,4] Strongly bound molecules with two, three, and up to six excitons are possible in many-valley semiconductors.[5,6] Correspondingly, an excitonic liquid in many-valley semiconductors exhibits metallic properties in all known cases. The main reason for the large difference in the basic physical properties of excitonic molecules is that many-valley semiconductors possess a degeneracy $g$-factor greater than 2 in the conduction band and in the valence band. The terms excitonic molecules, excitonic complexes, and polyexcitons, often used in the literature, should be treated as synonyms. Isoelectronic impurities and artificially produced quantum dots in many-valley semiconductors include not only electrons, holes, and excitons but also excitonic molecules consisting of a large number of excitons.

Excitonic molecules are smaller in size than the radius of a free exciton.[5,7] The luminescence spectra of bound excitonic molecules have been found to be similar to the spectra of free excitonic molecules.[8] Each line in the spectrum corresponds to a decrease in the number of excitons in the molecule by one, i.e., $E_z \rightarrow E_{z-1}$. The luminescence line corresponding to the $z$ excitons lies in the redder region of the spectrum than the line corresponding to the $z-1$ exciton. The distance between the spectral lines corresponding to the luminescence of excitonic molecules bound on quantum dots is generally larger for free molecules. It is of special interest to investigate the luminescence of bismuth-doped GaP crystals (GaP:Bi), since the isovalent bismuth impurity in these crystals can capture one electron and one hole with binding energy less than 0.05 eV (Ref. 8). In this paper, we treat this impurity as a quantum dot.

In the well-known semiconductors, such as Si, Ge, and GaP, the binding energy of free electrons and excitonic molecules is less than the spin-orbit splitting of the valence band. For this reason, the degeneracy factor of the hole ground state is 4. Since the energy required to trap an excitonic molecule by a quantum dot (per exciton) is higher than the spin-orbit interaction energy in the valence band (0.05 eV in GaP), we will disregard this interaction. Quantum dots can then capture six excitons, because the valence band is sixfold degenerate, and the conduction band possesses a camel-hump structure with an approximately 0.05-eV hump. From analysis of the luminescence spectra we obtain an energy $E_z$, which is related to the ground-state energy $F_z$ in the molecule by the relation

$$F_z = \frac{1}{z} \sum_{z'=1}^{z'=z} E_{z'},\tag{1}$$

where $z$ is the number of excitons in the molecule ($z \leq 6$). Correspondingly, the position $E_z$ in the spectrum is related to the ground-state energy in the following way:

$$E_z = zF_z - (z-1)F_{z-1}.\tag{2}$$

Let us consider a simple model of an excitonic molecule, where the single-electron wave function is identical for holes and electrons and is spherically symmetric.[5,7] We assume that $m_e = m_h = m_{ex}$. For simplicity, we use a hydrogen-like trial wave function

$$\Psi(R) = \sqrt{\frac{1}{\pi r^3}} e^{-\frac{R}{r}},$$

where $r$ is the effective radius of the excitonic molecule. We can then write the kinetic energy of the electron-hole pair as follows:

$$E_k = \frac{\hbar^2}{2m_h r^2} + \frac{\hbar^2}{2m_e r^2} = \frac{\hbar^2}{m_{ex} r^2}.$$

Since the electron and hole wave functions are identical, the charge density in the Hartree approximation is zero [$\rho(R) = 0$]. If the number of excitons in an excitonic molecule is large, then the exchange energy can be ignored, since the electrons and holes which encounter one another have, as a
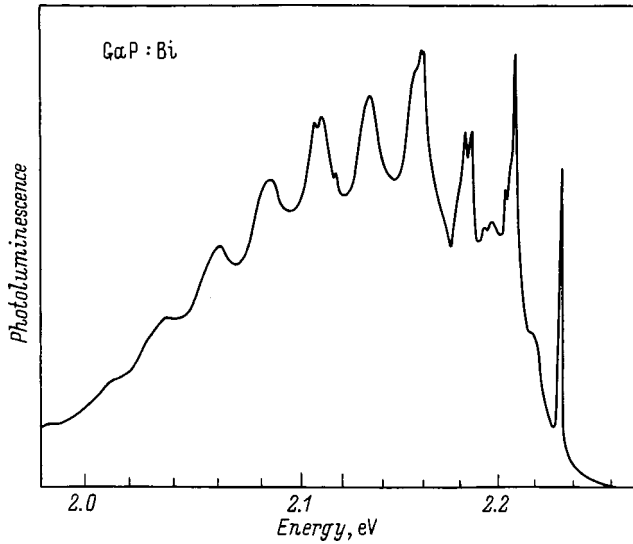
FIG. 1. GaP:Bi luminescence spectra from Ref. 8. The Bi density is approximately $10^{17}$ cm$^{-3}$, $T = 2$ K. The peaks in the luminescence spectrum at energies 2.06, 2.09, 2.12, and 2.14 eV correspond to different bound excitonic molecules.
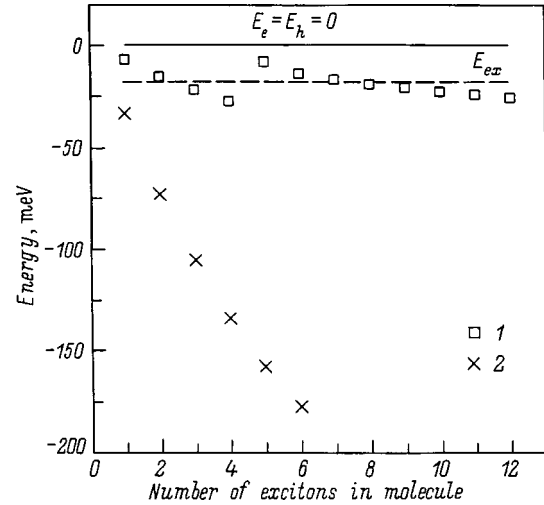


FIG. 2. Calculation of the energy of excitonic molecules. The positions of the lines in the luminescence spectra are measured from the energy of free electrons and holes ($E_z$). The solid line represents the energy of an electron–hole pair with $E_k = 0$, the dashed line corresponds to the binding energy of an exciton. $1$ — Energy in the spectrum of free excitonic molecules; $2$ — energy of excitonic molecules with a different number of electron–hole pairs (or excitons) bound on Bi atoms.

rule, different wave numbers (spins or valley numbers). The correlation potential energy per electron–hole pair ($e - h$) is approximately $e^2/\varepsilon L$, where

$$L = \sqrt{\frac{\varepsilon E_k}{6\pi e^2 n_{\text{ex}}}}.$$

The effective $e - h$ pair density in the molecule is

$$n_{\text{ex}} = \frac{2\pi}{9} z r^{-3}.$$

The energy per $e - h$ pair in an excitonic molecule can be written as

$$F_z = -\frac{e^2}{\varepsilon L} - A \int_0^l 4\pi R^2 |\Psi(R)|^2 dR + E_k$$

$$= -\frac{e^2}{\varepsilon L} - A \left[ 1 - \left( \frac{2l^2}{r^2} + \frac{2l}{r} + 1 \right) \exp\left( -\frac{2l}{r} \right) \right] + E_k, \quad (3)$$

where $l$ is the radius of the quantum well. The radius $r$ of the excitonic molecule is determined from the condition

$$\frac{dF_z}{dr} = 0.$$

The coefficient $A$ is chosen from the condition that if the Coulomb energy is removed in Eq. (3) for $F_z$ ($e = 0$), then $F_z$ should be very small in absolute magnitude but still negative (weak coupling). The width of the quantum well is chosen in such a way that $F_z$ obtained from Eq. (3) would be identical to the experimental value. For the experimental (Fig. 1) and theoretical (Fig. 2) results to agree, $l$ must be chosen to be 6 Å, i.e., on the order of 0.1 of the exciton radius.

The density of Bi atoms in the GaP crystal was approximately $10^{17}$ cm$^{-3}$, which makes tunneling the main method of diffusion of excitons bound at an impurity between bis-

muth atoms. The formation of excitons from electrons and holes produced by illuminating the sample is the fastest process, and its time is close to the cooling time of the light-generated $e - h$ pairs ($10^{-12}$ s). Capture of excitons on Bi impurity atoms requires a time on the order of $10^{-11}$ s, which is less than the lifetime of bound excitons. Therefore it can be assumed that illumination of the sample with a powerful light source directly produces excitons trapped at impurity atoms. These bound excitons move along the crystal by tunneling. If a bound exciton is found to be on an impurity atom where a trapped exciton is already present, then a trapped biexciton is formed. In principle, bound biexcitons can also diffuse by tunneling, but in this case the barrier is higher and the diffusing particle (biexciton) is heavier. Similar considerations also apply to the formation of other bound excitonic molecules. At low temperature tunneling can only increase the number of excitons in a bound molecule, because upon formation of a large molecule the energy is lost and the reverse process is impossible. The main recombination channel for bound excitonic molecules is Auger recombination.

Auger recombination in trapped excitonic molecules with the largest number of $e - h$ pairs and, correspondingly, the highest $e - h$ pair density, becomes the most intense recombination. Thus, in a simplified model of recombination only Auger recombination in hexaexcitons need be considered. Under these assumptions the system of kinetic equations can be written in the form

$$\frac{dn_{\text{ex}}}{dt} = I + 5\frac{n_6}{\tau} - 2\alpha_2 n_{\text{ex}} n_{\text{ex}} - \alpha_3 n_{\text{ex}} n_b - \alpha_4 n_{\text{ex}} n_3$$

$$- \alpha_5 n_{\text{ex}} n_4 - \alpha_6 n_{\text{ex}} n_5,$$

$$\frac{dn_b}{dt} = \alpha_2 n_{\text{ex}} n_{\text{ex}} - \alpha_3 n_{\text{ex}} n_b, \quad \frac{dn_3}{dt} = \alpha_3 n_{\text{ex}} n_b - \alpha_4 n_{\text{ex}} n_3,$$

$$\frac{dn_4}{dt} = \alpha_4 n_{ex} n_3 - \alpha_5 n_{ex} n_4, \quad \frac{dn_5}{dt} = \alpha_5 n_{ex} n_4 - \alpha_6 n_{ex} n_5,$$

$$\frac{dn_6}{dt} = \alpha_6 n_{ex} n_5 - \frac{n_6}{\tau}, \tag{4}$$

where $I$ is the intensity of the light exciting the luminescence, $\alpha_n$ is the trapping coefficient, $\tau$ is the lifetime of excitons in hexamolecules, and $n_b$, $n_3$, $n_4$, $n_5$, and $n_6$ are, respectively, the densities of biexcitons, triexcitons, tetraexcitons, pentaexcitons, and hexaexcitons that are trapped on quantum dots. The steady-state solution is

$$n_{ex} = \sqrt{\frac{I}{\alpha_2}}, \quad n_b = \frac{I}{\alpha_3 n_{ex}}, \quad n_3 = \frac{I}{\alpha_4 n_{ex}},$$

$$n_4 = \frac{I}{\alpha_5 n_{ex}}, \quad n_5 = \frac{I}{\alpha_6 n_{ex}}, \quad n_6 = I\tau. \tag{5}$$

A characteristic feature of this solution is that the ratios $n_b/n_5$, $n_4/n_5$, $n_3/n_4$, and $n_3/n_b$ do not depend on the intensity $I$ of the luminescence pump. The fact that the ratio of the intensities of the luminescence lines remains constant in a wide range of pump intensities has been established experimentally.

In summary, it can be regarded as proved theoretically and experimentally that in many-valley semiconductors up to six excitons are trapped in the ground state of a quantum dot and that the binding energy of the excitons increases with decreasing diameter of the quantum dot.

*)E-Mail: arog@les.ioffe.rssi.ru

[1] S. A. Moskalenko, Opt. Spektrosk. **5**, 147 (1958).
[2] M. A. Lampert, Phys. Rev. Lett. **1**, 450 (1958).
[3] L. V. Keldysh and Yu. V. Kopaev, Fiz. Tverd. Tela (Leningrad) **6**, 2791 (1964) [Sov. Phys. Solid State **6**, 2219 (1964)].
[4] L. V. Keldysh and A. N. Kozlov, Zh. Éksp. Teor. Fiz. **54**, 978 (1968) [Sov. Phys. JETP **27**, 521 (1968)].
[5] A. A. Rogachev, in *Handbook on Semiconductors*, edited by P. T. Landsberg (Elsevier/North-Holland Publ. Co., Amsterdam, 1992), Vol. 1, Chap. 9, p. 449.
[6] M. L. W. Thewalt, V. A. Karasyuk, and D. A. Muler, in *Proceedings of the 23rd Conference on the Physics of Semiconductors* (World Scientific, 1996), Part I, p. 341.
[7] A. A. Rogachev, Fiz. Tverd. Tela (St. Petersburg) **40**(5), 929 (1998) [Phys. Solid State **40**, 855 (1998)].
[8] P. J. Dean, in *Progress in Solid State Chemistry* (Pergamon Press, New York, 1973), Vol. 8, p. 1.

Translated by M. E. Alferieff

# High-frequency current instabilities in a silicon Auger transistor

E. V. Ostroumova and A. A. Rogachev

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

In an Auger transistor formed from an Al–$SiO_2$–$n$-Si heterojunction with a tunneling-thin oxide layer we have investigated high-frequency instabilities of $S$- and $N$-type in the collector current which arise during tunneling injection of hot electrons from the metal into the semiconductor. An Auger transistor is a new type of device in which a metal–insulator heterojunction is used as the wideband semiconductor emitter and the base of the transistor is induced by an electric field in the form of a self-consistent quantum well for holes on the silicon surface. The electrons injected from the metal into the semiconductor with a high kinetic energy (greater than 1 eV) during impact ionization generate electron–hole pairs in the region of the base–collector junction. This disrupts the current balance of the transistor and leads to the appearance of an unstable current of $S$- or $N$-type in the collector characteristics (in a circuit with a common emitter). The nature of the instability is connected with the large current gain in an Auger transistor ($\alpha > 1$). © *1999 American Institute of Physics.* [S1063-7826(99)02709-X]

The use of the heterostructure Al–$SiO_2$–$n$-Si with a tunneling-thin oxide layer[1,2] to create an Auger transistor is so far a unique realization of the idea of Krömer to create a transistor with a wideband hetero-emitter,[3] despite the fact that numerous theoretical efforts have been made to identify the right heteropairs. In an Auger transistor, current gain is achieved as a result of impact ionization, which is produced by hot electrons injected from a wideband emitter into the narrowband base.[3] The transistor structure Al–$SiO_2$–$n$-Si uses a metal–silicon oxide heterojunction instead of a semiconductor wideband emitter (Fig. 1), where the oxide should be transparent to electron tunneling (a few nanometers thick).[1,2] In this case, the kinetic energy of the electrons tunneling from the metal into the semiconductor is so large that generation of electron–hole pairs in the silicon becomes possible during impact ionization (Auger generation). Another device utilizing Auger generation in semiconductors is a Read diode.[4] This devices operates in the frequency range $10^{11} - 10^{12}$ Hz, but this is a two-contact device, which creates great difficulties in its application. In bipolar transistors, current gain is connected with diffusion or drift of carriers through the base of the transistor. In an Auger transistor based on the structure Al–$SiO_2$–$n$-Si, in contrast to a Read diode, the regions of generation and carrier drift are almost completely spatially separated. Impact ionization occurs near the base of the transistor, and drift, in the region of the base–collector junction. The rate of generation of electron–hole pairs is determined by the voltage supplied to the base.[6,7] The base of the transistor is formed under the action of the electric field of the oxide in the form of a self-consistent quantum well for holes on the surface in the $n$-type silicon layer and has a thickness about 10 Å for well depths up to 0.7 eV. Electrons tunnel-injected from the metal and therefore possessing a large kinetic energy substantially increase their kinetic energy in the field of the base of the transistor,

passing above the self-consistent surface quantum well for holes, and then lose their kinetic energy in the process of generation of electron–hole pairs and phonons upon impact ionization.[2] Additional heating and cooling of electrons takes place over its mean free path of $10^{-6}$ cm. At the same time, the drift region, i.e., the space charge region of the base–collector junction, is usually not less than $10^{-5}$ cm wide.

The superthin base of an Auger transistor allows the electrons to pass straight through the base ballistically (Fig. 1). Therefore, the time of flight of the electrons through the base cannot limit the working frequency of the Auger transistor or the impact ionization time (on the order of $10^{-13}$ s). The response time of an Auger transistor is determined virtually only by the charging time of the emitter–base capacitance. For a current density of $10^4$ A/cm$^2$ and emitter area of $0.2\,\mu$m$^2$ the limiting frequency can be increased to $10^{11} - 10^{12}$ Hz (Ref. 7). This frequency approaches the limiting working frequency of a Read diode, which at present possesses the shortest response time among semiconductor devices.

The kinetic energy received by the electron during ballistic transit of the base is equal to the depth of the well (0.7 eV) and is equal to roughly half of the threshold energy (1.5 eV) of impact ionization in silicon. The existence of deep self-consistent quantum wells on the surface of the silicon layer makes it possible to control the energy of the fast electrons taking part in the impact ionization and, in essence, to control the appearance of $S$- and $N$-type instabilities in the collector current of the Auger transistor (in a circuit with a common emitter).[5–7] For small currents in the base circuit the collector characteristics of the Auger transistor are similar to the characteristics of bipolar transistors. But with growth of the base current the collector characteristics exhibit a growth of the current gain $\beta$ that is characteristic of an Auger transistor; when $\beta$ becomes negative, instabilities
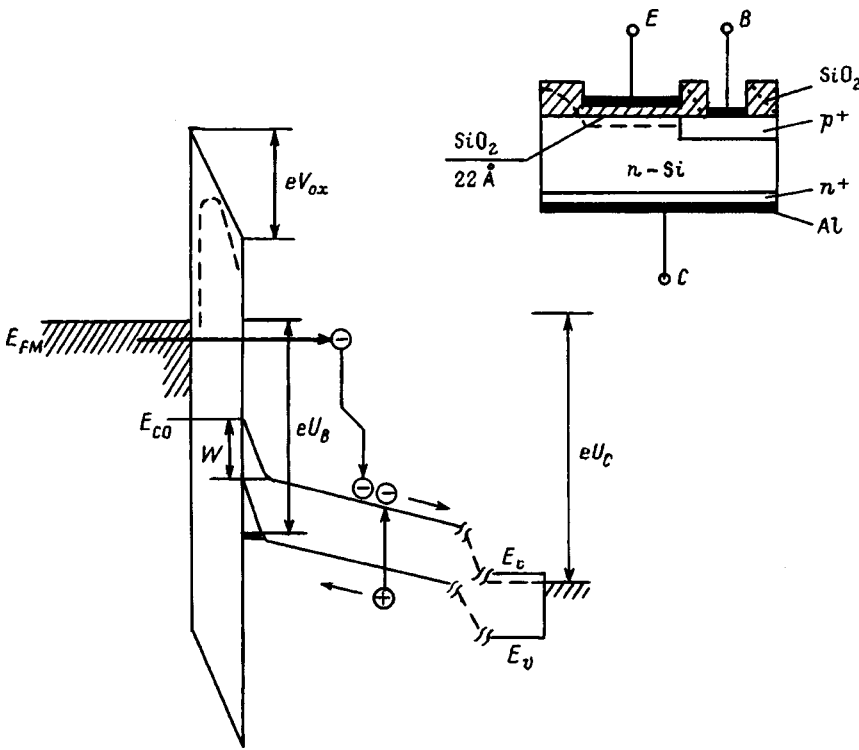
FIG. 1. Energy diagram of an Auger transistor based on an Al–SiO$_2$–$n$-Si heterojunction with a tunneling-thin oxide layer. The metal is negative-biased relative to the semiconductor. Inset shows a diagram of the transistor.

arise. The appearance of *S*-like instabilities of the collector current (Fig. 2) is typical already at collector current densities about $10^3$ A/cm$^2$. At higher current densities (greater than $5 \times 10^3$ A/cm$^2$) *N*-type instabilities are observed in the Auger transistor in the region of low voltages on the collector (Fig. 3). In some cases a transition from an *S*-type instability to an *N*-type instability is observed (Fig. 4).[6,7]

To describe tunneling currents of electrons and holes through a potential barrier, the semiclassical approximation is often used. Employing this approximation, the tunneling current can be written in the form

$$I = \frac{eN}{\tau} \exp\left[-(2/\hbar) \int p \, dx\right], \qquad (1)$$

where the integral $(2/\hbar)\int p \, dx$, which defines the tunneling probability, is taken over the classically inaccessible region; $N$ is the two-dimensional density of the electrons that can participate in tunneling, and $\tau$ is the characteristic collision time of the electrons located in the potential well. We can represent all tunneling currents flowing in an Auger transis-
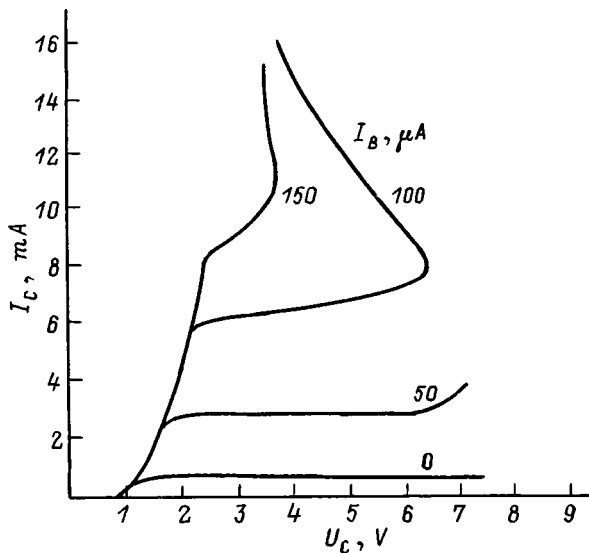


FIG. 2. Experimental *S*-shaped collector characteristics of an Auger transistor in a circuit with a common emitter; $T = 300$ K. The numbers alongside the curves give the values of the base current $I_B$.
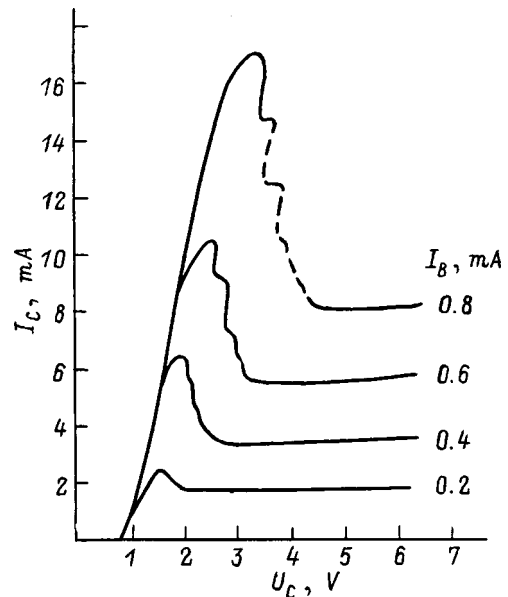


FIG. 3. Experimental *N*-shaped collector characteristics of an Auger transistor in a circuit with a common emitter; $T = 300$ K.
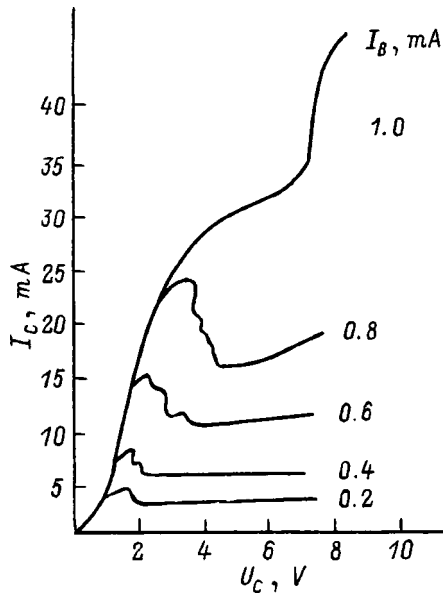
FIG. 4. Experimental collector characteristics of an Auger transistor in a circuit with a common emitter. For small voltages on the collector $S$-type instabilities go over to $N$-type instabilities; $T = 300$ K.

tor in a form analogous, for example, to the current for electrons tunneling from the metal to the semiconductor through the oxide layer:[7]

$$I_{mc} = \frac{eV}{16\pi^2\hbar d^2}\exp\left\{-\frac{4\sqrt{2m^*d}}{3\hbar V_{ox}}[V^{3/2}-(V-V_{ox})^{3/2}]\right\}.$$

(2)

Here $eV = \varphi_m$ is the work function of the electrons from the metal to the conduction band of the silicon-oxide layer; $m^* = 0.29m_0$ is the effective mass of the tunneling electrons in the silicon-oxide layer; $d$ is the thickness of the oxide layer, and $V_{ox} = (E_{FM} - E_{CO})/e$ is the voltage drop across the oxide layer. The calculations are described in greater detail in Refs. 6 and 7.

Generation of electron–hole pairs during impact ionization leads to amplification of the current since instead of one injected electron, three current carriers now arise simultaneously. Additional currents appear—the Auger electron and hole currents. While the Auger electron current flows into the collector, increasing the magnitude of the collector current, the Auger hole current flows from the generation region (base–collector junction) into the transistor base, and the current in the base circuit (hole current) flows from the base into the emitter. The base current $I_B = I_{ms} + I_{mv}$ and the Auger current ($I_{Auger}$) flow in the direction opposite the leakage currents $I_{ms}$ and $I_{mv}$:

$$I_B + I_{Auger} - I_{ms} - I_{mv} = 0.$$

(3)

This equation defines the current balance in an Auger transistor. The Auger current is not part of the emitter current $I_E = I_C + I_{ms} + I_{mv}$ but is proportional to the electron tunnel-

ing current from the metal into the conduction band of the semiconductor, $I_{mc}$, and depends on the impact ionization $A$: $I_{Auger} = AI_{mc}$ (Ref. 7). The ionization $A$ reflects the dependence of the Auger generation current on the energy of the injection electrons (see Fig. 1). Equation (3) has two solutions since the Auger hole current increases with increasing impact ionization, i.e., any value of the base current corresponds to two values of the base voltage and correspondingly two values of the collector current. Each value of the base current can be realized in two different cases: in the first, at a small Auger current and small base voltage ($I_B = I_{mc} + I_{mv}$), and in the second, at a large Auger current and a large base voltage ($I_B = I_{mc} + I_{mv} - I_{Auger}$). In the first case the collector current $I_C$ is small, and in the second, the collector current can be several orders of magnitude larger.

$N$-type instabilities (Fig. 3) appear when the collector voltage is small, and depend not only on the base and collector voltages, but also on the concentration near the silicon surface of electron–hole pairs formed earlier during impact ionization in a time less than the resorption time. $N$-type instabilities are similar to the instabilities observed in a Read diode. This is the highest-frequency current instability in an Auger transistor. There are two paths of resorption of the electron–hole cloud. The first is in the electric field of the collector, and the second is in the form of the leakage current, i.e., the hole tunneling current from the semiconductor into the metal. The first process leads to $N$-type instabilities in the collector current, and the second gives rise to $S$-type instabilities. Both types of instabilities are high-frequency instabilities. The nature of this phenomenon is purely quantum mechanical and is associated with the strong electric fields at the surface of the silicon layer.

[1] I. V. Grekhov, E. V. Ostroumova, A. A. Rogachev, and A. F. Shulekin, Pis'ma Zh. Tekh. Fiz. **17**(7), 44 (1991) [Sov. Tech. Phys. Lett. **17**(7), 476 (1991)].

[2] E. V. Ostroumova and A. A. Rogachev Fiz. Tekh. Poluprovodn. **28**(2), 1411 (1994) [Semiconductors **28**, 793 (1994)].

[3] H. Kroemer, in *Heterojunctions and Metal–Semiconductor Junctions*, edited by A. G. Milns, J. J. Feucht (Academic Press, New York, 1972).

[4] W. T. Read, Bell Syst. Tech. J. **37**, 401 (1958).

[5] E. V. Ostroumova and A. A. Rogachev, in *Proceedings of the 21st International Conference on Microelectronics (MIEL'97)*, Sept. 14–17, 1997 (Nish, Yugoslavia), Vol. 1, p. 227. Published by the Electron Devices Society IEEE, IEEE Catalog No. 97TH8232 (1997).

[6] E. V. Ostoumova and A. A. Rogachev, in *Fundamental Aspects of Ultrathin Dielectrics on* Si*-Based Devices: Towards an Atomic-Scale Understanding.* NATO Science Series (Kluwer Academic Publishers, Dordrecht, 1998), p. 383.

[7] E. V. Ostroumova and A. A. Rogachev, Microelectron. J. **29**, 701 (1998).

Translated by Paul F. Schippnick

# Application of superfast ($10^2 - 10^3$ °C/s) cooling of a solution–melt in the liquid-phase epitaxy of semiconductors

A. V. Abramov, N. G. Deryagin, and D. N. Tret'yakov

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted March 1, 1999; accepted for publication March 2, 1999
Fiz. Tekh. Poluprovodn. **33**, 1130–1133 (September 1999)

Use of a superhigh ($10^2 - 10^3$ °C/s) rate of cooling of a solution–melt substantially extends the possibilities of the method of liquid-phase epitaxy. In the case of the system Al–Ga–As we demonstrate the influence of superfast cooling on the main parameters of the as-grown layers, such as their thickness, composition, and carrier concentration. Possibilities of the method for obtaining various semiconductor heterostructures with high (up to 12%) lattice mismatch of the contacting materials are considered. © *1999 American Institute of Physics.*
[S1063-7826(99)02809-4]

## 1. INTRODUCTION

In the development of the method of liquid-phase epitaxy (LPE), recent years have witnessed a new trend based on the use of conditions far from equilibrium during crystallization. This is achieved by using superfast cooling of the solution–melt during growth. A specially developed technique which permits rapid cooling of a thin layer of the solution–melt immediately adjacent to the phase boundary makes it possible to control the cooling rate over a wide interval up to values in the range $(10^2 - 10^3)$ °C/s (Ref. 1).

The use of superfast cooling makes it possible to substantially expand the possibilities of the LPE method in comparison with the widely used methods of growing layers from a solution–melt. Thus, for example, cooling at rates in the range (300–600) °C/s has made it possible to obtain contiguous planar layers of GaAs on Si equal or superior in their degree of crystalline perfection to layers grown by molecular-beam or gas-phase epitaxy, and also to obtain solid solutions of the compositions $(Ge_2)_x(GaAs)_{1-x}$ and $(Si_2)_x(GaAs)_{1-x}$ (Ref. 2).

In the present paper we report the results of an experimental study of the influence of highly nonequilibrium conditions during crystallization on the main parameters of as-grown layers of semiconductor materials based on the system Al–Ga–As and demonstrates the potential of using superhigh cooling rates when cooling solution–melts in the fabrication of semiconductor layers.

## 2. EXPERIMENT

In our work we used liquid-phase epitaxy. This method allows one to perform crystallization under conditions of superfast cooling over a wide temperature range. In this case, superfast cooling is achieved by continuous and intense extraction of heat through the solid phase over the entire prescribed interval of epitaxial growth. This is achieved while moving the substrate in contact with the solution–melt along the ''cold'' surface of the graphite housing of the cassette, whose temperature decreases in the direction of motion. It is quite easy to control the rate of cooling here by prescribing the temperature distribution along the graphite housing, and also by varying the speed with which the substrate is moved along the surface. The cooling rate at different temperatures of the onset of crystallization (650−950) °C was measured by introducing a thermocouple into the solution–melt. These values showed that for temperature gradients along the surface of the graphite housing equal to (20−50) °C/cm the cooling rate can reach values in the range (300−600) °C/s. The time and temperature intervals of growth at enhanced cooling rates were (1.5–3.0) s and (300−600) °C, respectively. Further cooling of the system to room temperature took place at average rates of ∼0.5 °C/s. To grow AlGaAs structures, a buffer layer of $Al_xGa_{1-x}As$ with constant composition $x_b$ was first grown on the GaAs substrate with a cooling rate of 0.25 °C/min; then from this same melt crystallization at superhigh cooling rates was performed.

## 3. MEASUREMENT OF THE THICKNESS AND COMPOSITION OF THE AS-GROWN LAYERS

To study the effect of enhanced cooling rates on the thickness of as-grown layers, we examined the binary system Ga–As, where if one knows the parameters of the growth process, it is possible to calculate in a quite simple way the thickness of the deposited layer with high accuracy.

The thickness of the as-grown layers was measured by visual inspection of a cleavage of the structure under an optical microscope with an accuracy of $\pm 0.2$ $\mu$m. Results of measurements of the the thickness of the layers are presented in Table I. The calculated values of the layer thicknesses presented in the table were obtained for the case of equilibrium crystallization from a semi-infinite layer of solution–melt under diffusion-limited conditions imposed on macrotransport in the liquid phase.

In the calculation we took into account the temperature dependence of the diffusion coefficient and the solubility of As in the liquid phase. As can be seen from the table, the experimentally obtained values of the layer thicknesses greatly exceeds their calculated values. Since forced mixing

TABLE I. Thickness of the GaAs layer grown under different conditions.

| $T_0$, °C | $v$, °C / s | $t$, s | $h_{\exp}$, $\mu$m | $h_{\mathrm{cal}}$, $\mu$m |
|---|---|---|---|---|
| 950 | 610 | 1.0 | 15.0 | 7.0 |
| 850 | 460 | 1.24 | 7.0 | 1.46 |
| 800 | 415 | 1.3 | 4.5 | 0.58 |
| 750 | 365 | 1.4 | 2.8 | 0.21 |

*Note*: $T_0$ — initial growth temperature, $v$ — rate of cooling of solution–melt, $t$ — growth time of layer for superfast cooling, $h_{\mathrm{cal}}$ — calculated value of layer thickness, $h_{\exp}$ — measured layer thickness.

of the solution–melt was absent, an increase in the layer thickness is most probably associated with an increase in the rate of mass transport in the solution–melt due to natural convection induced by the steep temperature gradients arising in the liquid phase. The greatest influence here is probably exerted by the steep horizontal gradient along the substrate surface.

The profile across the thickness of the layer of the composition of the epitaxial $Al_xGa_{1-x}As$ solid solutions grown under conditions of superfast cooling was studied by secondary-ions mass spectroscopy (SIMS) on a CAMECA IMS-4f ion microscope. The mean rate of ion etching was determined from the depth of the etching craters measured with the aid of a Dektar-2000 mechanical profilometer.

The measurements showed that independent of the initial composition of the solid phase, the AlAs distribution across the thickness of the layer has a characteristic minimum. As an example, Fig. 1 shows the distribution of the AlAs content across the thickness of the layer for two limiting values of $x_b$. As can be seen from Fig. 1, the AlAs content is always observed to increase near the surface of the layer. Such a dependence is the result of the combined influence on the crystallization process of superfast cooling and natural convection. As a result of superfast cooling, the change in sign of the gradient of the composition of the solid phase takes place at a significantly higher temperature in comparison with the case of equilibrium cooling. Superfast cooling also expands the initial range of compositions of the solid phase at which the distribution of the AlAs content over the thickness of the layer is not monotonic.

Capacitance–voltage measurements of the carrier concentration, performed on the surface of the layers with the aid of a mercury probe, showed that intentionally undoped GaAs layers have $n$-type conductivity with carrier concentration $\sim 10^{16}$ cm$^{-3}$, and AlGaAs layers have $\sim 10^{15}$ cm$^{-3}$.

## 4. X-RAY STRUCTURE AND PHOTOLUMINESCENCE STUDIES

X-ray diffraction studies were performed on a TPC-1 high-resolution x-ray diffractometer. To collimate the primary x-ray beam and render it monochromatic, we used a perfect single crystal of Ge(001), which is asymmetric for reflection of (004) $CuK_{\alpha 1}$ radiation and provides a divergence of the monochromatized radiation of 1 arcsec. The diffraction reflection curves (DRC's) were recorded in the vicinity of the (004) reflection for structures grown on
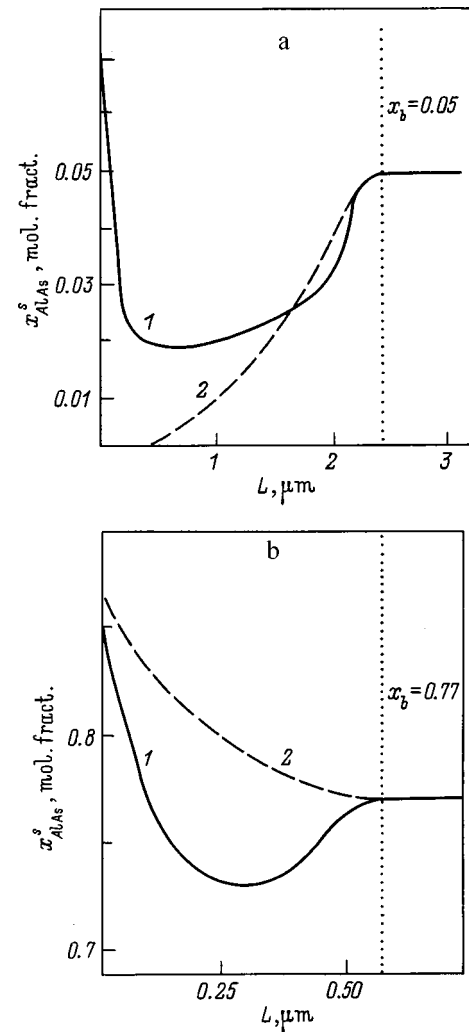


FIG. 1. Distribution of the AlAs concentration ($x$) with thickness of the $Al_xGa_{1-x}As$ layer ($L$ is distance from the surface of the layer). $T_0 = 750$ °C, $v = 320$ °C/s, temperature interval of growth 450 °C. *1* — experiment, *2* — calculation (for conditions of equilibrium cooling); a — concentration in the buffer layer $x_b = 0.05$, b — concentration in the buffer layer $x_b = 0.77$.

GaAs (001) substrates, and in the vicinity of the (333) reflection for structures grown on GaAs (111) substrates.

The degree of structural crystalline perfection of the samples was estimated from the half-widths and reflection coefficients of the diffraction peaks from the layer (layers), and of the substrate, from the presence and extent of the interference pattern.

Comparison of the shape and parameters of the calculated and experimental diffraction reflection curves, where the calculated curves were obtained in the dynamic approximation without allowance for the diffuse scattering by structural defects, allows us to state that crystallization of the obtained structures takes place in an epitaxial growth regime that is near the two-dimensional (2D) ideal. In this case there are only defects of point and cluster type, which are localized in the epitaxial layer. For the homoepitaxial structures GaAs / GaAs(001) the half-width of the diffraction maximum is 10.3 arcsec.

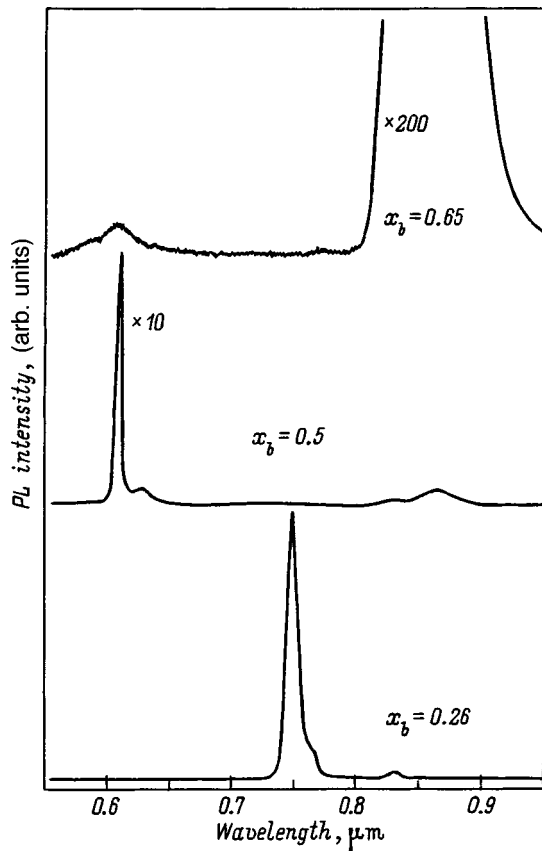A photoluminescence (PL) study was performed after the samples were pumped with an LG-106 argon laser ($\lambda$

FIG. 2. Photoluminescence spectra of $Al_xGa_{1-x}As$ structures with different AlAs content $x_b$ in the buffer layer, measured at $T = 77$ K.



FIG. 3. Photoluminescence spectra of a sample with $x_b = 0.5$, recorded at $T = 77$ K. *1* — initial spectrum, *2* — modified spectrum.

$\cong 0.5\ \mu m$) at 77 K. The excitation density was varied in the range $10-1000\ W/cm^2$.

Figure 2 shows photoluminescence spectra of samples differing in the AlAs content in their buffer layer. The photoluminescence spectrum of the sample with AlAs content in the buffer layer, $x_b \cong 0.26$ shown in Fig. 2, is typical of all structures with $x_b \leq 0.45$. As can be seen from the figure, the spectrum contains two maxima. The short-wavelength peak corresponds to radiation from the region of the sample with minimum AlAs content, and the long-wavelength maximum is due to radiation from the region of the substrate. The experimentally observed minimum values of the half-width of the short-wavelength photoluminescence maximum $\Delta h\nu$ are about 12 meV. Such $\Delta h\nu$ values are indicative of the high quality of the $Al_xGa_{1-x}As$ epitaxial layers grown under conditions of superfast cooling of the solution–melt.

The photoluminescence spectra of structures with $x_b \cong 0.6$ are characterized by the presence of a low-intensity short-wavelength band due to carrier recombination in the indirect-band material $Al_xGa_{1-x}As$.

The most interesting effects in the photoluminescence spectra were observed when examining structures with $x$ in the region of the minimum at which the dependence of the widths of the direct and indirect band gaps on the composition $[E_g^\Gamma(x)$ and $E_g^x(x)$, respectively] intersect. A typical photoluminescence spectrum of a structure of such type with $x_b = 0.5$, recorded at an excitation density $\sim 50\ W/cm^2$, is shown in Fig. 2. A more detailed photoluminescence spec-

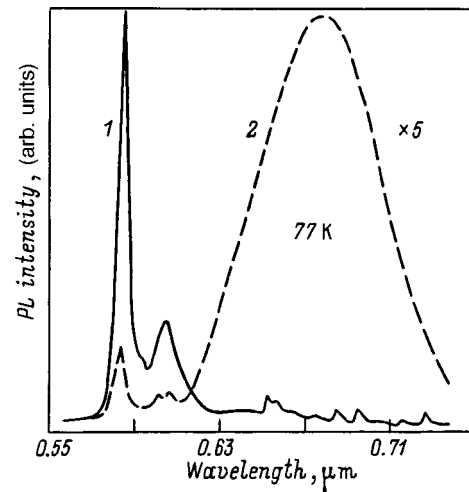trum of this structure is shown in Fig. 3. As can be seen from Fig. 3, the spectrum consists of a group of intense peaks located in a region near $\lambda \sim 0.61\ \mu m$. A similar structure was observed earlier for $Al_xGa_{1-x}As$ layers grown by gas-phase epitaxy from metallo-organic compounds, and was interpreted as photoluminescence of a bound exciton and several phonon echoes. Such spectra are characteristic of high-quality layers.

An interesting peculiarity of the investigated structures with $0.5 \leq x_b \leq 0.55$ was the irreversible change in their photoluminescence characteristics that took place when the pump density was increased to 1 kW/cm$^2$ at 77 K. The photoluminescence spectrum of a sample subjected to such ''laser annealing'' is also shown in Fig. 3. Modification of the spectrum concluded with an abrupt (roughly twentyfold) decrease in the intensity of the exciton line and phonon echoes in the region $\lambda \sim 0.61\ \mu m$ and in the appearance of a wide band ($\Delta h\nu \sim 200$ meV) with $\lambda_{max} \sim 0.69\ \mu m$. At room temperature changes in their photoluminescence characteristics were not observed. Possible mechanisms of such behavior at 77 K were considered earlier and linked with optically induced annealing of defects, and also with a recombination–generation process of impurity diffusion with formation of an arsenic vacancy ($V_{As}$)–impurity donor complex.

## 5. POTENTIAL OF USING SUPERHIGH COOLING RATES IN THE PREPARATION OF SEMICONDUCTOR LAYERS FROM SOLUTION–MELTS

The possibilities of the method, in which heterojunctions of different types of semiconductor materials were grown, were studied. The degree of crystalline perfection was estimated from the half-width of the diffraction reflection curves (Table II). It can be seen from the table that this method allows one even to obtain GaSb layers on GaP substrates, where the lattice mismatch reaches 12%. The possibility of obtaining GaAs/Si and other heterostructures with large lattice mismatch points to the fundamental possibility of creating III–V/GaAs/Si type heterostructures, including those in which an entire set of different materials is combined in one

TABLE II. Parameters of various heterostructures grown under conditions of superfast cooling of the solution–melt.

| Heterostructure | Mismatch of lattice periods, % | Half-width of DRC, arcsec. | Layer thickness, $\mu$m |
|---|---|---|---|
| GaSb/InAs | 0.6 | 73 | 1.0 |
| GaAs/GaP | 3.6 | 130 | 1.0 |
| $Al_{0.3}Ga_{0.7}As$/GaP | 3.6 | 118 | 1.0 |
| Ge/GaP | 3.7 | 145 | 1.5 |
| GaAs/Si | 4.0 | 150 | 1.5 |
| GaAs/GaAs/Si | 4.0 | 110 | 1.5 |
| Ge/Si | 4.1 | 143 | 1.0 |
| InP/GaP | 7.4 | 277 | 1.0 |
| GaSb/GaP | 11.7 | 155 | 50.0 |

heterostructure. As can be seen from Table II, increasing the number of layers grown in a GaAs/Si heterostructure leads to an increase in their degree of crystalline perfection. With the aid of this method, even for large lattice mismatch of the contacting materials, we obtained contiguous layers with thickness less than 0.1 $\mu$m. Thus, GaAs layers grown on GaP substrates had a thickness of 500 Å with a thickness of the transitional layer of 10 Å while the minimum thickness of GaAs layers on Si was 20 Å.

## 6. CONCLUSIONS

In this study we have demonstrated the possibility of using superfast cooling of a solution–melt in LPE. In the instance of the Al–Ga–As system we have shown that the application of superfast cooling makes it possible to significantly expand the possibilities of LPE for controlling the parameters of the as-grown layers. In addition, use of the given technique makes it possible to obtain heterostructures with significant (up to 12%) lattice mismatch of the contacting materials and metastable solid solutions.

The simplicity of realization of this technique and the high quality of the layers grown by it ensure the future wide application of this technique in the fabrication of semiconductor materials and devices.

[1] A. V. Abramov, N. G. Deryagin, and D. N. Tret'yakov, Semicond. Sci. Technol. **11**, 607 (1996).
[2] A. V. Abramov, N. G. Deryagin, and D. N. Tret'yakov, Semicond. Sci. Technol. **9**, 1815 (1994).

Translated by Paul F. Schippnick

# GaInAsSb/GaSb heterostructures grown in the spinodal decay region by liquid-phase epitaxy from Sb-enriched solution−melts

V. I. Vasil'ev, D. Akhmedov, A. G. Geryagin, V. I. Kuchinskiĭ,[*] I. P. Nikitina,
V. M. Smirnov, and D. N. Tret'yakov

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

Nearly isoperiodic solitary $Ga_{1-x}In_xAs_ySb_{1-y}$/GaSb heterostructures, in which the composition of the solid solution should be found inside the region of spinodal decay ($x \leqslant 0.4$), were grown by liquid-phase epitaxy from solution−melts enriched with antimony. On the basis of the results of a study of structural and luminescence properties of $Ga_{1-x}In_xAs_ySb_{1-y}$/GaSb heterostructures we have determined the main conditions ensuring reproducible growth of epitaxial layers, homogeneous in the composition of their solid solutions in the region where the existence of processes of spinodal and binodal decay have been theoretically predicted. It is shown that the magnitude and sign of the deformation which the layer undergoes during growth and also the thickness of the layer are the main factors influencing the properties of the growing GaInAsSb solid solutions in the spinodal-decay zone. © *1999 American Institute of Physics.* [S1063-7826(99)02909-9]

## INTRODUCTION

The quaternary solid solutions $Ga_{1-x}In_xAs_ySb_{1-y}$, which are isoperiodic with GaSb, are promising materials for optoelectronics in the middle infrared region. However, it is expected that the possibilities of creating devices based on these compounds are limited because of the existence in this system of a theoretically calculated, extended immiscibility zone and a region of spinodal decay inside it. In light of this circumstance, it has been assumed to be impossible to obtain solid solutions $Ga_{1-x}In_xAs_ySb_{1-y}$ isoperiodic with GaSb for $x > 0.29$ (Refs. 1 and 2). The largest values of $x$ and $y$ on the GaSb side of the isoperiodic cross section that were obtained earlier by liquid-phase epitaxy (LPE) were $x = 0.23$ and $y = 0.20$ when growing the solid solutions on GaAs(100) substrates, and $x = 0.26$ and $y = 0.23$ on GaSb(111)B substrates,[2,3] fell into the region of binodal decay but under the significantly more nonequilibrium conditions of gas-phase epitaxy were found near the spinodal decay boundary ($x = 0.3$). In the present paper we report the synthesis, with use of LPE from Sb-enriched solution−melts, the solid solutions $Ga_{1-x}In_xAs_ySb_{1-y}$ ($x \leqslant 0.4$), which are nearly isoperiodic with GaSb in the spinodal decay region.

## EXPERIMENT

Undoped epitaxial layers of quaternary solid solutions were grown from supercooled Sb-enriched solution−melts at a constant temperature. This technique of preparing the considered solid solutions was first proposed by us and is described in our published papers.[3–5] In the present work the layers were grown at temperatures from 560 to 600 °C. The degree of supercooling $\Delta T = T_l - T_g$ (where $T_l$ is the growth temperature) was varied within the limits from 8 to 120 °C.

The composition of the solid solutions in layers more than 1 $\mu$m thick was determined with the help of a Camebax x-ray micro-analyzer or by secondary-ion mass spectroscopy (SIMS) in layers whose thickness was less than 1 $\mu$m. The magnitude of the lattice mismatch of the epitaxial layer and the substrate ($f_\perp$) was measured by two-crystal x-ray diffractometry. Two-crystal rocking curves (RC) were obtained on the symmetric reflections (400) and (333) as functions of the substrate orientation. To record the photoluminescence spectra, we used a cooled PbS photoresistance. The thickness of the epitaxial layers was measured on a DECTAK profilometer.

## RESULTS AND DISCUSSION

Our studies showed that the chosen conditions of epitaxial growth can yield homogeneous layers having a mirror-smooth surface and maximal values of $x$ which depended, as expected, on the orientation of the substrates. For GaSb(100) and (111)B these values were 0.32 and 0.4, respectively. It was also found that for layers whose $x$ values corresponded to compositions lying well within the region where spinodal decay should be observed it is necessary to allow for the strong influence on the crystallization of the layers of two more factors that determine their degree of homogeneity. This follows from the fact that on the basis of the rocking curves and photoluminescence spectra these layers can be clearly separated as to their degree of homogeneity into two groups. On the rocking curves obtained from layers of the first group, only one peak, corresponding to the solid solution GaInAsSb, was observed (Fig. 1). The sign of the lattice mismatch for these layers was negative, which is analogous to their pulling in the growth plane. The values of the lattice mismatch in these layers varied from $-5 \times 10^{-4}$ to $-2 \times 10^{-3}$, from which it follows that at the growth temperature the values of the lattice mismatch were shifted by an additional $1 \times 10^{-3}$. An additional peak with a half-width of 20−
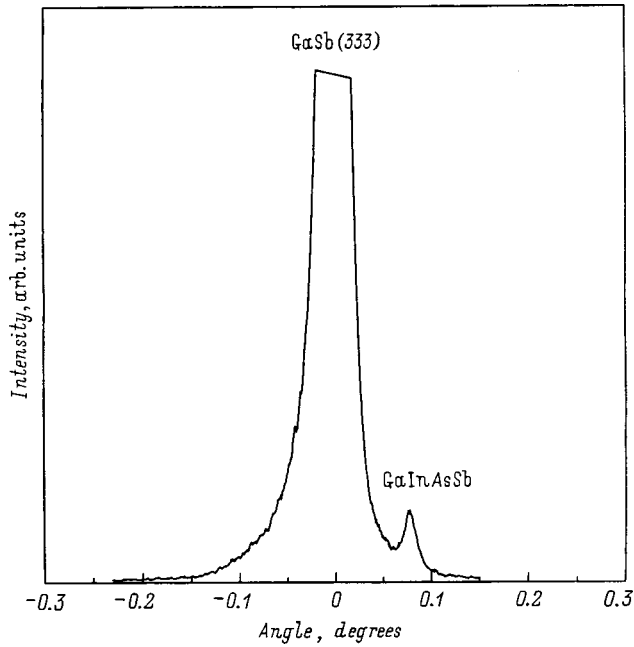
FIG. 1. X-ray diffraction rocking curves (Cu$K_{\alpha1}$ radiation) near the GaSb (333) reflection for a $Ga_{0.59}In_{0.41}As_{0.38}Sb_{0.62}$ homogeneous epitaxial layer, grown on a GaSb (111) substrate ($f_\perp = 1.46 \times 10^{-3}$).



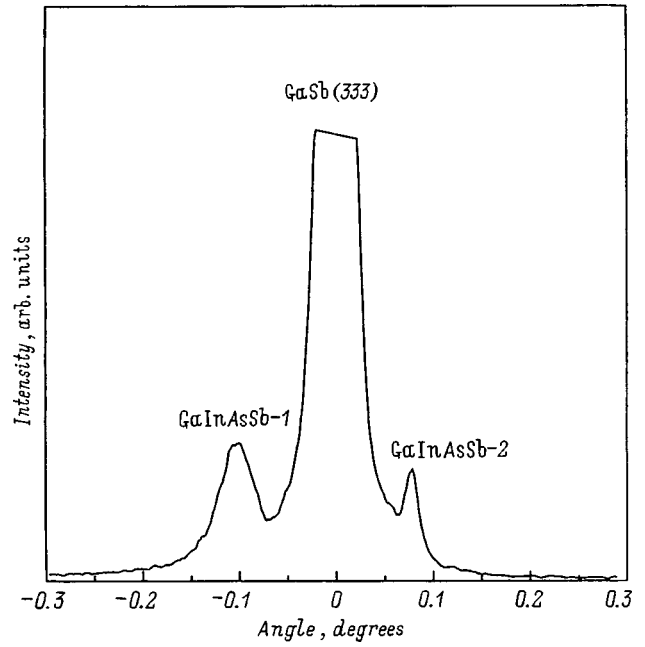FIG. 3. X-ray diffraction rocking curves (Cu$K_{\alpha1}$ radiation) near the GaSb (333) reflection for a $Ga_{0.63}In_{0.37}As_{0.36}Sb_{0.64}$ epitaxial layer, grown on a GaSb (111) substrate with indications of spinodal decay ($1 - f_\perp = +1.94 \times 10^{-3}$, $2 - f_\perp = -1.55 \times 10^{-3}$).

25 meV due to interband recombination was also present in the photoluminescence spectra of these samples (Fig. 2). The thickness of the homogeneous layers of the solid solutions, whose composition corresponds to the spinodal decay zone, does not exceed 500 Å. Note that regardless of the degree of

homogeneity of the layers, their rate of growth decreased rapidly from 3 $\mu$m/min at $x = 0.2$ to 500 Å/min at $x = 0.4$ as the composition shifted into the region of spinodal decay. On the rocking curves obtained from layers of the second group,
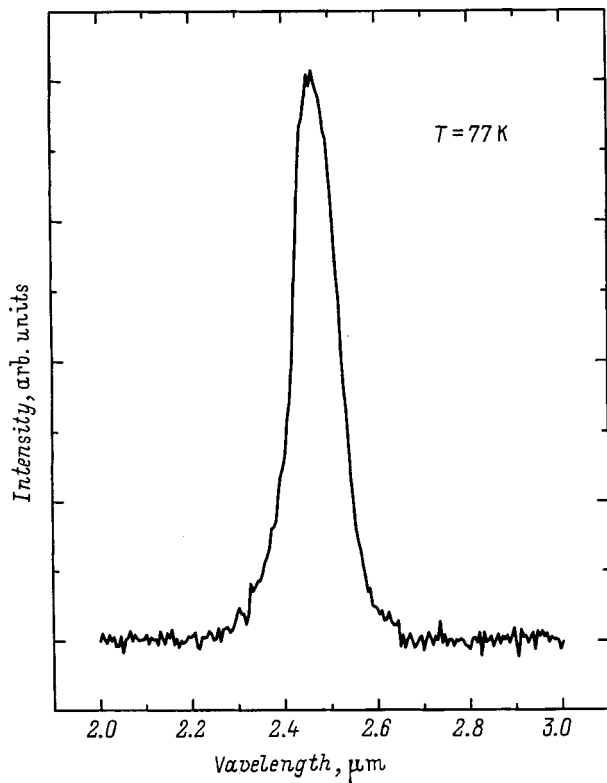


FIG. 2. Photoluminescence spectrum of a $Ga_{0.59}In_{0.41}As_{0.38}Sb_{0.62}$/GaSb heterostructure.
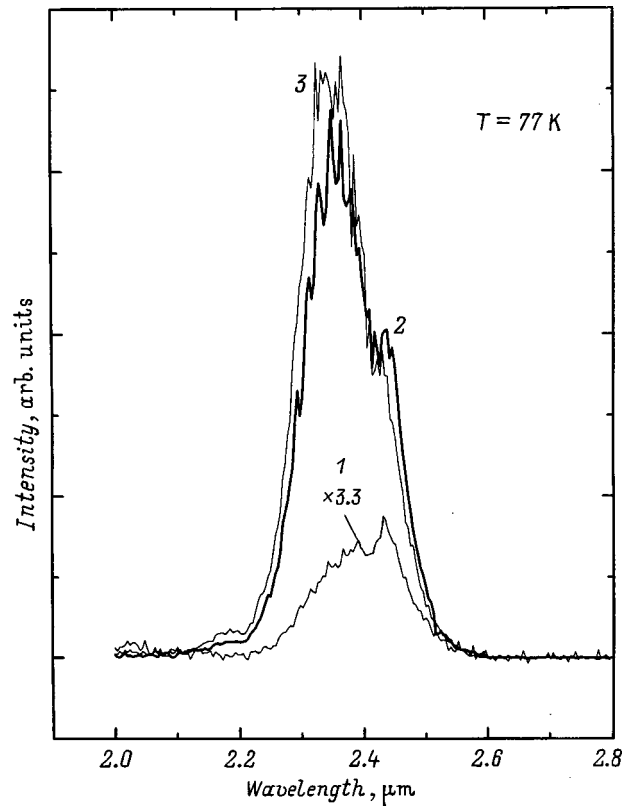


FIG. 4. Photoluminescence spectra of epitaxial layers of the solid solution $Ga_{0.63}In_{0.37}As_{0.36}Sb_{0.64}$/GaSb with different thicknesses, Å: *1* — 1000, *2* — 360, *3* — 200.

two peaks were observed for the solid solutions with negative and positive lattice mismatch (Fig. 3). In other words, in these layers, whose composition also corresponds to the spinodal decay zone, the solid solution undergoes decay. The decay of the solid solution is accompanied by a broadening of the interband recombination band in the photoluminescence spectra to 50 meV. The thickness of these layers is usually greater than 500 Å. It should be emphasized that decay of the solid solution was not observed in layers whose composition lay outside the spinodal decay zone, regardless of their thickness.

More detailed studies of the effect of thickness on the structural and optical properties of the layers for given compositions of the liquid phase, growth temperature, and degree of supercooling shows that when its value exceeds $\sim 200$ Å, an additional peak from the solid solution with positive lattice mismatch appears in the x-ray rocking curves, and its intensity grows with thickness. Simultaneously, a decrease in the intensity and an increase in the half-width of the interband recombination band in the photoluminescence spectra was observed (Fig. 4). This indicates that with growth of the $Ga_{1-x}In_xAs_ySb_{1-y}$ layers (i.e., with increasing distance from the heteroboundary) in the region of compositions where spinodal decay should occur, the positive influence of the tensile stress along the heteroboundary weakens. Thus, the results obtained in this work and in our previous papers[3,6] allow us to conclude that in the epitaxial growth of solid solutions with compositions $Ga_{1-x}In_xAs_ySb_{1-y}$ from Sb-enriched solution–melts it is possible to simultaneously control the properties of heterostructures by varying the degree of inhomogeneity not only in the direction perpendicular to the heteroboundary, but also in the direction parallel to it, where this is achieved by prescribing the lattice mismatch at the heteroboundaries and the thickness of the layer and its composition.

## CONCLUSIONS

In the present paper we have demonstrated the possibility of varying the degree of inhomogeneity of epitaxial layers by prescribing the lattice mismatch and the growth time. We have found that a tensile stress in the layer parallel to the heteroboundary and a small thickness ($0.02-0.05$ $\mu$m) are prerequisites for obtaining homogeneous solid solutions in the spinodal decay zone.

*)E-mail: viv.kuch@pop.ioffe.rssi.ru

[1] A. M. Litvak and N. A. Charykov, Zh. Fiz. Khim. **66**, 923 (1992).
[2] E. Tournie, J.-L. Lazzari, F. Pitard, C. Alibert, A. Joulie, and B. Lambert, J. Appl. Phys. **68**, 5936 (1990).
[3] A. G. Deryagin, N. N. Faleev, V. M. Smirnov, G. S. Sokolovskii, and V. I. Vasil'evm, IEE Proceedings Optoelectron. **144**, 438 (1997).
[4] V. I. Vasil'ev, V. V. Kuznetsov, and V. A. Mishurnyĭ, Izv. Akad. Nauk SSSR, Ser. Neorg. Mater. **26**, 23 (1990).
[5] V. I. Vasil'ev, V. V. Kuznetsov, V. A. Mishurnyi, V. V. Sazonov, and N. N. Faleev, in *Proceedings of the First International Conference on Epitax. Growth* (Budapest, 1990), p. 659.
[6] V. I. Vasil'ev, F. G. Deryagin, V. I. Kuchinskiĭ, V. M. Smirnov, G. S. Sokolovskiĭ, D. N. Tret'yakov, and N. N. Faleev, Pis'ma Zh. Tekh. Fiz. **24**(6), 58 (1998) [Tech. Phys. Lett. **24**(6), 45 (1998)].

Translated by Paul F. Schippnick

# Optoelectronic images of polycrystalline thin-film solar cells based on CuInSe$_2$ and CuInGaSe$_2$ obtained by laser scanning

G. A. Medvedkin*[)]

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

L. Stolt and J. Wennerberg

*Uppsala University, Angström Solar Center, P. O. Box 534, SE-75121 Uppsala, Sweden*

The laser scanning technique was used to obtain two- and three-dimensional optoelectronic images of polycrystalline solar cells based on thin films of CuInSe$_2$ and CuInGaSe$_2$. Topograms obtained with the aid of the laser-beam-induced current reveal microregions with reduced photovoltaic efficiency and provide a detailed picture of the distribution of hidden inhomogeneities over the entire active surface of the solar cell. Gradation of the microdefects with intensity and size was achieved by post-experimental graphic and false-color processing of the obtained three-dimensional images. © *1999 American Institute of Physics.*[S1063-7826(99)03009-4]

## INTRODUCTION

For polycrystalline optoelectronic devices, an analysis of homogeneity represents an important task since the films consist of a large number of grains and their photoresponse is obviously inhomogeneous near each crystallite. Total illumination of a solar cell gives only the integrated output characteristics, and the average photoresponse is decreased due to distributed macro- and microdefects. Laser scanning of the surface provides information at each point of the device and can also give a detailed distribution of the photovoltaic efficiency over the entire receiver area. The goal of the present work was to examine invisible defects and inhomogeneities of solar cells on the micro- and mini-scales.

## EXPERIMENTAL PART

High-efficiency polycrystalline thin-film solar cells with the device structure Mo/Cu(In,Ga)Se$_2$/CdS/ZnO/ZnO:Al were prepared on sodium-calcium carbonate glass. We examined solar cells with 12–13% efficiency and with the absorber materials CuInSe$_2$ (CIS) and CuIn$_{0.7}$Ga$_{0.3}$Se$_2$ (CIGS).

We used a laser scanner which allowed us to measure the photoresponse (the $V_{oc} i_{sc} V_R$ regimes) and obtain matrices of experimental data for post-experimental construction of two- and three-dimensional images of these solar cells. The setup consisted of two He–Ne lasers ($\lambda = 632.8$ and 1152 nm), an optical focusing channel, a specular deflecting system, and an electronic amplifying circuit intended for detecting a phase-modulated or constant optical signal. The focusing and polarization optics narrowed the probe beam down to a diameter of $50-80 \, \mu$m, and the scanned area with minimum optical distortion had dimensions of $30 \times 30$ mm. The red and infrared lasers excited regions in the solar cell at different depths and therefore gave information about the surface layer and about layers far from the CdS/CIS (CIGS) heterojunction. Since the crystallites and their faces were oriented at various angles in the films, the implemented definite linear polarization of the incident beam created a higher-contrast image of the microdefect distribution. Scanning and data collection were controlled by a personal computer using the program Solscan; further processing, including graphical processing, was done on a Mac computer using the program Matlab or on a PC with the programs Origin and Excel.

## 1. RESULTS AND DISCUSSION

A grid topogram, obtained in red light for the CIGS active plane of the solar cell, is shown in Fig. 1. Hidden microdefects varying in size from 40 to 400 $\mu$m are present, as it turned out, in a larger quantity on one active band than
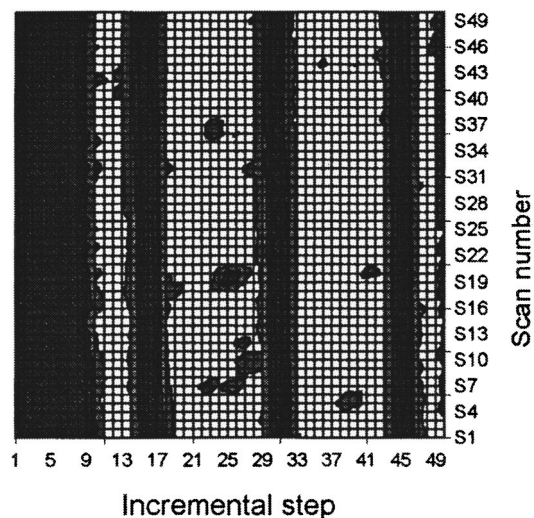


FIG. 1. Topogram of a CIGS solar-cell receiver surface in red light. The dark regions correspond to a reduced photoresponse and the light regions correspond to maximum photoresponse. The vertical bands correspond to the current collector of the Al grid and the the dark spots correspond to defects.
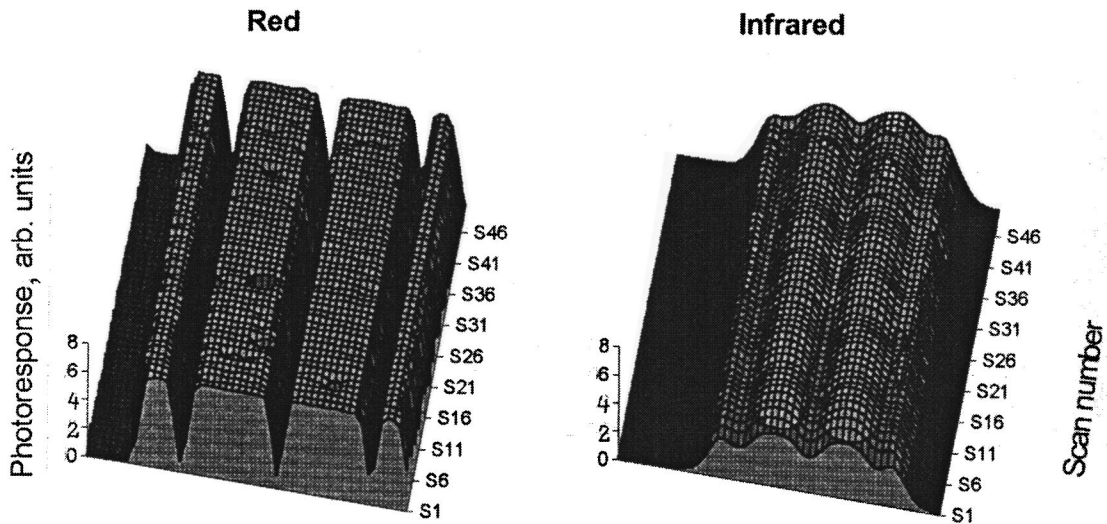
**Red**

**Infrared**

FIG. 2. 3D optoelectronic images of a CIGS solar cell in red and infrared light.

on the other. Thus, the topogram reveals the distribution, size, and shape of the photoactive defects on each receiver surface making an individual contribution to the total output signal of the solar cell.

A three-dimensional optoelectronic image of a polycrystalline thin-film solar cell makes it possible to demonstrate the degree of perfection of the device and lack thereof more distinctly. Figure 2 shows 3D optoelectronic images of the CIGS solar cell in red and infrared light, respectively. A series of defects varying from 10 to $100\,\mu$m in size was reliably detected by the red laser beam (Fig. 2, left image). These photoactive defects are distributed in the upper layer of the absorber film. The infrared beam, however, does not see these defects since it carries information about bulk properties of the absorber-film. In the case of the CIGS wideband semiconductor, photogeneration occurs not only in the bulk of the film, but also on the rear Mo contact. The infrared photoresponse of the examined CIGS solar cells has a smooth distribution, which is evidence of a high bulk homo-geneity of the absorber-film. The absence of a negative signal over the entire active area confirms that no generation takes place on the rear contact.

Differences are observed in the optoelectronic images of solar cells of CIS type possessing $I-V$ characteristics with ''red'' rollover on the direct branch and without it. Figure 3 shows the stronger fluctuations of the photoresponse over the entire sensitive area of the solar cells having rollover of their $I-V$ characteristics. The structures with chemical deposition of CdS possess a rollover effect due to the low concentration of free electrons and the high concentration of trap states in the CdS layer.[3,4] Therefore, the concentration of free electrons in the CdS layer should sense the modulation of the current as a result of photogeneration. Thus, the influence of the concentration of the dopant level in the CdS buffer layer is manifested in the $I-V$ curves and in laser scanning.

Longitudinal inhomogeneities in CIS and CIGS solar cells were also found when using red and/or infrared laser excitation.[5] Such inhomogeneities are characteristic of the
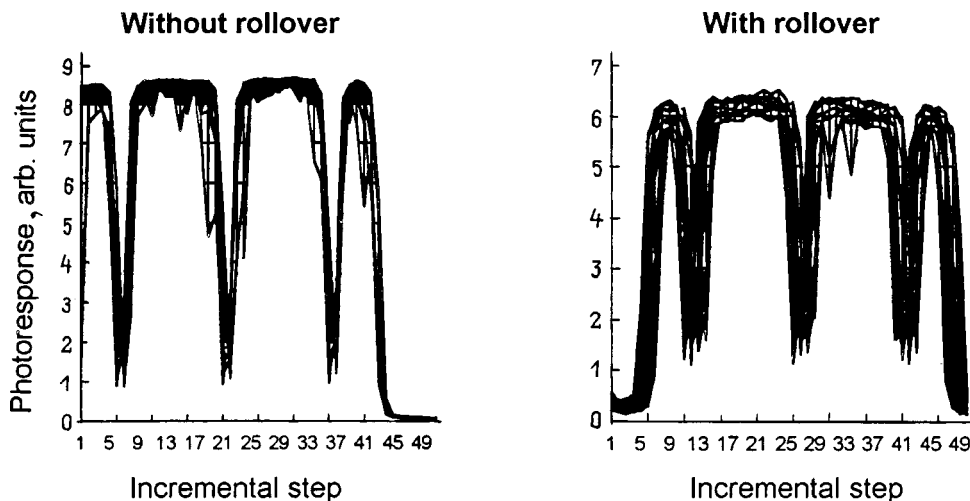
**Without rollover**

**With rollover**

FIG. 3. Transverse multiscanning of a CIS solar cell with a ''good'' $I-V$ characteristic and a $I-V$ characteristic with ''rollover.''

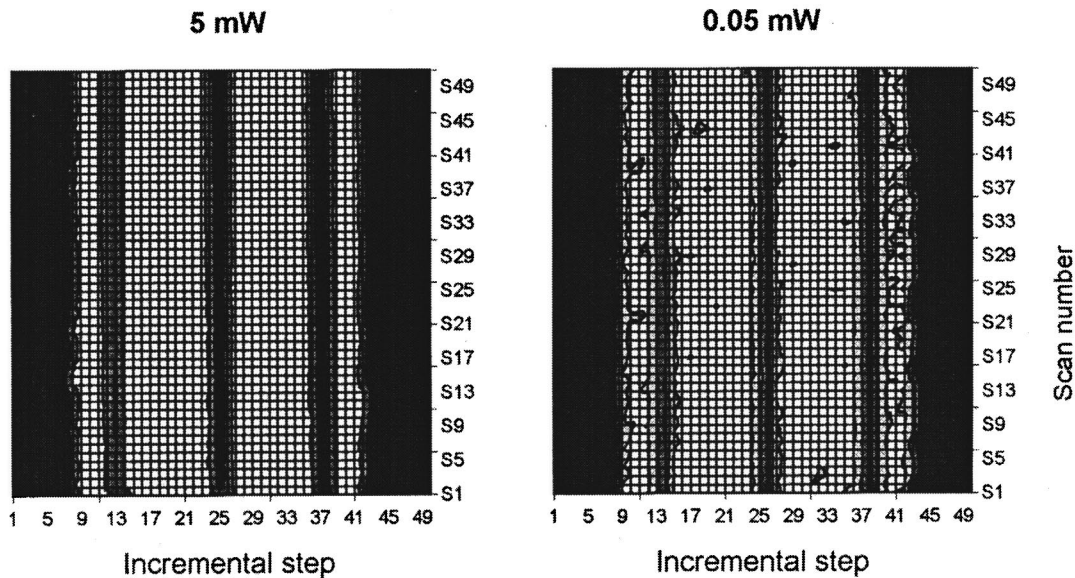**5 mW**                                              **0.05 mW**

FIG. 4. Topogram of photoactive microdefects in CIS solar cells in red light. Recombination channels are saturated at a high level of illumination (left) and unsaturated at a low level of illumination (right).

bulk of the absorber-film, are manifested on the millimeter and centimeter scale, and are stretched out along the grid stripes.

The presence of electron traps or metastable states in the CdS layers and compounds with ordered vacancies, through which recombination proceeds intensely, usually causes the total photoresponse to decrease. Controlling the balance between generation and recombination by varying photogeneration of nonequilibrium charge carriers, it is possible to saturate or free up the recombination channels and, consequently, reveal the regions with enhanced concentration of recombination centers. We exploited this peculiarity in order to detect local inhomogeneities in the CIS and CIGS solar cells. Experiments with a phase-modulated signal allowed us to use a wide power range of the monochromatic light equivalent to the power of sunlight $(0.5 - 600\,\mathrm{Suns})$ to reveal local inhomogeneities with the aid of the contrast image. Figure 4 shows a grid topogram in red light: regions with an enhanced recombination rate are visible at a low level of illumination and fall out as the light power grows while the recombination channels saturate. Such defects lower the photoresponse in certain spectral ranges corresponding to the energies of the levels. As was reported earlier,[6–8] comparison of the optoelectronic image (after laser scanning) and a typical surface image observed in an optical microscope does not give satisfactory agreement when comparing the distribution of inhomogeneities such as pinholes (or similar ''mechanical'' microdefects) in the solar cell. We confirmed this statement with our own microphotographs made with the aid of an optical microscope. Therefore, the microdefects observed with the help of a current induced by a laser beam may be associated with the regions of reduced no-load voltage as a result of enhanced concentration of recombination centers.

## CONCLUSIONS

Photoactive microdefects varying in size from 10 to $100\,\mu$m as well as elongated inhomogeneities in the millimeter and centimeter ranges were identified and differentiated in strength and size with the help of two- and three-dimensional optoelectronic images obtained by red and infrared laser scanning of thin-film solar cells.

*)E-mail: gen@medv.ioffe.rssi.ru

[1] L. Stolt, in *Proceedings of the 9th International Photovoltaic Science and Engineering Conference*, Miyazaki, Japan (1996), Vol. 9, p. 135.

[2] J. Hedström, H. Ohlsén, M. Bodegård, A. Kylner, L. Stolt, D. Hariskos, M. Ruckh, and H. W. Schock, in *Proceedings of the 23rd IEEE Photovoltaic Spec. Conf.*, Louisville (1993), p. 364.

[3] J. Hou, S. J. Fonash, and J. Kessler, in *Proceedings of the 25th IEEE Photovotaic Specialists Conf.*, Washington, D.C. (1996).

[4] A. Kylner, J. Lindgren, and L. Stolt, J. Electrochem. Soc. **143**, 2662 (1996).

[5] G. A. Medvedkin and J. Wennerberg, in *Proceedings of the International Conference POLYSE'98*, Schwabisch Gmund, Germany (1998) O4.

[6] I. L. Eisgruber and J. R. Sites, Progress in Photovolt., Res. Appl. (UK) **4**, No. 1, 63 (1996).

[7] I. L. Eisgruber, R. J. Matson, J. R. Sites, and K. A. Emery, in *Proceedings of the 1st World Conference on Photovoltaic Energy Conversion*, Hawaii, USA (1994), Vol. 1, p. 283.

[8] R. J. Matson, K. A. Emery, I. L. Eisgruber, and L. L. Kazmerski, in *Proceedings of the 12th European Photovoltaic Solar Conference*, Bedford, UK (1994), p. 1222.

Translated by Paul F. Schippnick

# Conductance of a quantum wire in a parallel magnetic field

V. A. Geĭler*) and V. A. Margulis

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1141–1143 (September 1999)

Ballistic electron transport in a three-dimensional quantum wire with elliptic cross section is investigated. The potential of the single-particle Hamiltonian of the system under consideration was chosen to be parabolic. Using the Landauer–Buttiker formalism, we find an expression for the conductance at zero temperature. We show that the number and width of the steps in the dependence of the conductance on the electron energy are determined by the ratio of the characteristic frequencies of the potential. In the case of nonzero temperature we show that the conductance consists of two terms. The first is monotonic and depends quadratically on the energy; the oscillating second term gives sawtooth-shaped peaks. The height of the conductance steps is equal to the conductance quantum, and the width of the plateau depends on the energy, the field, and the frequency ratio. We stress that the picture of the conductance is extremely sensitive to the ratio of hybrid frequencies. © *1999 American Institute of Physics.*
[S1063-7826(99)03109-9]

As was shown in Refs. 1 and 2, in a three-dimensional quantum wire whose diameter is on the order of the Fermi wavelength of the electron $\lambda_F$, the dependence of the conductance $G$ on the electron energy $\varepsilon$ (or, which is physically equivalent, on the diameter of the wire) has a step-like character. The height of each step in the dependence $G(\varepsilon)$ is equal to the conductance "quantum" $G_0 = 2e^2/h$ (quantization of conductance). The quantization effect, which arises in the ballistic transport regime, is affected by two factors: first, the temperature of the electron gas, and second, the geometry of the system. In addition, a magnetic field applied to the wire intensifies the dimensional confinement of the electron in a plane, which leads to a dependence on the magnetic field $B$ of the physical presentation of the effect.

The conductance of such a ballistic nanostructure (quantum wire), which connects two macroscopic electron reservoirs, can be described at zero temperature using the Landauer–Buttiker formula

$$\frac{G}{G_0} = \sum_{\alpha\alpha'} T_{\alpha\alpha'}, \tag{1}$$

where $T_{\alpha\alpha'}$ is the probability of the transition from the state $|\alpha\rangle$ to the state $|\alpha'\rangle$. As the temperature is raised, the quantization steps smear out; therefore, the conductance quantization effect is observed only at very low temperatures $T \simeq 1$ K.

The geometry of a quantum wire also has an effect on ballistic electron transport.[3–5] It is important to allow for both the finite length of the wire, which leads to reflection of electron modes back into the nanostructure,[5] and the shape of the cross section of the wire.[3] Different models have been used to describe the potential that keeps the electrons inside the wire.[6–10] However, as was noted in Refs. 9 and 11, the simple parabolic approximation fits better into a self-

consistent scheme which takes into account the Coulomb effects that were examined in detail in Refs. 11–13.

In this study we examined in the model of a parabolic potential ballistic electron transport in a quantum wire situated in a longitudinal magnetic field,

$$V(x,y) = \frac{m^*}{2}(\omega_1^2 x^2 + \omega_2^2 y^2). \tag{2}$$

Here $m^*$ is the effective mass, and the frequencies $\omega_j$ ($j = 1,2$) of the potential define the semiaxes of the elliptical cross section of the wire $l_j = (1/2)\sqrt{\hbar/m^*\omega_j}$.

The spectrum in this model is described by the formula

$$\varepsilon_\alpha = \hbar\Omega_1(n_1 + 1/2) + \hbar\Omega_2(n_2 + 1/2) + p^2/2m^*. \tag{3}$$

Here $\alpha = |p,n_1,n_2\rangle$, $p$ is the momentum of the electron along the $z$ axis, the quantum numbers $n_1, n_2 = 0,1,2,\ldots$; and the hybrid frequencies $\Omega_{1,2}$ are given by the formula

$$\Omega_{1,2} = \frac{1}{\sqrt{2}}\{\omega_c^2 + \omega_1^2 + \omega_2^2 \pm [(\omega_c^2 + \omega_1^2 + \omega_2^2)^2 - 4\omega_1^2\omega_2^2]^{1/2}\}^{1/2},$$

where $\omega_c$ is the cyclotron frequency.

The coefficients $T_{\alpha\alpha'}$ for the states $|\alpha\rangle$ have the form $T_{\alpha\alpha'} = \delta_{n_1 n_1'}\delta_{n_2 n_2'}$. Using Eq. (1), we obtain for the conductance of the wire at $T=0$ the expression

$$\frac{G(\varepsilon)}{G_0} = \sum_{n=0}^{N}\left[\frac{\Omega_1}{\Omega_2}(n+\delta)\right] + N + 1, \tag{4}$$

where $N$ and $\delta$ denote, respectively, the integer and fractional part of the number $(2\varepsilon - \hbar\Omega_1 - \hbar\Omega_2)/2\hbar\Omega_1$, where $\varepsilon$ is the energy of the electron. Since $\Omega_1 \neq \Omega_2$, the quantity $\Omega_1\delta/\Omega_2$ varies as $\varepsilon$ or $B$ varies. This means that both the number of terms and their magnitude vary. Consequently, the width of the plateau and the number of steps in the de-
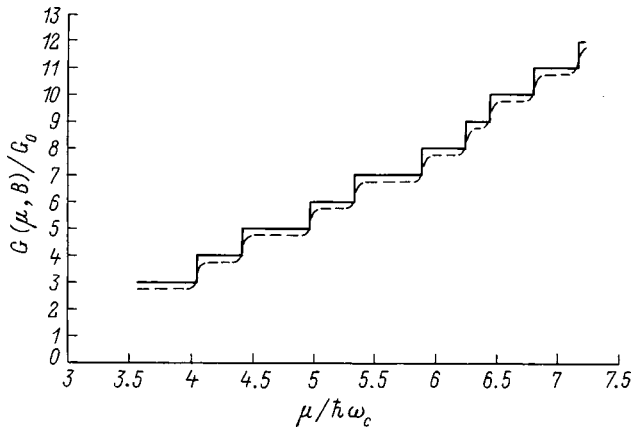
FIG. 1. Dependence of the conductance on the chemical potential $\mu$. Calculational parameters: $\Omega_1 = 1.7 \times 10^{13} \, \text{s}^{-1}$, $\Omega_2 = 0.83 \times 10^{13} \, \text{s}^{-1}$, $B = 3$ T; solid line — $T = 0$ K, dashed line — $T = 1$ K.

pendence $G(\varepsilon)$ depend on the relationship between the frequencies in the potential (2) and the magnitude of the magnetic field. This result is represented in Figs. 1 and 2 by the solid line.

Let us now consider thermal smearing of the conductance steps. We make use of the following observation in the calculations. The number of conductance quanta for an electron with energy $\varepsilon$ is equal to the number of electron states with energy less than or equal to $\varepsilon$ for a two-dimensional gas of oscillators with frequencies $\Omega_1$ and $\Omega_2$. The classical partition function for such a gas is[14]

$$Z^{-1} = 4 \sinh\left(\frac{\hbar\Omega_1}{2T}\right) \sinh\left(\frac{\hbar\Omega_2}{2T}\right). \tag{5}$$

The number of states $\nu(\varepsilon) = G/G_0$ can be expressed in terms of the partition function (5) by means of the following formula derived in Ref. 14:

$$\nu(\varepsilon) = \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} Z(\zeta) e^{\varepsilon\zeta} \frac{d\zeta}{\zeta}, \tag{6}$$

where $\alpha > 0$ and $\zeta = 1/T$. We see from (6) that the integral here is determined by the poles that lie on the imaginary axis. For commensurate frequencies the poles are multiple poles at the points $n\Omega_1 = m\Omega_2$, where $n$ and $m$ are integers,
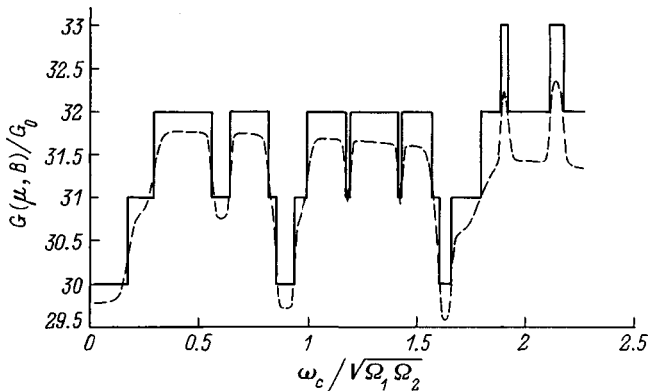


FIG. 2. Dependence of the conductance on the magnetic field. Chemical potential $\mu = 10^{-13}$ erg.

and for incommensurate frequencies all poles are simple. Since a real number has probability 1 of being irrational, in what follows we restrict the discussion to the more interesting case where the frequencies are incommensurate. In this case, calculation of the integral (6) gives the following expression for $\nu(\varepsilon) = G(\varepsilon, 0)$:

$$\frac{G(\varepsilon, 0)}{G_0} = \frac{1}{2\pi\hbar^2\Omega_1\Omega_2}\left(\varepsilon^2 - \frac{\hbar^2\Omega_1^2 + \hbar^2\Omega_2^2}{4}\right) + \frac{1}{2\pi}$$

$$\times \sum_{n=1}^{\infty} (-1)^{n+1}\left[\frac{\cos(2\pi\varepsilon/\hbar\Omega_1)}{\sin(\pi n\Omega_2/\Omega_1)}\right.$$

$$\left. + \frac{\cos(2\pi\varepsilon/\hbar\Omega_2)}{\sin(\pi n\Omega_1/\Omega_2)}\right]. \tag{7}$$

Here the Fourier series is determined by the contribution from the poles on the imaginary axis, and the first term in the expression is determined by the contribution of the third-order pole at zero. It is convenient to expand the conductance in a Fourier series in order to take into account the dependence of the conductance on temperature:

$$\frac{G(\varepsilon, T)}{G_0} = \int_0^{\infty} G(\varepsilon, 0) \frac{\partial f}{\partial \mu} d\varepsilon. \tag{8}$$

Here $f(\varepsilon)$ is the Fermi function. Assuming that the chemical potential of the electron gas in the wire $\mu \gg T$, it follows from Eq. (8) that

$$\frac{G(\mu, T)}{G_0} = \frac{1}{2\hbar^2\Omega_1\Omega_2}\left[\mu^2 + \frac{\pi^2 T^2}{3}\right.$$

$$\left. - \frac{1}{4}\hbar^2(\Omega_1^2 + \Omega_2^2)\right] + \pi T \sum_{n=1}^{\infty} (-1)^{n+1}$$

$$\times \left[\frac{1}{\hbar\Omega_1} \frac{\cos(2\pi n\mu/\hbar\Omega_1)}{\sinh(2\pi^2 nT/\hbar\Omega_2)\sin(\pi n\Omega_2/\Omega_1)}\right.$$

$$\left. + \frac{1}{\hbar\Omega_2} \frac{\cos(2\pi n\mu/\hbar\Omega_2)}{\sinh(2\pi^2 nT/\hbar\Omega_2)\sin(\pi n\Omega_1/\Omega_2)}\right]. \tag{9}$$

Dependences constructed with formula (9) are represented in Figs. 1 and 2 by the dashed lines. It can be seen from these figures that thermal smearing causes the plateaus of the conductance steps to slope and smooths out the thresholds.

As follows from (4) and Figs. 1 and 2, the height of the conductance steps for any ratio of the frequencies is equal to the conductance quantum, and the length of the plateau varies as a function of the energy of the magnetic field and the ratio of frequencies. The number of steps in the same energy interval (or the same interval of field values) depends on the position of the interval on the corresponding axis and on the ratio of frequencies $\omega_{1,2}$. Thus, a relatively small variation of the field (or frequencies) can lead to a fundamental change in the conductance picture (see Figs. 1 and 2). Also, at $T \neq 0$ the plateaus of the steps are slightly shifted downward and acquire a slope.

Note that the use of Eqs. (7) and (9) instead of the more commonly used starting expressions (4) and (8) is more con-

venient for analysis. In (7) and (9) we have separated the monotonic part of the dependence $G(\varepsilon)$ and the oscillating part – the series in (7) and (9), which consist of sawtooth-shaped peaks that account for the steps. As the temperature is increased, the series rapidly decay. At $T > 10$ K the monotonic part is 2–3 orders of magnitude greater than the oscillating part, and in this temperature region the steps are smoothed out.

*)E-mail: geyler@mrsu.ru

———————

[1] U. Landman, W. D. Luedtke, and N. A. Burnham *et al.*, Science **248**, 454 (1990).

[2] E. N. Bogachek, A. M. Zagoskin, and I. O. Kulik, Fiz. Niz. Temp. **16**, 1404 (1990) [Sov. J. Low Temp. Phys. **16**, 796 (1990)].

[3] C. J. Ford, S. Washburn, M. Buttiker *et al.*, Phys. Rev. Lett. **62**, 2724 (1989).

[4] E. N. Bogachek, A. G. Scherbakov, and U. Landman, Phys. Rev. B **56**, 14917 (1977).

[5] E. Tekman and S. Ciraci, in *Science and Engineering of One- and Zero-Dimensional Semiconductors*, edited by S. P. Beaumomp and C. M. Sotomajor Torres (Plenum Press, New York, 1990), p. 99.

[6] C. S. Chu and R. S. Sorbello, Phys. Rev. B **40**, 5941 (1989).

[7] M. Buttiker, Phys. Rev. B **40**, 7906 (1990).

[8] T. Martin and S. Feng, Phys. Rev. Lett. **64**, 1971 (1990).

[9] C. W. J. Beenakker and H. Van Houten, in *Solid State Physics*, edited by H. Ehrenrech and D. Tornbull (Academic Press, New York, 1991), Vol. 44, p. 83.

[10] B. J. Van Wees, H. Van Houten, C. W. J. Beenakker *et al.*, Phys. Rev. Lett. **60**, 848 (1988).

[11] V. B. Shikin, Zh. Éksp. Teor. Fiz. **101**, 1599 (1992) [Sov. Phys. JETP **74**, 852 (1992)].

[12] V. B. Shikin, JETP Lett. **50**, 212 (1989).

[13] V. B. Shikin, T. Demel', and D. T. Khaĭtman, Zh. Éksp. Teor. Fiz. **96**, 1406 (1989) [Sov. Phys. JETP **69**, 797 (1989)].

[14] V. A. Geyler and V. A. Margulis, Phys. Rev. B **55**, 2543 (1997).

Translated by Paul F. Schippnick

# Quantization of magnetic induction in a 2D system under conditions of the quantum Hall effect

G. G. Zegrya

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

In this paper we show that in a 2D system situated in an external transverse magnetic field $H$ the magnetic induction $B$ can under certain conditions, as a result of the de Haas–van Alfvén effect, take only certain discrete values, i.e., it is ''quantized.'' In this case the dependence $B(H)$ consists of jumps and a plateau. At the plateau $B(H) = $ const. As a consequence of quantization of the magnetic induction $B(H)$, quantization of the Hall resistivity $\rho_{xy}(H)$ is possible. © *1999 American Institute of Physics.* [S1063-7826(99)03209-3]

## 1. INTRODUCTION

Systems of two-dimensional (2D) electrons in a transverse quantizing magnetic field continue to attract attention in connection with observation of the quantum Hall effect (QHE).[1] During the last several years, attention has been focused on investigating the ground state of 2D electrons at low temperatures and large values of the external magnetic field.[2–9] It is assumed that under certain conditions the ground state of the 2D electrons in a strong quantizing magnetic field is ferromagnetic.[8] However, a qualitative and quantitative analysis of the effect of strong magnetism of the 2D electrons on the kinetic phenomena of such systems has not been carried out. Because of the possibility of strong magnetism of the 2D electrons, which magnetic field quantity—the magnetic induction $B$ or the magnetic field intensity $H$—should enter into the expression for the Hall ($\sigma_{xy}$) and the ohmic ($\sigma_{xx}$) conductivity of the 2D electrons in the quantum Hall effect?

The electrons in a 2D system move along Larmor orbits in the $xy$ plane (the magnetic field is directed along the $z$ axis—QHE geometry). Therefore, the magnetic field created by them is averaged over regions with dimensions on the order of the Larmor radius $r_L$ (Ref. 11). The mean distance between the electrons is less than or on the order of $r_L$. Consequently, the field acting on the 2D electrons is the mean microscopic field, i.e., the induction $B$.[10–13] Thus, for large values of the magnetic susceptibility $\chi$, the magnetic induction $B$ should be substituted for the magnetic field strength $\chi$ in the formulas for $\sigma_{xx}$ ($\rho_{xx}$) and $\sigma_{xy}$ ($\rho_{xy}$), and also for the magnetic moment $M$. Here the dependence of the induction $B$ on the macroscopic magnetic field strength $H$ [$B(H)$] is found from the equation[13]

$$H = B - 4\pi M(B). \tag{1}$$

In order to solve this equation, it is necessary to know the dependence $M(B)$.

## 2. THE de Haas–van Alfvén EFFECT FOR 2D CHARGE CARRIERS IN THE FERMI-LIQUID MODEL

In strong magnetic fields two-dimensional electrons can be considered as a unique sort of Fermi liquid. In such a 2D system excitation interaction between the electrons plays an important role at low temperatures. The de Haas–van Alfvén effect for 2D electrons in a transverse magnetic field is considered below in the framework of the isotropic Fermi-liquid model. (The de Haas–van Alfvén effect for a 2D gas of carriers at $T=0$ was first considered by Peierls.[14] Peierls's result was analyzed in detail in Refs. 11 and 15. It states that the magnetic moment of the 2D electrons $M$ varies with a jump each time the Fermi level $E_F$ intersects a Landau level or when the Landau filling factor is equal to an integer value. The de Haas–van Alfvén effect at finite temperatures was considered in Ref. 16 in the framework of the model of noninteracting electrons.) We assume that at the large concentrations of 2D carriers of interest to us the long-range part of the Coulomb interaction is screened. Consequently, the isotropic Fermi-liquid model can be applied to the 2D electrons.[17,18] On the whole, in application to spatially homogeneous states, the Landau theory of a Fermi liquid[19] is also valid in the case of the Coulomb interaction between particles.[17]

Since elementary excitations possessing spin have a magnetic moment, a Fermi liquid should be a magnet.[11,19,20] In the isotropic case, the liquid can be a paramagnet or a ferromagnet.[20] Below we show that Fermi-liquid effects have a substantial influence on the paramagnetic part of the susceptibility.[18,19]

Following Refs. 11 and 18, we obtain the following expression for the oscillating part of the magnetic moment of the 2D electrons:

$$M = \frac{\pi}{1+Z_0} \frac{eT}{\hbar c d} \sum_{k=1}^{\infty} \frac{1}{\sinh(2\pi^2 kT/\hbar\omega_c)} \sin\left(2\pi k \frac{B_0}{B}\right), \tag{2}$$

where $d$ is the width of the quantum well, $B_0 = \Phi_0 n_e \equiv hcn_e/e$, $\Phi_0$ is the magnetic flux quantum, $n_e$ is the two-dimensional electron density, $e$ is the charge of the electron,
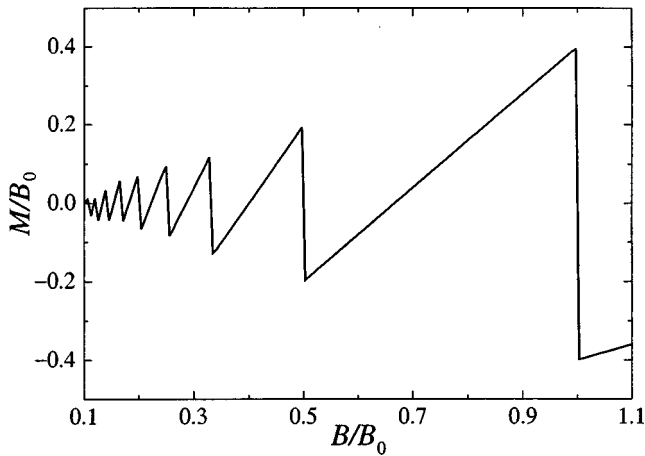
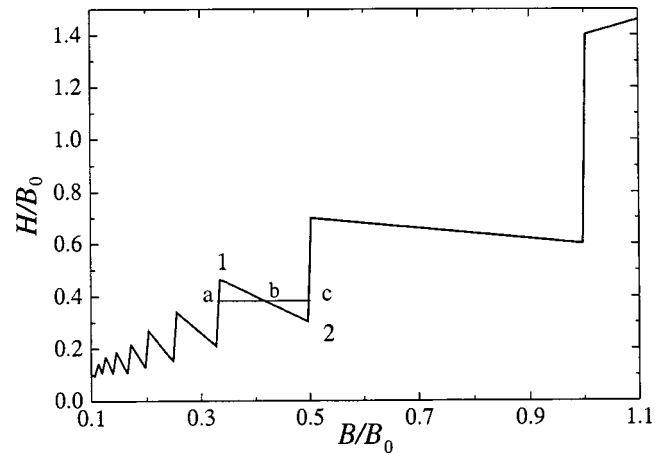FIG. 1. Dependence of the magnetic moment $M$ on the magnetic induction $B$.



FIG. 2. Dependence of the macroscopic magnetic field intensity $H = B - 4\pi M(B)$ on the magnetic induction $B$. The letters a, b, and c represent the values $B_a$, $B_b$, and $B_c$.

$c$ is the speed of light, $h$ is Planck's constant, $\omega_c = eB/mc$, $m$ is the effective mass of the electron, $T$ is the temperature in energy units, and in Eq. (2) we have expressed the ratio $E_F/\hbar\omega_c$ in terms of $B_0/B$ $[E_F/\hbar\omega_c = B_0/B]$.

The quantity $Z_0$ is proportional to the interaction constant of the exchange interaction between the electrons; it is negative and can be close to $-1$ (Ref. 11). In this region the oscillating parts of the magnetic moment and the magnetic susceptibility grow steeply. Such behavior of $M$ and $\chi$ can be explained as follows. According to the Pauli principle, the electrons tend to occupy states with opposite spin. If the interaction between the electrons depends on their spins and the spins are aligned ($Z_0 < 0$), then for an appreciable value of $Z_0$ the decrease in the energy due to the interaction can exceed the increase in the kinetic energy due to redistribution of the electrons over the available states.[11] In this case a new state of the 2D system with nonzero total spin can arise.[11]

Below we consider the case in which $Z_0$ is very close to $-1$ ($Z_0 + 1 \ll 1$). In this case the magnetic susceptibility of the 2D system is large. Correspondingly, the magnetic moment $M$ is also large.

This leads to an important consequence. The amplitude of the oscillations of the magnetic moment $M$ grow dramatically. Figure 1 shows the dependence of the magnetic moment on the magnetic induction $B$, calculated according to formula (2) (in the calculation we chose the value $1 + Z_0 = 2 \times 10^{-4}$). The magnetic moment as a function of $B$ varies discontinuously. Jumps take place at values of the magnetic induction defined by the relation $B = B_0/p$, where $p$ is an integer, $p = 1,2,3, \ldots$ or, equivalently, when the Fermi level intersects a Landau level.[14] In this event, according to Eq. (1), we obtain an oscillating dependence of the magnetic field intensity $H$ on the induction $B$, as shown in Fig. 2. It follows from this figure that there is a region of $H$ values, where each $H$ value corresponds to three different values of $B$ ($B_a$, $B_b$, and $B_c$). Such nonuniqueness testifies to an instability of the state analogous to the instability that occurs for the equation of state of a van-der-Waals material $P(V)$.[11,14,21] In the case under consideration here, as in the case of the van-der-Waals equation, the thermodynamic in-

equality $\partial H/\partial B > 0$, which forbids segments 1–2 in Fig. 2, holds (Ref. 13). Just as the van-der-Waals curve describes a first-order phase transition from a gas to a liquid, the curve $H(B)$ in Fig. 2 describes a sequence of phase transitions with discontinuous variation of the magnetic induction $B$.[21,22] Each time such a point as $a$ is reached (see Fig. 2), the induction varies abruptly from the value $B_a$ to the value $B_c$. In this case the difference $B_c - B_a$ is equal to $B_0/p$, i.e., it is a multiple of the quantity $B_0 = hcn_e/e$, where $p$ is an integer.

Thus, the condition of the induction jumps is the appearance of such segments on the $H(B)$ curve, where $\partial H/\partial B < 0$:

$$\partial H/\partial B = 1 - 4\pi \frac{\partial M(B)}{\partial B} \equiv 1 - 4\pi\chi < 0, \qquad (3)$$

$\chi > 1/4\pi$.

An induction jump takes place in a constant external field $H$ and its occurrence is determined by equality of the free energies in the given field. The equilibrium curve of the magnetic field intensity as a function of the induction corresponds to placing the horizontal line segment $ac$ (Fig. 2) so that the areas of the triangles $a1b$ and $b2c$ are equal[23] (Maxwell's rule). In this case, the curve $B(H)$ consists of plateaus and jumps (Fig. 3). Here the distance between the plateaus is a multiple of the quantity $\tilde{B}_0 = B_0/p \equiv \text{const}$ ($p$ is an integer). Thus, we have found that in a 2D system, as a consequence of the de Haas–van Alfvén effect the magnetic induction $B$ can take only certain discrete values (see Fig. 3)

$$B = B_0/p. \qquad (4)$$

For a given density of 2D electrons $n_e$ the maximum value of the "quantum" of magnetic induction is equal to $B_0$. Note that temperature corrections for the 2D electron density, and consequently for $B_0$, are exponentially small for $T \ll E_F$. Therefore, the exactness of quantization of the magnetic induction is quite high. Thus, as the external magnetic field strength $H$ varies, the 2D system of electrons undergoes a sequence of phase transitions with a discontinuous change in the magnetic induction $B$.
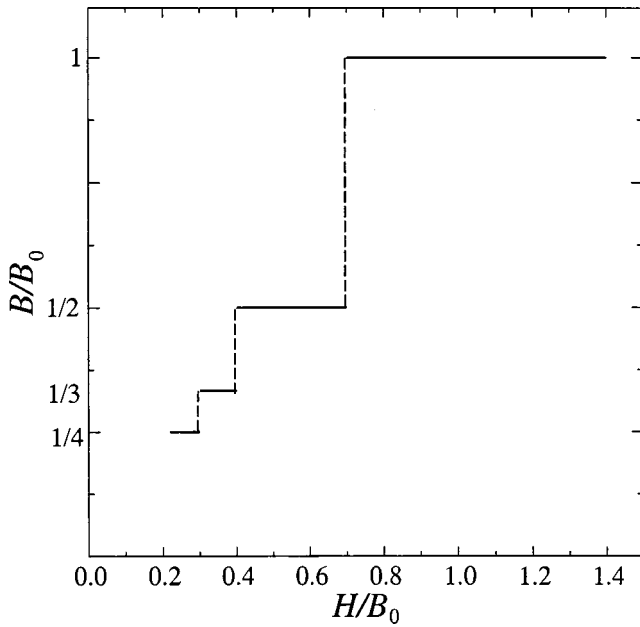
FIG. 3. The dependence $B(H)$, consisting of plateaus ($B_0/p$ = const) and jumps ($B_0 = n_e hc/e$, $p$ = 1,2,3, . . . ).



FIG. 4. Dependence of the Hall resistivity $\rho_{xy}$ on the intensity of the external magnetic field $H$.

It is important to note that the temperature and also the interaction between the electrons have no influence on the period of the oscillations of the magnetic moment.[11,12] This means that the distance between plateaus on the curve $B(H)$ does not vary under these conditions. As the temperature is raised, the amplitude of the oscillations of the magnetic moment varies. The amplitude of the oscillations of $M$ falls, and the width of the plateaus in the dependence $B(H)$ narrows. Thus, we can formulate our main result: in a 2D system, as a result of the de Haas–van Alfvén effect the magnetic induction $B$ as a function of the external magnetic field $H$ has the form of plateaus and jumps (Fig. 3). The distance between plateaus in the dependence $B(H)$ is independent of the temperature and scattering of electrons by phonons and impurities. However, the width of a plateau depends on these factors—with growth of the temperature, the width of the plateau decreases and at some temperature the plateau disappears. On the other hand, the width of the plateaus should increase with increasing impurity concentration.

## 3. INFLUENCE OF THE de Haas–van Alfvén EFFECT ON THE KINETIC COEFFICIENTS OF THE 2D ELECTRONS

The field that figures in the expressions for the kinetic coefficients of two-dimensional electrons $\sigma_{xx}$ and $\sigma_{xy}$ as the acting magnetic field is usually the external magnetic field $H$. The de Haas–van Alfvén effect leads to a large value of the magnetic moment and large magnetic susceptibility of the 2D electrons. In this case, the induction $B$ inside the 2D system differs from the external magnetic field $H$. Consequently, in the expressions for $\sigma_{xx}$ and $\sigma_{xy}$ $H$ should be replaced by $B$.

If we substitute the value of $B$ from Eq. (4) in the expression for the Hall resistivity $\rho_{xy}$, we obtain
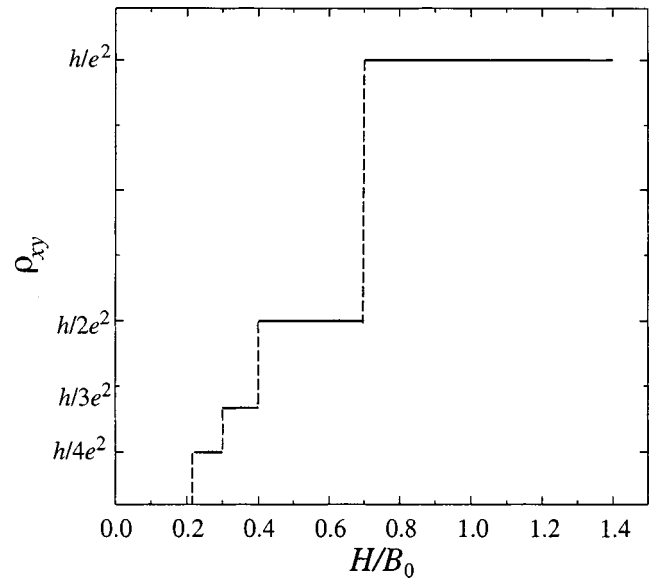
$$\rho_{xy} = \frac{B}{n_e e c} = \frac{h}{e^2 p}. \qquad (5)$$

Thus, quantization of the Hall resistivity is possible because of the quantization of the magnetic induction $B$. Both of these phenomena (quantization of the magnetic induction and quantization of the Hall resistivity) arise as a result of the de Haas–van Alfvén effect.

The dependence of $\rho_{xy}$ on the external magnetic field intensity $H$ has a form analogous to that of $B(H)$ (Fig. 3) and consists of jumps and plateaus (Fig. 4). On the plateaus $\rho_{xy}$ = $h/e^2 p$ = const. The distance between plateaus is a multiple of $h/e^2$. The dependence of $\rho_{xy}$ on the external magnetic field $H$ exactly replicates the dependence of the induction $B(H)$. With growth of the temperature, the width of the plateaus in the dependence of $\rho_{xy}$ on $H$ decreases. At high temperatures the plateaus disappear. The width of the plateaus of $\rho_{xy}(H)$ is also affected by scattering of electrons by phonons and impurities. Since the magnetic induction $B = B_0/p$ is constant on any one of the plateaus of $\rho_{xy}(H)$, the chemical potential $\mu$ is also constant there. Since the chemical potential is found in this case in the region of localized states, we have $\rho_{xx} = 0$. At the moment the induction undergoes a jump, delocalized states arise and $\rho_{xx} \neq 0$. After the jump, $B, \mu$ = const on the plateau and, as before, $\rho_{xx} = 0$.

Thus, quantization of the Hall resistivity $\rho_{xy}$ and the zero value of the ohmic resistivity $\rho_{xx}$ are possibly connected with quantization of the magnetic induction $B$. In this case we obtain integer quantization of the Hall resistivity $\rho_{xy}$.

[1] K. von Klitzig, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

[2] B. I. Halperin, P. A. Lee, and N. Read, Phys. Rev. B **47**, 7312 (1993).

[3] S. L. Sondhi, A. Karlhede, S. A. Kivelson, and E. H. Rezay, Phys. Rev. B **47**, 16,419 (1993).

[4] H. A. Fertig, L. Brey, R. Côté, and A. H. MacDonald, Phys. Rev. B **50**, 11,018 (1994).

[5] K. Moon, H. Mori, Kun Yang *et al.*, Phys. Rev. B **51**, 5138 (1995).

[6] Yu. A. Bychkov, T. Maniv, and I. D. Vagner, Phys. Rev. B **53**, 10,148 (1996).

[7] S. V. Iordanskiĭ and S. G. Plyasunov, Zh. Éksp. Teor. Fiz. **112**, 1899 (1997) [JETP **85**, 1752 (1997)].

[8] S. V. Iordanskiĭ, JETP Lett. **66**, 191 (1997).

[9] W. Apel and Yu. A. Bychkov, Phys. Rev. Lett. **78**, 2188 (1997).

[10] D. Schönberg, Philos. Trans. R. Soc. London, Ser. A **255**, 85 (1962).

[11] A. A. Abrikosov, *Fundamentals of the Theory of Metals* (North-Holland, Amsterdam, 1988).

[12] I. M. Lifshits, M. Ya. Azbel', and M. I. Kaganov, *Electron Theory of Metals* (Consultants Bureau, New York, 1973).

[13] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1960).

[14] R. E. Peierls, Phys. Z. **81**, 186 (1933).

[15] K. Huang, *Statistical Mechanics* (Wiley, New York, 1963).

[16] W. Zawadzki, Solid State Commun. **47**, 317 (1983).

[17] V. P. Silin, Zh. Éksp. Teor. Fiz. **33**, 495 (1957) [Sov. Phys. JETP **6**, 387 (1958)].

[18] Yu. A. Bychkov and L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **41**, 1592 (1961) [Sov. Phys. JETP **14**, 1132 (1962).

[19] L. D. Landau, Zh. Éksp. Teor. Fiz. **30**, 1058 (1956) [Sov. Phys. JETP **3**, 920 (1956).

[20] A. A. Abrikosov and I. M. Khalatnikov, Usp. Fiz. Nauk **66**, 177 (1958).

[21] J. H. Condon, Phys. Rev. **145**, 526 (1966).

[22] A. B. Pippard, Proc. R. Soc. London, Ser. A **272**, 192 (1963).

[23] E. M. Lifshitz and L. P. Pitaevskiĭ, *Statistical Physics*, Pt. 2, (Butterworth Heineman, Oxford, 1995).

Translated by Paul F. Schippnick

## OBITUARY

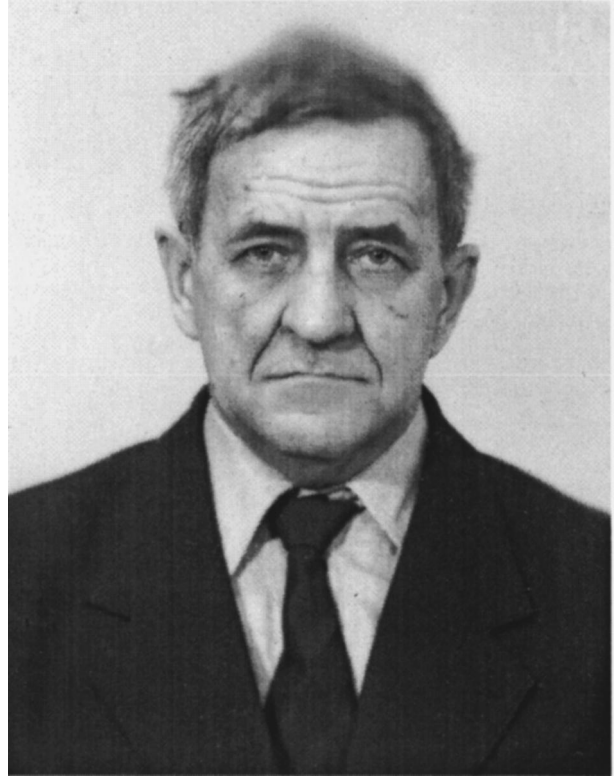# In memory of Aleksandr Aleksandrovich Lebedev

After a protracted and serious illness, Chief Scientist of the A. F. Ioffe Physicotechnical Institute of the Russian Academy of Sciences, Doctor of Physical and Mathematical Sciences Aleksandr Aleksandrovich Lebedev departed this life on March 22, 1999.

The scientific career of A. A. Lebedev was inseparably linked with the physics of semiconductors, and many of his works were of a pioneering nature. Lebedev worked at the A. F. Ioffe Physicotechnical Institute (PTI) for 48 years. Leaving the physics department of Leningrad State University, he arrived at the Physicotechnical Institute in 1951. He initiated the growth of pure germanium single crystals, actively participated in the creation of the first domestic semiconductor devices, and double-injection compensated-silicon diodes were first created at his initiative. The results of his studies of photoelectric phenomena in compensated silicon are well known to specialists and were used in the development of photoelectric devices. His studies of photoelectric phenomena in new ternary semiconductor compounds made it possible to determine or refine the band structure of those compounds and create new polarization-sensitive photodetectors. Lebedev was one of the founders of deep-level transient spectroscopy in semiconductors. The identification of impurities with deep levels controlling the lifetime of charge carriers made it possible to optimize the technology of fabrication of semiconductor devices intended for various purposes.

Lebedev facilitated the growth and development of a large number of young scientists and received the well-deserved acclaim of the world scientific community. Author and co-author of more than three hundred scientific papers and monographs, he was awarded medals ''for valiant labor during the Great Patriotic War'' (Second World War) and in honor of the thirtieth and fiftieth anniversaries of the victory in the Great Patriotic War.

Aleksandr Aleksandrovich Lebedev was always a gallant, sympathetic, and kind colleague and person. His memory as a talented physicist, leader, and friend will always remain in our hearts.

*Colleagues, Friends, and the Editorial Board of the journal Physics and Technology of Semiconductors*

Translated by Paul F. Schippnick

# Eightieth anniversary of the Ioffe Institute

B. B. D'yakov

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

September 23, 1918 is regarded as the birthday of the A. F. Ioffe Physicotechnical Institute. For that is the date of the decree of the People's Commissariat for Education of the Russian Soviet Federated Socialist Republic specifying the creation of the State Roentgenological and Radiological Institute, which included a physicotechnical branch. In 1921 that branch, which had been located in several rooms at the Petrograd Polytechnical Institute, became the State Physicotechnical Roentgenological Institute. It moved into its own small building, which is still referred to as the main building among the numerous many-storied buildings in the Institute, in 1923, when its staff amounted to just a few dozen people. The Institute subsequently retained the ''physical'' and ''technical'' in its name, which was supplemented in 1939 by ''of the Academy of Sciences of the USSR'' (after a change in its departmental affiliation) and in 1960 by the name of its founder, A. F. Ioffe, following his death.

The first staff members of the Physicotechnical Institute included scientists who became leaders in physics. They included future Nobel prize winners: P. L. Kapitsa (1919–1921),[1] N. N. Semenov (1920–1931), L. D. Landau (1927–1932), and I. E. Tamm (1942–1946). The Nobel Prize in chemistry for 1956 was awarded to Semenov for work in chemical kinetics that was mainly done at the Institute.

Ioffe also attracted to the Institute already established scientific authorities, who were the founders of new branches in physics in this country. These include: the electrophysicist A. A. Chernyshev (1918–1931), the crystallographer Yu. V. Vul'f (1918–1925), the acoustician N. N. Andreev (1926–1938), and the thermal physicist M. V. Kirpichev (1928–1930). The next generation of outstanding physicists was represented by Ya. I. Frenkel (1921–1952), G. A. Gamow (1925–1933), and M. P. Bronshtein (1930–1938).

The scientists at the Institute, especially Ioffe, created what, even many years later, could be called the ''cradle of Soviet physics,'' universally known scientific schools and new institutes that have made major contributions to the development of physics and its applications. These include the Ukrainian Physicotechnical Institute (Kharkov), the Institute of Chemical Physics, Laboratory No. 2 (the Institute of Atomic Energy), the Leningrad Institute of Nuclear Physics (Gatchina), and the Semiconductor Institute, organized by Ioffe in 1954 and merged with the Ioffe Institute in 1972.

The following brief review emphasizes an historical perspective on the work that has led to the development of new ideas in physics and to new research areas at the Institute. Other important research at the Institute is mentioned briefly. The range of this research is so broad that some limits on what could be described here were inevitable.

## EARLY RESEARCH

From the first years of the Institute's existence, the work of its staff was marked by fundamental discoveries. By that
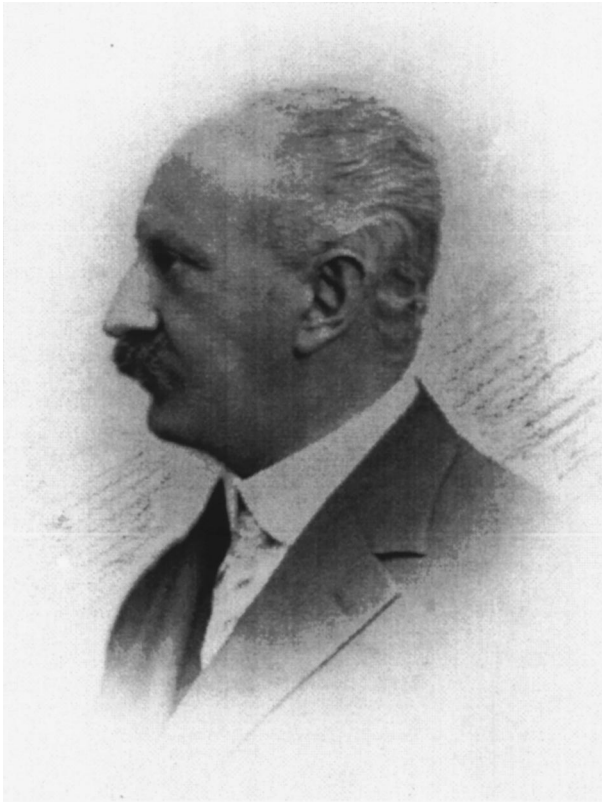
A. F. Ioffe in 1932.

time Ioffe had already formulated his ideas on ionic conductivity in crystals and Frenkel had validated them theoretically by advancing the idea of vacancies (''holes'' in the author's terminology) in a crystal lattice and their role in conductivity. (Later the combination of an interstitial atom and a hole came to be called a ''Frenkel defect.'')

At that time research on magnetic phenomena became a tradition at the Institute. Kapitsa and Semenov proposed measuring the magnetic moment of atoms using the deviation of a beam of the neutral atoms in a magnetic field (independently of Stern and Gerlach). Ya. I. Dorfman's concept of paramagnetic and ferromagnetic resonances, advanced in 1923, was confirmed by his own experiments which demonstrated the nonmagnetic origin of the internal ''molecular'' field in a ferromagnetic material and by Frenkel's theory, which interpreted ferromagnetism as the orientation of the magnetic moment of an electron in a ferromagnetic crystal. Landau continued to study these questions and developed a quantum theory of the diamagnetism of free electrons in metals. Bronshtein and Frenkel's work on the possibility, in principle, of observing the resonance absorption of radiation, whose frequency corresponds to the transition energy of electrons in a magnetic field, should also be mentioned, along with I. K. Kikoin and M. M. Noskov's discovery of the photoelectromagnetic effect.

Nuclear physics research began in this country with the work of D. V. Skobel'tsyn (1924–1937) at the Institute on cosmic rays using a Wilson cloud chamber mounted in a magnetic field (in accordance with an idea of Kapitsa's). Soon the Institute became not only the first center for nuclear physics in this country, but, in the full sense of the word, an organizer of research in this field. The leading Soviet scientists whose efforts subsequently created the country's nuclear shield included the Institute staff members (''fiztekhovtsy'') Yu. B. Khariton (1921–1931), I. V. Kurchatov (1925–1943), I. K. Kikoin (1927–1937), A. P. Aleksandrov (1930–1946), Ya. B. Zeldovich (1931), and L. A. Artsimovich (1933–1946), who came to nuclear physics from other branches of physics. (Kurchatov, for example, was known for his classical papers of 1929–1933 on ferroelectricity.) The first world-class results obtained in this area in our country are also associated with the names of people who worked at the Institute. A list of them speaks for itself: formulating the proton-neutron hypothesis of nuclear structure (D. D. Ivanenko), the discovery of nuclear isomerism (I. V. Kurchatov, B. V. Kurchatov, and L. I. Rusinov, together with L. N. Mysovskii from the Radium Institute), spontaneous fission (G. N. Flerov together with K. A. Petrzhak from the same institute), startup of the first cyclotron in Europe (M. A. Eremeev), and development of the electrocapillary model of the nucleus (Ya. I. Frenkel, independently of Bohr and Wheeler).

As for the technical area, nothing characterizes it better than: the creation of television systems (L. S. Termen and A. P. Konstantinov) and electronic musical instruments (L. S. Termen), inventions in high-voltage technology (A. A. Chernyshev), the creation of the first radar systems in this country (Yu. B. Kobzarev, P. A. Pogorelko, and N. Ya. Chernetsov) and the fundamental work by D. A. Rozhanskii in this area, work on protecting ships against mines (under the leadership of A. P. Aleksandrov with the participation of many leading scientists at the institute) that was begun prior to the Second World War, and, finally, the first semiconductor devices in this country and materials with record parameters for that time (B. T. Kolomiets, Ya. P. Maslakovets, B. V. Kurchatov, and Yu. A. Dunaev).

## ESTABLISHMENT AND DEVELOPMENT OF SCIENTIFIC SCHOOLS

Undoubtedly, the major contribution of scientists from the Institute has been to the development of the physics and technology of semiconductors, beginning with the work of Ioffe on the mechanism for their conductivity. In 1931, suitable scientific subdivisions were set up at the Institute. Since then, this area has been associated with dozens (to recall only the most famous) of names and discoveries made within the Institute's walls. The foundations of the modern science of semiconductors were laid precisely then: the theory of the ''barrier layer'' ($p-n$ junction) based on the tunnel mechanism (Ioffe and Frenkel), the separation of semiconductor conductivity into intrinsic and impurity mechanisms (V. P. Zhuze and B. V. Kurchatov), etc. Studies in another area of semiconductor physics, thermoelectricity, also began at that time. The theory was developed by Ioffe in 1950. In the 1930s, the All-Union conferences on semiconductors (of which there were six before the war) began with the efforts of Institute scientists and immediately attained world status.

In fact, research conducted at the Institute encompassed

all the then-known branches of solid state physics. Certainly, in terms of its significance for physics and for the Physicotechnical Institute, here we should note the concept of the quasiparticle, i.e., the exciton, advanced by Frenkel in 1931, which represented an excited state in a crystal. The significance of the idea was demonstrated many times (first of all, right at the Institute), including an experimental proof of the existence (1951) and the manifold properties of the exciton by E. F. Gross (1944–1972) and his school. The subsequent discovery of the exciton line spectrum and experiments on magnetooptics, electrooptics, piezo-spectroscopy, and exciton dynamics provided the impetus for the development of a new discipline, the optics and spectroscopy of semiconductors. This area is developing successfully at the Institute, even now, through experimental and theoretical work on the optical orientation of electrons, nuclei, and excitons in semiconductors, as evidenced by the discovery of new phenomena in bulk semiconductors and nanostructures. In the 1970s there was the discovery of ordering in the angular momenta of hot electrons by polarized light, a phenomenon which made it possible to measure ultrashort carrier relaxation times in semiconductors after illumination.

One of the leading topics at the Institute has been research on the mechanical durability of solids. Following a realization of the remarkable ideas of Ioffe in attaining the tensile strength of table salt, A. P. Aleksandrov and S. N. Zhurkov (1930–1997) came very close to the theoretical tensile strength of glass and quartz filaments by annealing them. Aleksandrov and P. P. Kobeko (1924–1952) studied polymers (the effect of temperature on polymer deformation, the relaxation states of polymers in terms of a concept of sharp changes in the structure of ''large'' molecules due to thermal fluctuations) and N. N. Davidenkov (1925–1962) studied the brittle fracture of steels at low temperatures and the transition to plastic failure at high temperatures.

In 1931 A. V. Stepanov (1924–1972) began to study defect formation and the development of an electrical potential at the facets of ionic crystals during plastic deformation; this is considered a direct continuation of Ioffe's work. Stepanov advanced the fundamental idea of a double role for plastic deformation in crystalline materials, having established that plastic deformation both hardens crystals and simultaneously facilitates their fracture. Later, these assertions formed the basis of the modern dislocation theory of strength in solids.

The modern scientific literature refers repeatedly to a series of papers by Frenkel and T. A. Kontorova in the 1930s, in which they developed a theory of plastic deformation and slip. These papers became the foundation of the physics of dislocations, but also of the physics of solitons.

In the 1950s and later, Z. N. Zhurkov and his school developed the idea of the kinetic theory of strength in solids, Stepanov's method for growing crystals became widespread, and substantial progress made in understanding the nature of dislocations, plasticity, and the mechanical properties of materials at temperatures ranging from liquid helium to the melting points of these materials.

The first papers of P. I. Lukirskii (1918–1954) and N. N. Semenov on the photoelectric effect and the scattering of electrons in 1923–1926 were the beginning of research on physical electronics, which became one of the major traditional areas of research at the Institute. Lukirskii developed and introduced an original spherical capacitor method for analyzing electron velocity distributions. His work also promoted the development of electron optics (the theory of focussing in electric and magnetic fields) at the Institute and were continued by G. A. Grinberg (1918–1930; 1941–1991) and his school, which began in the 1940s. The work of the founders of physical electronics at the Institute was continued after the war in studies of electron emission and nonequilibrium surface ionization, as well as by the discovery of negative alkali metal ions and the manifold enhancement in the degree of ionization by means of an electric field.

The succession and traditions in scientific subjects are also illustrated by the development of work on mass spectrometry. Mass spectrometric techniques were first developed at the Institute (and in the USSR) by V. N. Kondrat'ev (1923–1931) in a study of the ionization of salt vapor and chemical reactions in 1924. This area of research developed successfully in the 1950s. Mass spectrometers were perfected and they were used to study fundamental problems (precision measurements of physical constants, measuring the helium isotopes in the earth's mantle), as well as practical ones (industrial process monitoring).

The war years did not halt the life of the Institute as a research establishment. Those who remained in besieged Leningrad (under the leadership of P. P. Kobeko) and those who worked in the evacuation used the rich arsenal of results from their earlier research (demagnetization of warships, radar stations, armor plate, thermal electrical generators, cyclotron, etc.), but also obtained new results (an antigangrene preparation, isotope separation, nuclear constants of fissile materials). During the war years, fundamental physics was represented by Frenkel's research that led to his classic monograph *The Kinetic Theory of Liquids*. Also noteworthy are the experiments of Lukirskii on the formation of facets along the surface of a sphere of sodium chloride and Frenkel's 1945 theoretical paper on the formation and dynamics of crystal surfaces, which is very important for modern concepts of semiconductor nanostructures.

## NEW AREAS OF RESEARCH

In the postwar years the Institute strengthened its leading position in many areas relating to armaments, which were distinguished both by their very important concrete applied significance (separation of light isotopes and uranium isotopes, high-temperature coatings for reentry vehicles, semiconductor devices for the first nuclear submarines, etc.) and by the discovery and development of areas of fundamental research that were new to the Institute, but without which it is impossible to imagine today's Physicotechnical Institute. These include controlled thermonuclear fusion and high-temperature plasma diagnostics, astrophysics and space studies, and holography.

Work on electromagnetic isotope separation gave an impetus to research on collisions of atomic particles and their interactions with electrons and protons. These studies are the

main source of information on the structure of atoms and molecules. These results are widely used in plasma physics, astrophysics, gas dynamics, and quantum electronics.

As for astrophysical research, in the 1930s it was represented by the theoretical work of Bronshtein, Landau, and Frenkel, and in the 1950s a broad program of experimental astrophysics research was introduced at the Institute, including studies of near-earth space and the distant cosmos, paleoastrophysics, and gamma-ray astronomy. The focus of theoretical study was such unique objects as neutron stars, quasars, and molecular clouds with a large red shift. It was shown that contrary to some hypotheses, the fundamental physical constants (fine structure constant, ratio of the proton and electron masses) do not change on a cosmological time scale and have the same values in causally unconnected regions of the universe.

Work on thermonuclear fusion began at the initiative of I. V. Kurchatov and B. P. Konstantinov (1958). The primary effort was in the development of particle, microwave, and optical diagnostic techniques for high-temperature plasmas. In the 1970s the first tokamaks were built and yielded encouraging results on the compression and containment of plasma, as well as plasma heating. These studies began to increase in scale.

## SEMICONDUCTOR PHYSICS

The history of progress in the traditional area of research, semiconductor physics, merits special attention, but even the most outstanding work confirming the priority of the Physicotechnical Institute cannot be represented in a single sketch. More detailed information on the work of the Institute in this area (as in the other areas mentioned here) is contained in a collection prepared by the leading Institute scientists of the current generation and published for the jubilee.[2)]

During the early postwar years Ioffe initiated research on the properties of intermetallic compounds. The most important was the discovery by A. R. Regel and N. A. Goryunova of some valuable classes of semiconductors, especially type III–V compounds. Research was begun on their electrical properties. Another important task was to develop methods for purifying, growing, doping, and synthesizing new semiconductor compounds, with 3, 4, or more components. Work with diamond-like semiconductors make it possible to systematize the semiconductor properties of various compounds of this type.

Research closely tied to semiconductor work at the prewar Institute was continued under the leadership of Ioffe at the Semiconductor Institute (IPAN). In 1954–56 the deep penetration of kilovolt ions into semiconductors was discovered. This phenomenon stimulated the development of the implantation method, which has become an efficient instrument for the manufacture of microelectronics components. In the late 1950s and early 1960s work began on the magnetism of nonmetals, and garnet-ferrites with vanadium and bismuth that had an enormous Faraday effect were synthesized for the first time. Magnetic ordering was observed in a number of ferroelectric materials. The nature of weak ferromagnetism

was studied and the transparent ferromagnetic material Rb-$NiF_3$ was synthesized for the first time. Research on ferro- and antiferromagnetic resonances and spin waves was begun in the early 1960s.

A. F. Ioffe initiated research on the physics of disordered systems which led to outstanding results. A wide class of glassy chalcogenide semiconductors was found (N. A. Goryunova, B. T. Kolomiets, 1955), which gave impetus to extensive research on these materials. These semiconductors now have technological applications and have drawn universal attention to the fundamental problem of disordered (amorphous) systems.

Studies of nonequilibrium processes have had a substantial effect on the development of semiconductor physics and technology: the discovery of the convective current instability, the observation of a dependence of the gap width on the concentration of injected carriers, the discovery of stimulated radiation in GaAs $p-n$ structures (1962).

Important results have been obtained from studies of shallow impurity centers in semiconductors. One of the models based on this research describes the energy structure of impurity levels associated with nonequivalent extrema of the bands in III-V semiconductor compounds. Another explains the effect of isovalent doping on the electrical, optical, and mechanical properties of semiconductors and thereby makes it possible to change these properties in a controlled manner without introducing electrically active impurities.

The ''semiconductor revolution,'' under the banner of which science, technology, and society have been developing for 50 years, has been associated with the production of the first semiconductor point-contact diodes in this country (1950–1951), high-purity single crystals of germanium and silicon, junction diodes and $p-n$ junction transistors (1953), III–V semiconductor compounds (1950), power converters (1957), and the discovery of stimulated emission in gallium arsenide (1962).

Work on the creation and study of $p-n$ junctions in Ge began in 1950. After that point-contact diodes and transistors were made in 1950–1951, Institute scientists under the leadership of V. M. Tuchkevich obtained the first single crystals of Ge in the USSR and made the first $p-n$ junction transistors, high-voltage Ge-diodes, and the first electronic circuits using transistors and transistor radios in this country. In 1954 the first planar photodiodes of an original design were produced in the USSR with better sensitivity and active area than the existing ones.

In that same year, the first Soviet high-power devices were produced: germanium semiconductor rectifiers for currents of up to 500 A with a reverse voltage of up to 150–200 V. After this success with germanium devices, studies of single-crystal silicon, the other leading material of the semiconductor revolution, were undertaken for solar energy conversion and high-voltage diodes and thyristors.

After semiconductor devices based on germanium and silicon had been created and successfully brought into industrial production, further study and searches for new materials revealed that gallium arsenide is one of the best materials for creating devices with exceptional parameters at high temperatures.

The Scientific-Education Center of the Institute.

Controlling the most important parameters of semiconductor crystals, i.e., using heterostructures in semiconductor physics and semiconductor electronics, was then addressed. Under the leadership of Zh. I. Alferov, this topic has become one of the major areas of research at today's Physicotechnical Institute.

In the 1960s the concept of a laser with a two-sided heterostructure was formulated; this made it possible to attain a high density of injected carriers and a population inversion as a result of "two-sided" injection. An "ideal" heterostructure, a self-consistent lattice for GaAs in the form of solid solutions of AlGaAs, was observed; in it predicted effects (such as carrier superinjection, an optical and electron clipping) could take place. The unique injection properties of wide-gap emitters and the phenomenon of superinjection were confirmed experimentally. Besides luminescence and carrier diffusion in a smooth heterojunction, stimulated emission was observed in double AlGaAs structures.

The search for new heterostructures with matched lattices has opened up the possibility of obtaining various kinds of heterojunctions using quaternary III–V solid solutions, which permits independent variation of the lattice constant and gap width. InGaAsP compounds have been studied in detail and successfully used in a large number of applications, including photocathodes and lasers for fiber optic communication in the infrared and visible. It is generally recognized that further progress in this area has relied on the efforts of Institute scientists: lasers have been built with distributed feedback provided by a diffraction grating on the surface of a waveguide layer (not only to provide distributed feedback, but also for extraction of well-collimated light). Research on heterostructures has led to the discovery of fundamental physical phenomena, a sharp improvement in the characteristics of essentially all the known semiconductor devices, and the creation of completely new types of devices.

One of the most modern devices based on nanostructures is the quantum dot laser. Quantum dot structures offer promise for other devices as well, especially for creating new types of transistors. The first work in this area has already been done at the Physicotechnical Institute.

## THE INSTITUTE AT THE THRESHOLD OF THE NEW CENTURY

The structure of the Physicotechnical Institute reflects the priorities of the 20th century. A. F. Ioffe, himself, tried to encompass everything new and the new areas which developed in basic and applied physics. As noted above, this effort was later embodied in the creation of the Nuclear Research Institute at Gatchina and new areas of research at the Institute, including astrophysics, high-temperature plasma physics, physical gas dynamics, and holography. Thus, research by Institute scientists encompasses essentially the entire range of contemporary physics and continues as an example of an optimum combination of basic research with concrete technical applications.

These studies have been judged worthy by the scientific community and the government. Of the Institute staff, more than one hundred have been made academicians and corresponding members of the USSR Academy of Sciences (Russian Academy of Sciences) and the Academies of Sciences of the former republics of the Soviet Union. About 150 scientists have been awarded prestigious national and academic prizes. Twenty nine Institute scientists were winners of the Lenin Prize,[3] beginning with A. A. Chernyshev for his work on high-voltage and high-frequency technology (1930). At present, five academicians and ten corresponding members of the Russian Academy of Sciences and more than 800 doctors and candidates of science work at the Institute. Of the leading scientists at the Institute, special mention should be made of five, who, along with their scientific activity, have had enormous responsibility for the leadership of the Institute, contributing to its development as an institution and as a scientific collective: A. F. Ioffe headed the Institute from 1918 to 1952; he was replaced by A. P. Komar (1930–1936 and 1950–1972), who held the post of director until 1957;

B. P. Konstantinov (1926–1930 and 1940–1969) headed the Institute for ten years, from 1957 to 1967, and passed his post on to V. M. Tuchkevich (1935–1997), who led the Institute for the next 20 years; and Zh. I. Alferov, the current director who took over the Physicotechnical Institute in 1987.

The Institute participates in major international projects in plasma physics, solid state physics, and astrophysics, and maintains ties with the leading research centers of the world. This research is supported by additional funds from a number of national and international foundations, including some specially created for the support of joint research by scientists from the former USSR and the European Community (INTAS) and the United States (CRDF), as well as the International Scientific-Technical Center supported by the United States government and the Japanese government.

Certainly, the most indisputable achievement of the Physicotechnical Institute has been the creation of a national school of physics that has not been locked in one research area, regardless of how productive, or in one institute, but has reached the level of the world scientific community. An important part of this process has been a system of preparing scientific personnel whose foundation was laid by Ioffe in 1919. In that year, Ioffe organized the physical-mechanical faculty of the Polytechnical Institute, whose first graduates quickly became staff members of the Institute. (They included Yu. B. Khariton and G. A. Grinberg.) At present, the system of continuous physics training based at the Physicotechnical Institute includes a physicotechnical secondary school, the chair of optoelectronics at the Electrotechnical University, and the physicotechnical faculty at St. Petersburg Technical University.

It is the purpose of the newly founded Scientific-Education Center at the Institute to ensure the unity of these educational degrees. One of the stages in constructing the special building of the Center was completed in September, 1998, on the eve of the eightieth anniversary of the Physi-

cotechnical Institute. The plenary meeting of the International Conference on Physics on the Verge of the 21st Century was held there, with the most famous Russian and foreign physicists participating, including recent Nobel laureates. The conference program encompassed all areas of scientific activity at the Institute and more than 200 scientific papers and poster presentations were given, distributed over plenary sessions and three topical symposia: plasma physics and astrophysics; solid state physics; and semiconductor heterostructures. Some of the papers presented in the last symposium have been published in this issue of this journal.

Continuing the tradition begun with the famous prewar conferences on nuclear physics and semiconductor physics, the Institute has many times been the organizer of prestigious national and international conferences, symposia, and schools, and has also initiated a number of scientific activities which have become standard, such as the International Symposium on Nanostructures: Physics and Technology and the International Conference on Fullerenes and Atomic Clusters, which have taken place since 1993. This brief history of the Physicotechnical Institute gives only a general idea of the origins and vital activity of a complex and continuously developing organism of the sort that a modern scientific research center has become. The brilliant achievements and traditions of the past, and the broad range of modern research offers hope that the future of physics will continue to be tied to the A. F. Ioffe Physicotechnical Institute.

[1]Here and in the following, the dates in parentheses indicate the years a person worked at the Institute.
[2]Ioffe Institute: 1918–1998. Development and Research Activities, St. Petersburg (1998).
[3]Of these, 28 were for work at the Physicotechnical Institute.

# Silicon-germanium—a promise into the future?[1]

H. G. Grimmeiss*

*Lund University, Solid State Physics, Box 118, S-22100 Lund*
Fiz. Tekh. Poluprovodn. **33**, 1032–1034 (September 1999)

Comparing different semiconductor technologies and taking into account their actual performance/ cost ratio, the current understanding is that the silicon technology is and will be the basis of IC fabrication also in the beginning of the next century. Due to new applications there is, however, a market's demand to overcome the cutoff frequency limit inherent to silicon technology. Hitherto obtained results suggest that existing solutions may be challenged by SiGe or other silicon-based heterostructure devices. As an example of such heterostructure devices, in general, the first practical bandgap-engineered silicon device, namely, the SiGe heterojunction bipolar transistor, is discussed in more detail. However, to succeed commercially, the SiGe technology has to outperform the silicon technology and undersell the III–V technologies. To this end, the development of a heterojunction silicon-based CMOS technology will be of utmost importance. First results are encouraging and, hence, the very near future will probably show whether or not SiGe is a promise into the future. © *1999 American Institute of Physics.* [S1063-7826(99)00209-4]

Information technology is still one of the most important global technologies. Ever since the invention of the transistor, we have witnessed a fantastic growth in semiconductor electronics, an important area within information technology. It is fair to say that this impressive progress has only been possible due to a unique combination of basic conceptional advances, the perfection of new materials and the development of new device principles.

The first transistors were fashioned not from silicon but from germanium. There were important reasons for this priority, among them the much greater ease of purifying germanium and the higher mobility of electrons and holes within it. But due to the smaller energy bandgap, the performance of germanium transistors degrades rapidly with temperature. In addition, no material was found that would provide diffusion masking for germanium.

Recognizing these problems, the role of germanium as a major player ended in the mid-1950s and semiconductor technology switched to silicon. Since then silicon has been the dominant material in semiconductor technology. Only a few years later, in 1958, the integrated circuit was invented providing the base for the fantastic growth of microelectronics.

Microelectronics means the multifunctional integration of electronic devices and since the beginning of microelectronics, the objectives have been to further increase the packing density, the complexity, the signal processing rate, and the multifunctionality.

Various attempts have been made to accomplish these goals. Restricting ourselves to technology, three approaches are of particular interest to be mentioned briefly:

1) Three dimensional, stacked device structures would allow circuits of particularly high complexity and multifunctionality. Though first results are encouraging, this approach will acquire further interest in the future to overcome crippling flaws.

2) The continuous reduction of the structure size within silicon technology over the past 30 years or so has been extremely successful in improving the performance of microelectronics and is in large measure responsible for the breakthrough of information technology. Yet from a circuit designer's perspective, silicon is hardly the perfect material. Compared with some of the other semiconductors, it is quite poor in terms of how fast charge carriers can move through the crystal lattice. This sluggishness limits the speed at which all-silicon devices can operate.

3) In contrast to silicon, many III–V semiconductor compounds are known to exhibit high carrier mobilities and saturation velocities. Almost endless variations on their chemical composition and, hence, energy bandgap are possible. Bandgap engineering, in fact, is a powerful tool for growing different kinds of heterostructures and, hence, creating faster transistors. Even the most developed all-silicon technology will therefore never be able to compete with III–V semiconductor compounds at least not with respect to signal processing rates and light-emitting properties.

Due to their performance, compound III–V semiconductors should therefore without any doubt provide the best possibilities for accomplishing the objectives of mainstream microelectronics.

However, in spite of these performance advantages, why are more than 95% of all manufactured ICs fabricated in silicon? The answer is economics. ICs are simply much more difficult and expensive to fabricate from III–V compounds than from silicon. Cost estimates performed by different companies show (Ref. 1) that the cost/area increases by a factor of about 5 when going from silicon to GaAs and by more than a factor of 10 when going from 6-inch silicon wafers to 3-inch InP wafers. These figures are still more in favor of silicon considering that the 300-mm silicon wafer is on the market and that the 450-mm wafer (probably together with a 70-nm technology) is to be expected within the next

10 years (Ref. 2). World-wide investments in silicon technology is enormous and yields are high in the silicon world (CMOS).

On the other hand, high-quality oxides are scarce in the III–V world, impeding device integration and high-purity, large-diameter crystals are difficult to grow. In addition, chip yields are lower due to inherent defects implying higher production costs, as mentioned earlier. Since market shares are in most cases determined by performance/cost ratios, most companies are convinced that semiconductor technology is most probably predictable through Moore's law another 10 to 15 years (Ref. 3) and that silicon-based ICs will dominate the market also in the beginning of the next century.

However, as already pointed out, silicon is not perfect. Due to material properties, there is a cutoff frequency limit inherent to silicon technology and considerably lower than with III–V semiconductors. Due to new applications in information technology asking for still faster circuits, there is a market's demand to overcome the cutoff frequency limit of the silicon technology since these applications will only generate a large market if low-cost systems can be offered. Anticipating that mobicom applications up to about 3–5 GHz will be covered by all-silicon technologies, there are numerous other wireless applications such as wireless local area network (WLAN), hyperLAN and broadband satellite communication which are of interest. In-home digital networks coupled to public networks are another example. All these applications need frequencies up to about 15 GHz. If fiber communication or some radar applications are included, circuit designers like to have frequencies up to about 40 GHz available. These frequencies are difficult to obtain by mainstream silicon technologies.

One way of overcoming the cutoff frequency limit of silicon devices is using the performance advantage of III–V compound semiconductors based on heterostructures. In this context, silicon-germanium alloys have been discussed since the mid-1950s (Ref. 4). Well aware of the fact that III–V compound semiconductors already have shown their superior performance for such applications, the vision is to outperform silicon technology by using the performance advantage of III–V compounds, and to undersell the III–V technologies.

One of the first and probably also one of the simplest bandgap-engineered silicon devices is the SiGe heterojunction bipolar transistor (HBT). The HBT may therefore serve as an example for demonstrating a few important differences in the performance of all-silicon devices with regard to silicon-based heterojunction devices in general.

The *npn* HBT, for instance, differs from a silicon bipolar junction transistor (BJT) only with regard to the base which in the HBT consists of a strained $Si_{1-x}Ge_x$ epitaxial layer of either constant or varying composition (Fig. 1). The only constraint for a practical SiGe HBT design is that the integrated germanium content be low enough to guarantee film stability. Since the difference in the bandgap energy $\Delta E_g$ is almost completely absorbed by a lowering of the conduction band edge, electrons have to overcome a much lower energy barrier in the HBT than in a BJT when being injected into the base (Fig. 2). This implies, somewhat simplified, that the
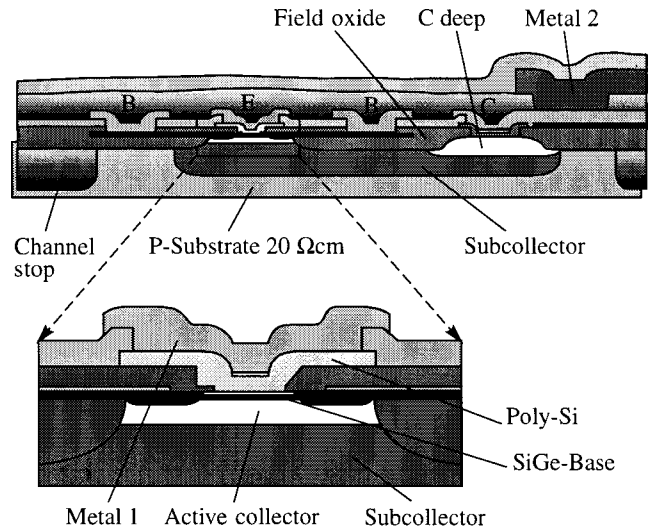


FIG. 1. Schematic diagram of a SiGe hetero-bipolar transistor (Courtesy J. Arndt, Temic).

(common-emitter) current gain $\beta$ can be approximated by the expression $\beta=[(D_nN_EW_E)/(D_pN_BW_B)]Z$, where $Z=[(N_cN_v)_B/(N_cN_v)_E]\exp(\Delta E_g/kT)$. Here $D_{n,p}$ is the diffusion constant of electrons and holes, respectively; $N_{E,B}$ is the emitter and base doping, respectively; and $W_{E,B}$ is the emitter and base layer thickness, respectively. The other symbols have their usual meaning.

In most transistor designs, $D_n/D_p$ is close to 1 and $W_E/W_B$ between 3 and 5. To achieve a current gain of about 150 in a silicon BJT, the $N_E/N_B$ ratio must be chosen between 20 and 30 since $Z=1$. Since the emitter doping can be increased above $10^{20}$ cm$^{-3}$ only slightly, this restriction puts a limit on the base doping and, hence, on the base resistance $R_B$.

The situation is very different with the HBT. There, the current gain depends exponentially on the germanium content at the base-collector junction, implying that $Z$ is about 100 for $x=0.2$ and about 400 for $x=0.25$. For many applications, current gain is less important than speed. Consider-
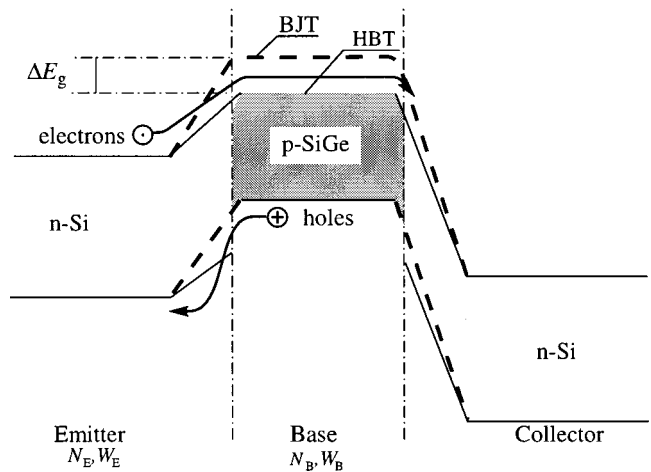


FIG. 2. Schematic energy diagram of a SiGe *npn* hetero-bipolar transistor (Courtesy J. Arndt, Temic).

ing that the maximum oscillator frequency at which power amplification is 1, can be approximated by $f_{max}=0.2\sqrt{f_T(R_B C_{BC})}$, a larger current gain can be traded against a higher $f_T$ and a lower base resistance by reducing the base layer thickness and simultaneously increasing the base doping. Here $C_{BC}$ is the collector-base capacitance, and $f_T$ is the frequency at which the magnitude of the incremental short-circuit common-emitter current gain extrapolates to unity. The quantity $f_T$ is given by $f_T=1/2\pi\tau_{EC}$ and, hence, increases with decreasing base layer thickness, since $\tau_{EC}$ is the transit time of the carrier from the emitter to the collector.

Recent studies at different laboratories have demonstrated that the upper bound on cutoff frequency in aggressively designed SiGe HBTs is well above 100 GHz, more than a factor of 2 higher than the best achieved in silicon. The frequencies $f_T=116$ GHz, $f_{max}=165$ GHz and noise figures of 0.5 dB at 2 GHz, as well as power added efficiencies larger than 70% have been reported and there are good reasons to believe that there is still more to come. Simulations show that cutoff frequencies close to 200 GHz may well be achieved in SiGe HBTs, suggesting that visions in the past may still become reality. Assuming that the operating frequency is about $0.1f_T$, the realization of such circuits would indeed open a wide range of new applications for silicon-based technologies.

Though the SiGe HBT is a very interesting and attractive device, bipolar devices and circuits comprise only a small fraction of the total market. Realizing that CMOS (together with BiCMOS) dominates about 80% of the market, SiGe technology will probably never merge with mainstream technology if it does not contribute to CMOS technology. Although in a first approach, SiGe HBTs will probably be combined with standard CMOS technology, several attempts have been made to use the performance advantage of silicon-based heterostructures also for CMOS applications. Due to the limited scope of this paper, only two examples will be mentioned.

One approach is based on the vertical MOS which, for instance, has been demonstrated by Siemens as an all-silicon ROM in MultiMediaCards (1997: 64 Mb, next generation: 256 Mb). Although the channel length is defined by the epitaxial channel layer, these devices still use a channel length of 200 nm. The advantage of the vertical MOS over a conventional planar MOS is the 4 times smaller area of $2F^2$ compared with $8F^2$ for a DRAM and a 4-$\mu$s read access time of the Multimedia Card application.

A heterojunction at the source side of a vertical MOS keeps the barrier constant in the off-state and is expected to avoid, among other things, short channel effects such as drain-induced barrier lowering. Simulations suggest that vertical hetero-MOS with channel length down to about 20 nm may exhibit on-currents of up to 20 000 $\mu$A/$\mu$m$^2$, off-state currents smaller than $10^{-3}$ $\mu$A/$\mu$m$^2$, a transconductance larger than 3500 mS/mm, and an intrinsic carrier transit time less than 1 ps (Ref. 5).

Another approach of heterojunction field effect transistors (HFET) has been demonstrated by several laboratories such as the Daimler-Benz Research Laboratories, Ulm, Ger-
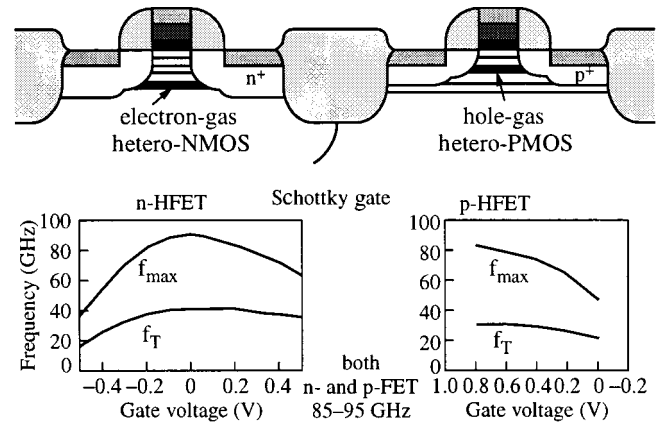


FIG. 3. Schematic diagram of a lateral hetero-CMOS and examples for the frequency response of the $n$-FET and $p$-FET (Courtesy U. Konig).

many and IBM. The Daimler-Benz design is based on the high-mobility performance of strained silicon and germanium channels in otherwise conventional planar layer structures (Fig. 3). Both $n$-HFET and $p$-HFET with $f_{max}$ and cutoff frequencies above 85 GHz and 40 GHz, respectively, have already been demonstrated. Similar data have been obtained by IBM. Interestingly, due to the large mobility increase of holes in the germanium 2DHG, both types of devices exhibit similar frequency performances, in contrast with conventional silicon MOS. Simulations suggest that $f_{max}$ of SiGe HFETs may well approach values above 200 GHz for channel length below 100 nm.

We will soon find out whether or not hetero-CMOS will overcome the cutoff frequency limit of silicon or whether they are challenged by other approaches such as gigascale single-electron devices (Ref. 6). Nevertheless, SiGe (or perhaps even SiGeC) alloys suggest many interesting possibilities for the future. Several companies offer already SiGe circuits on the market with considerable performance advantages over conventional devices. The very near future will show whether or not SiGe has succeeded in outperforming silicon technology and underselling the III–V technologies.

[1] U. Konig, in *Proceedings of Nobel Symposium 99, Heterostructures in Semiconductors* (Arild, Sweden 4–8 June 1996), edited by H. G. Grimmeiss (World Scientific, Singapore, 1997), p. 90.
[2] A. M. Wacker (private communication).
[3] A. W. Wieder, Research and Innovation, I/98, p. 39, published by Siemens AG, Corporate Technology Dept., Munich.
[4] H. Kroemer, Proc. IRE **45**, 273 (1957).
[5] K. De Meyer (private communication).
[6] K. Yano *et al.*, ISSCC 98, p. SA 21.7:A.

## SEMICONDUCTOR HETEROSTRUCTURES

## Heterostructure solar cells

V. M. Andreev*)

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1035–1038 (September 1999)

Heterostructure solar cells based on III-V compounds are studied. Record-high efficiencies are obtained for solar cells based on AlGaAs/GaAs heterostructures: 24.6% for 100-fold concentration of sunlight in outer space (AM0) and 27.5% for 100-fold concentration of the light on the ground (AM1.5). A substantial increase in radiation resistance is obtained for solar cells with a built-in Bragg mirror. Cascaded solar cells with efficiencies of up to 32% for 100 suns (AM1.5) are created and studied; in these cells the upper wide-gap materials are infrared transparent elements based on GaAs, while the lower narrow-band elements are made of GaSb or an InGaAs solid solution. © *1999 American Institute of Physics.* [S1063-7826(99)00309-9]

Photoelectric conversion is one of the most promising methods for converting solar energy. Since the launching of the satellite Sputnik-3, on which solar batteries were used for the first time, the photoelectric method has been the major method for obtaining electrical energy on spacecraft and is in ever wider use on earth. Silicon is currently the main material in solar cells. The efficiency of silicon-based solar cells is 15–16% under direct solar irradiation near the earth. The efficiency of these cells is about 20% on the earth under direct solar irradiation and up to 25–27% at 30 to 50-fold concentration of the energy of the sunlight. When the illumination intensity is increased further, the efficiency of silicon cells decreases because of the elevated operating temperature and higher resistive losses.

Solar cells based on heterostructures have high efficiencies and high radiation resistance.

The important advantage of heterostructure photoconverters is their ability to efficiently convert highly concentrated sunlight (up to 1000–2000 times), which opens up the possibility of a substantial (proportional to the degree of concentration) reduction in the area and cost of solar cells and, thereby, a reduction in the cost of ''solar'' electric power.

Solar cells based on $n$-GaAs-$p$-AlGaAs heterostructures were first proposed and created at the Ioffe Institute.[1,2] The use of a wide-gap ''window,'' made of a thin layer of AlGaAs solid solution (Fig. 1a) that is essentially completely transparent to sunlight, ensures passivation of the surface of the photoactive region[3,4] and makes it possible to achieve efficiencies close to the theoretical maxima. Low-temperature liquid-phase epitaxy was used to create[5,6] multilayer AlGaAs/GaAs heterostructures that yielded record-high efficiencies for solar cells with a single $p-n$-junction: 24.6% for 100 ''suns'' under the conditions of outer space and 27.5% for 100–200 ''suns'' on earth. These efficiencies were achieved by reducing the thickness of the frontal layer of $Al_{0.9}Ga_{0.1}As$ to 30–50 nm, crystallizing high-quality material in the active region, and creating a

back potential barrier, made of $Al_{0.1}Ga_{0.9}As$ which increased the efficiency of collecting the charge carriers generated by the light (Figs. 1b and 2).

In recent years, MOC-hydride epitaxy (gaseous phase epitaxy from metal-organic compounds) has come into widespread use for fabricating AlGaAs/GaAs heterostructures.[7–11] Solar cell heterostructures with a built-in Bragg mirror (Fig. 1c) made by this method are of considerable interest. In the structure developed at the Ioffe Institute,[9,11] instead of a back potential barrier, a multilayer dielectric mirror which consists of 24 alternating layers of GaAs (60 nm) and AlAs (70 nm) was fabricated. The reflectivity from this sort of mirror is $\sim 95\%$ in the spectral interval $750-900$ nm. This reflects the part of the sunlight that is not absorbed in the base layer into the active region and makes it possible to maintain a high efficiency of collection of the current carriers generated by the light while reducing the thickness of the base layer and reducing the diffusion displacement lengths. A consequence of its use is a higher radiation resistance for the solar cells.

Because of their high efficiency and elevated radiation resistance, solar cells based on AlGaAs/GaAs heterostructures are widely used in solar batteries in outer space. The solar battery of the base module of the Mir space station is equipped with heterostructure solar cell with a total area of about 60 m$^2$ that was manufactured by the technology developed at the Ioffe Institute. Over its more than ten years of operation, the power of this battery has fall by only 20%.

Cascade solar cells[12–16] based on multilayer heterostructures with two or more $p-n$ junctions in materials with different band gaps can further enhance the efficiency (Fig. 3). In these solar cells, the ''upper'' $p-n$-junction, made of the wider-band material, is intended for efficient conversion of the short wavelength part of the sunlight, while the ''lower'' $p-n$-junction, made of the narrow-gap material, is optimized for conversion of long wavelength light that has passed through the wide-gap element. Theoretical estimates
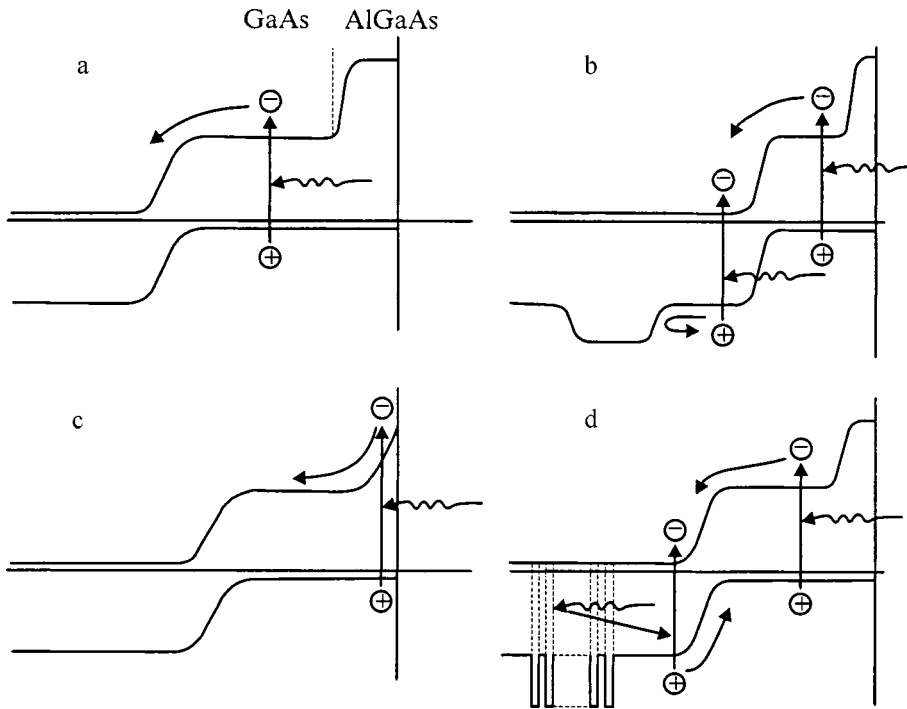
FIG. 1. Energy band diagrams of solar cells based on AlGaAs/GaAs heterostructures: a — structure with a $p-n$-junction in GaAs and with a frontal wide-gap ''window'' of $p$-AlGaAs; b — structure with a back potential barrier in the $n$-region; c — structure with a frontal $p$-AlGaAs layer of variable composition; d — structure with a Bragg mirror in the $n$-region.

show that in complicated photoconverters of this sort, efficiencies of greater than 40% are possible.

A substantial enhancement in the efficiency of cascade cells has been demonstrated[12,13] in a design with mechanical fastening (Figs. 3c and 3d). GaAs is used for the wide-gap element in this design and either InGaAs or GaSb are used as the material for the narrow-gap element. Here the GaAs element is transparent for infrared radiation with wavelengths exceeding 0.9 $\mu$m, while the narrow-gap element ensures efficient conversion of the long-wavelength portion of the sunlight within the spectral range 0.9-1.8 $\mu$m. Solar cells of

this sort have efficiencies of 32-33% (Ref. 13) for 50 to 100-fold concentration of solar radiation when the air mass is present (AM1.5).

A monolithic design for a cascade solar element (Figs. 3a and 3b), in which the ''upper'' and ''lower'' elements are connected by a $p^+-n^+$-tunnel junction, which ensures low



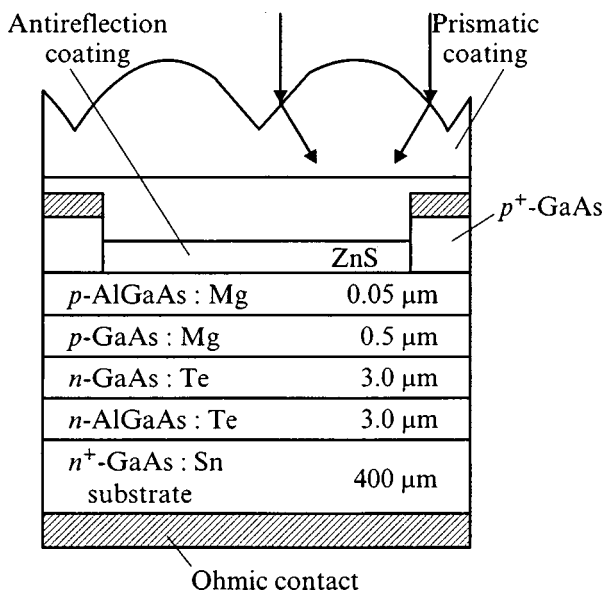| | |
|---|---|
| $p$-AlGaAs : Mg | 0.05 $\mu$m |
| $p$-GaAs : Mg | 0.5 $\mu$m |
| $n$-GaAs : Te | 3.0 $\mu$m |
| $n$-AlGaAs : Te | 3.0 $\mu$m |
| $n^+$-GaAs : Sn substrate | 400 $\mu$m |

FIG. 2. Sketch of a heterostructure solar cell with a back potential barrier of $n$-AlGaAs:Te and a prismatic front coating used to reduce the optical losses due to shading by the front strip contacts.
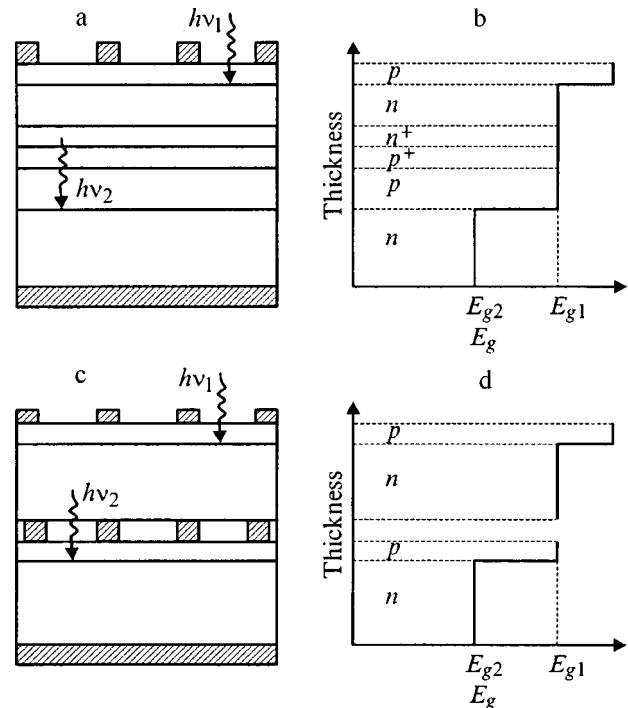


FIG. 3. Schematic representation (left) and distribution of the width of the band gap $E_g$ (right) for two types of cascade solar cells: a, b — monolithic construction with a $p^+-n^+$-junction; c, d — cascaded cells with mechanical joint connected by a contact grid.
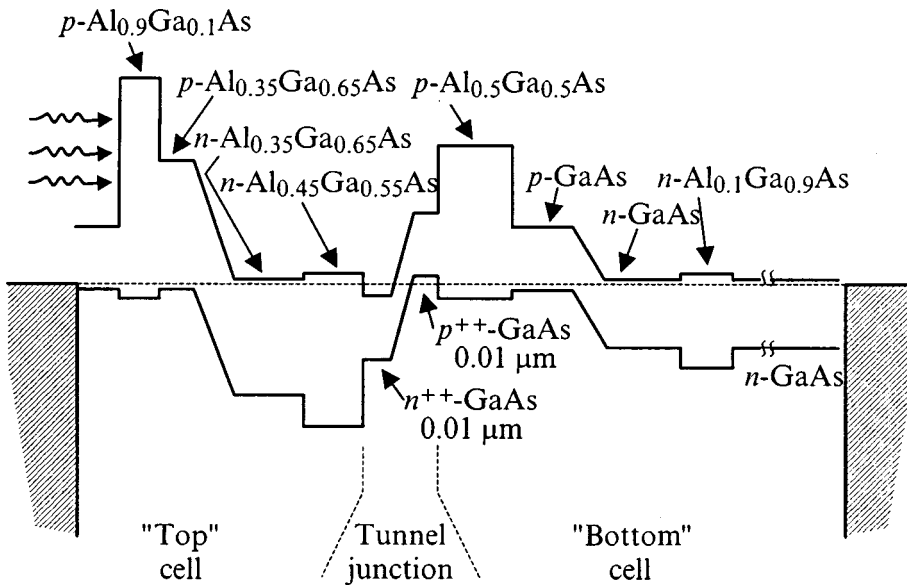
FIG. 4. Band energy diagram of a monolithic cascade solar cell based on an AlGaAs/GaAs heterostructure fabricated by low-temperature liquid-phase epitaxy.

resistive losses, is of even greater interest. Cascade cells of this sort (Figs. 4 and 5) have been created on the basis of AlGaInP/GaAs[14,15] and AlGaAs/GaAs[16] heterojunctions by MOC-hydride epitaxy[14,15] and liquid-phase epitaxy.[16] The maximum attainable efficiencies for these cells are about 30% (Ref. 15) for solar radiation passing through the atmosphere (AM1.5). When MOC-hydride epitaxy is used for this purpose, it becomes possible to use substrates of relatively cheap and more durable germanium for fabricating multilayer III-V heterostructures (Fig. 5), which opens up the possibility of reducing the expense and specific power output of solar batteries. In addition, as Fig. 5b shows, an additional (third) $p-n$-junction can be created in the germanium in these structures, thereby raising the efficiency in a three-cascade photoconverter of this type. Therefore, record-high efficiencies have been obtained in heterostructure solar cells

and there are possibilities for further enhancing their efficiency and radiation resistance, as well as reducing their cost.

Power generators with sunlight concentration based on these photocells have recently been developed at the Ioffe Institute for use on the earth and in space.[17,18] The ground-based power systems were based on high-power (5–10 kW), high-efficiency photocells and concentrators (mirrors and Fresnel lenses) that concentrate sunlight by up to 500 times. These devices are equipped with systems for tracking the sun's position and are completely automatic, which makes them a promising source of electrical energy for consumers who are far from central electricity supply lines.

The concentrator battery modules for use in space contain a lens panel, a honeycomb frame, and stamped plates with solar cells mounted at the focus of the Fresnel lenses. In addition to their higher efficiency, using these solar batteries in space offers the following additional advantages: the possibility of protecting the photoconverters from ionizing radiation using the components of the concentrator system, the possibility of selecting a thermal operating regime for the photoconverters that ensures thermal annealing of radiation defects, and the possibility of performing photon and injection annealing of solar cells operating at elevated photocurrent densities.
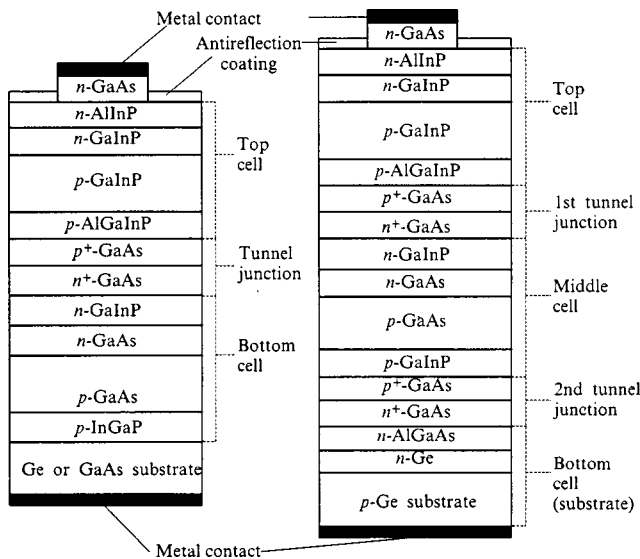
*)E-mail: andreev@scell.ioffe.rssi.ru



FIG. 5. Structures of monolithic cascade solar cells based on AlGaInP/GaAs/Ge heterostructures with two (left) and three (right) photoactive $p-n$ junctions.

[1] Zh. I. Alfërov, V. M. Andreev, N. S. Zimogorova, and D. N. Tret'yakov, Fiz. Tekhn. Poluprovodn. **3**, 1633 (1969) [Sov. Phys. Semicond. **3**, 974 (1969)].

[2] Zh. I. Alfërov, V. M. Andreev, M. B. Kagan, I. I. Protasov, and V. G. Trofim, Fiz. Tekhn. Poluprovodn. **4**, 2378 (1970) [Sov. Phys. Semicond. **4**, 1535 (1970)].

[3] V. M. Andreev, T. M. Golovner, M. B. Kagan, N. S. Koroleva, T. L. Lyubashevskaya, T. A. Nuller, and D. N. Tret'yakov, Fiz. Tekhn. Poluprovodn. **7**, 2289 (1973) [Sov. Phys. Semicond. **7**, 1365 (1973)].

[4] Zh. I. Alfërov, V. M. Andreev, D. Z. Gaburov, V. R. Larionov, V. D. Rumyantsev, and V. B. Khalfin, Fiz. Tekhn. Poluprovodn. **11**, 1765 (1977) [Sov. Phys. Semicond. **11**, 1014 (1977)].

[5] V. M. Andreev, A. B. Kazantsev, V. P. Khvostikov, E. V. Paleeva, S. V. Sorokina, and M. Z. Shvarts, in *Conference Record of the 1st World Conference Photovoltaic Energy Conversion*, Hawaii, USA (1994), p. 2096.

[6] V. M. Andreev, V. P. Khvostikov, E. V. Paleeva, and M. Z. Shvarts, in *Proceedings of the Symposium Compound in Semiconductors*, St. Petersburg (1996), p. 449.

[7] S. P. Tobin, S. M. Vernon, S. J. Wojtczuk, C. Bajgar, M. M. Sanfacon, and T. M. Dixon, in *Proceedings 21st IEEE Photovoltaic Specialists Conference*, N.Y. (1990), p. 158.

[8] S. P. Tobin, S. M. Vernon, M. M. Sanfacon, and A. Mastrovito, in *Conference Record 22nd IEEE Pohotovoltaic Specialists Conference*, Las Vegas, USA (1991), p. 147.

[9] V. M. Andreev, V. V. Komin, I. V. Kochnev, V. M. Lantratov, and M. Z. Shvarts, in *Conference Record of the 1st World Conference Photo voltaic Energy Conversion*, Hawaii, USA (1994), p. 1894.

[10] V. M. Andreev, V. S. Kalinovsky, V. V. Komin, I. V. Kochnev, V. M. Lantratov, and M. Z. Shvarts, in *Proceedings of the 4th European Space Power Conference*, Poitiers, France (1995), p. 367.

[11] V. M. Lantratov, M. Z. Shvarts, and I. V. Kochnev, Electrochem. Soc. Proc **97−21**, 125 (1997).

[12] L. M. Fraas, J. E. Avery, J. Martin, V. S. Sundaram, G. Girard, V. T. Dinh, T. M. Devenport, J. W. Yerkes, and M. J. O'Neill, IEEE Trans. Electron Devices **37**, 444 (1990).

[13] V. M. Andreev, L. B. Karlina, V. P. Khvostikov, V. D. Rumyantsev, M. Z. Shvarts, and S. V. Sorokina, in *Proceedings of the 13th European Photovoltaic Solar Energy Conference*, Nice, France (1995), p. 329.

[14] K. A. Bertness, S. R. Kurtz, D. J. Friedman, A. E. Kibbler, C. Kramer, and J. M. Olson, Appl. Phys. Lett. **65**, 989 (1994).

[15] M. Yamaguchi and S. Wakamatsu, in *Proceedings of the 25th IEEE Photovoltaic Specialists Conference*, Washington, USA (1996), p. 9.

[16] V. M. Andreev, V. P. Khvostikov, E. V. Paleeva, V. D. Rumyantsev, S. V. Sorokina, M. Z. Shvarts, and I. V. Vasil'ev, in *Proceedings of the 26th IEEE Photovoltaic Specialists Conference*, Anaheim, USA (1997), p. 927.

[17] V. M. Andreev, B. A. Grilikhes, and V. D. Rumyantsev, *Photoelectric Converters of Concentrated Sunlight* [in Russian], Nauka, Leningrad (1989).

[18] V. M. Andreev, B. A. Grilikhes, and V. D. Rumyantsev, *Photovoltaic Conversion of Concentrated Sunlight*, Wiley, New York (1997).

Translated by D. H. McNeill

# Quantum dot lasers: the birth and future trends

N. N. Ledentsov*

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

Despite its early age, laser based on arrays of self-organized quantum dots modified all the basic commandments of the heterostructure laser. Excitonic gain mechanism and discrete energy spectrum in a quantum dot provide principally new ways to control optical properties of the media. Extension of the spectral range using the same substrate will probably soon lead to the appearance of quantum dot lasers on the market.  © *1999 American Institute of Physics.*
[S1063-7826(99)00409-3]

## 1. INTRODUCTION

The advantage of a discrete energy spectrum and efficient overlap of electron and hole wave functions in a quantum dot (QD)[1–4] was recognized already in the early 1980s. When the first papers[3,4] on the possibility of using QDs as active media of a semiconductor laser with strongly improved and temperature insensitive parameters appeared, many scientists and engineers started searching for ways to fabricate quantum dots and to study their properties. However, more than a decade passed until first lasers based on self-organized QDs have been fabricated in 1993 (Ref. 5) and were proven to demonstrate the predicted properties.[6]

## 2. METHODS OF FABRICATING QUANTUM DOTS

Currently the most promising method of fabricating QDs is based on the effect of spontaneous nanoislanding during heteroepitaxial growth. Flat (2D) nanoislands are usually formed by submonolayer deposition and the driving force relates to the surface stress discontinuity at the island edges. The elastic relaxation of the surface stress along the island boundary makes formation of uniform in size nanoislands energetically favorable.[7] After overgrowth 2D islands represent ultrathin nanoscale ''pancakes'' inserted in a wide-gap matrix. The localization energy of carriers and excitons in these islands is relatively small, except for materials with large electron and hole masses are used (II–VI materials, group-III nitrides). In view of the small average thickness of the insertion, a possibility of stacking strained 2D islands by keeping the average strain in the epilayer low exists. Arrays of 2D islands usually provide much narrower absorption or gain peaks.

In the case of 3D islands[7] the driving force relates to the elastic relaxation of the volume strain of the island formed on a lattice-mismatched substrate. Possibility of stable with respect to ripening 3D islands appears if the total surface energy of the island is smaller than the surface energy of the corresponding area of the wetting layer occupied by it. The latter is possible if the strain-induced renormalization of the surface energy of the facets is taken into account.[8]

Oscillator strength in a small QD is not a function of the QD volume. Thus, dense arrays of very small QDs ($10^{12}$ cm$^{-2}$) provide much higher modal gain, as compared to a more dilute array of larger QDs (typically about $10^{10}$–$10^{11}$ cm$^{-2}$). On the other hand, larger QDs can provide much higher localization energy. This gives some flexibility in constructing the device. In the case where one is interested in keeping light maximum absorption or gain values, 2D islands are preferable. High-temperature stability of the threshold current and a maximum long-wavelength shift of the emission (e.g., 1.3 or 1.5-$\mu$m range using GaAs substrates) are realized for 3D islands.

Dense arrays of QDs can demonstrate lateral ordering due to their interaction via the strained substrate.[8] Stacked 3D QD deposition demonstrate vertically-correlated growth.[9,10] The 2D islands demonstrate either correlated or anticorrelated growth, depending on the relative thickness of the spacer layer.[11]

Several other promising methods of fabricating QDs using self-organization phenomena exist (see, e.g., Ref. 12 and the references cited there):

—spontaneous quasiperiodic faceting of crystal surfaces and heteroepitaxial growth of faceted surfaces;

—spontaneous phase separation in semiconductor alloys during growth or slow cooling;

—spontaneous alloy decomposition upon high-temperature annealing.

## 3. EDGE-EMITTING AND VERTICAL-CAVITY QUANTUM DOT LASERS

Evident progress in using QDs has been achieved in the area of semiconductor lasers. Two basic device geometries have been applied. In one case, the light propagates along the plane with QDs, and the resonator represents conventional Fabri–Perot cavity with natural cleavages as mirrors (see Fig. 1, on the right). In the other case, the light is emitted perpendicular to this plane (see Fig. 1, on the left), while the cavity is confined in vertical direction by multilayer stacks of layers that form distributed Bragg reflectors.

The first approach allows fabrication of high-power lasers utilizing advantages of ultralow threshold current density due to QDs, which can prevent dislocation growth and suppress laser mirror overheating by nonradiative surface re-
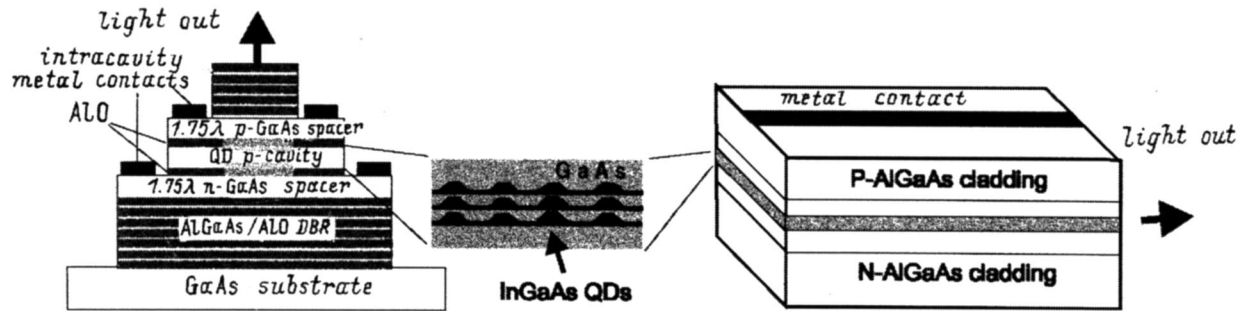
FIG. 1. QDs are used as active media of semiconductor heterostructure lasers in edge-emitting (right) and vertical cavity (left) geometry.

combination due to localization of carriers in QDs.[13] In the second approach lasers with ultralow total currents can be fabricated, and, even more exciting, lasers based on single QD can be potentially realized.

The important events in the QD laser field can be briefly listed here.

—Photopumped QD laser has been first realized by Ledentsov *et al.* in 1993 at the A. F. Ioffe Physicotechnical Institute.[5]

—First QD injection laser has been fabricated in 1994 by a joint collaboration between Technical University of Berlin and A. F. Ioffe Institute. Lasing via the QD ground state and the temperature-insensitive threshold current have been demonstrated.[6]

—Room temperature (RT) operation via quantum dots has been demonstrated.[13–17]

—Ultrahigh material and differential gain in QD lasers have been manifested.[18]

—RT lasing with 60 A/cm$^2$ has been realized.[19]

—Continuous-wave RT high-power operation of a QD laser (1.5 W; Ref. 20) was realized (see Fig. 2). For copper heat sink and waveguide layer thickness of about 0.3 $\mu$m these structures show comparable results to the state of the art quantum well (QW) devices.

—Low-threshold InAs QD laser on InP substrate[21] emitting in the range 1.84–1.9 $\mu$m has been fabricated.

—Significant progress in theoretical understanding of QD lasers with realistic parameters has been achieved.[22,23]

—Operation of QD lasers in the visible spectral range has been demonstrated.[24]

—Vertical-cavity surface-emitting lasers (VCSELs) based on QDs with good properties have been demonstrated by Huffaker *et al.*[25]

—The joint team from the A. F. Ioffe Institute, Air Force Institute of Technology, Ohio, USA, and Technical University of Berlin demonstrated a QD vertical-cavity laser with parameters which fit to the best values for devices of similar geometry based on QWs (Ref. 26) (see Fig. 3).

—1.31-$\mu$m lasing at room temperature with a threshold current density of 240 A/cm$^2$ is demonstrated for the device based on InGaAs QDs in a GaAs host.[27]

In general, the basic parameters of edge-emitting and vertical-cavity QD lasers approached those for QW devices, while the QD laser story is only starting.

## 4. UNEXPECTED RESULTS

### 4.1. Far-Infrared emission in quantum dot lasers

In ultrathin layers, or QWs, there exists a continuum of states at any energy above the subband energy, since the in-plane motion of charge carriers is not limited. If the carrier is excited to the second subband, it relaxes to the first
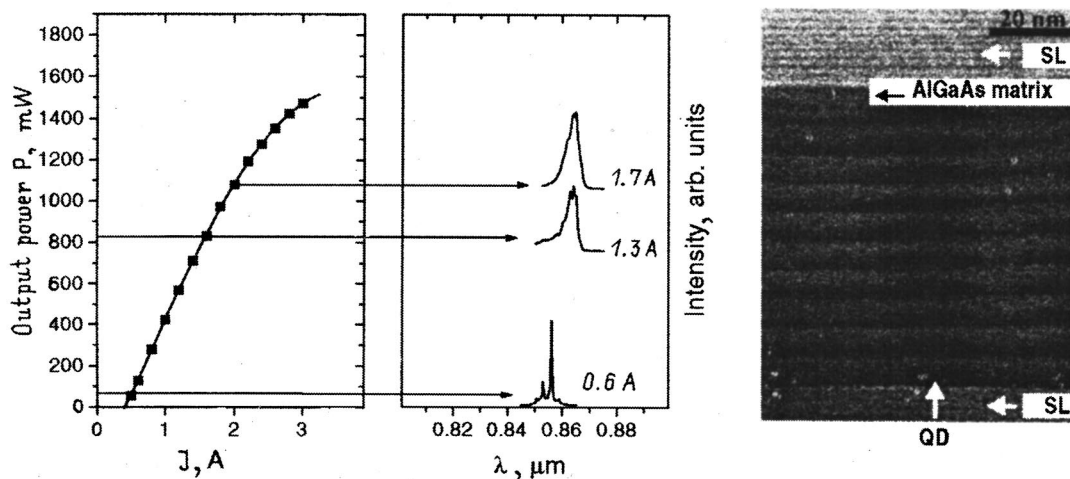


FIG. 2. High-power operation of edge-emitting InAs/AlGaAs QD laser (left). Transmission electron microscopy image of the active region of the high power QD laser (right). Stripe length 850 $\mu$m, width 100 $\mu$m, waveguide region 0.3 $\mu$m.
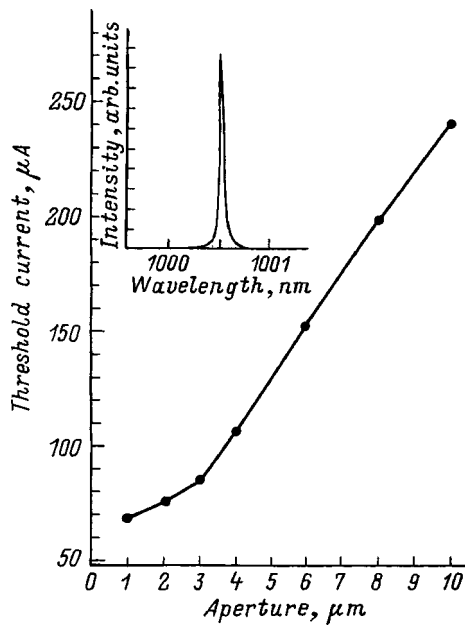
FIG. 3. Threshold current ($J_{th}$) of the QD VCSEL. The emission spectrum at $1.3J_{th}$ is shown in the inset. The quantum efficiency is $\eta = 16\%$ at 10 $\mu$m.

subband via emission of a descrete quantum of energy—an optical (LO) phonon. Because of the continuum nature of electron states in a QW, there always exist states in the first subband to which electrons can scatter within 1 ps. In contrast, in QDs the relaxation time to the ground sublevel takes typically 10–40 ps. The electron must emit a combination of different phonons to match the energy difference. This increases the relative importance of the competing relaxation mechanism via emission of far-infrared (FIR) photons.[28]

In Ref. 28 the FIR emission was observed in QW and QD lasers. The intensity of the FIR spontaneous emission was about an order of magnitude higher in the QD case. Moreover, the FIR emission in QDs has a threshold character since it requires a much higher intensity of the FIR emission in the QD case. Hopefully, this will make it possible to create a new generation of FIR lasers.

### 4.2. Extension of the spectral range of GaAs-based devices to 1.8 $\mu$m

QDs make it possible to cover strategically important spectral ranges of 1.3 and 1.55 $\mu$m using GaAs substrates. This is particularly important for VCSELs, where high-quality monolithic AlAs–GaAs Bragg reflectors and developed oxide technology are available only on GaAs substrates. Recently it was discovered that associates of InAs QDs formed at low substrate temperatures[29] emit light at wavelengths up to 1.8 $\mu$m at 300 K.

### 4.3. Resonant waveguides

The experimentally measured absorption coefficient for structures with stacked CdSe QDs in a ZnSe matrix in the direction perpendicular to the planes with nanoislands approaches $\alpha = 10^5$ cm$^{-1}$ (Ref. 30). High absorption coefficients and lack of exciton screening in dense arrays of QDs

result in ultrahigh QD exciton (or even higher biexciton) gain values under generation of nonequilibrium carriers.[31]

Resonant waveguides are based on the effect of resonant enhancement of the refractive index ($n$) along the contour of the absorption (or gain) curve. To have a significant impact on the waveguiding properties of the media the absorption peak must be strong enough ($\Delta n \sim 0.5$ for $\alpha \sim 10^5$ cm$^{-1}$). For resonant waveguiding it is not necessary to have external cladding of the active region with QDs by layers with significantly lower refractive indices. Practically, it means that lasers can be created in materials having no suitable lattice-matched heterocouple with lower refractive index.

### 4.4. Self-adjusted cavities

In VCSELs the effect of strong resonant modulation of the refractive index serves for self-adjustment of the cavity mode and lasing spectrum. As the material gain of a single QD reaches ultrahigh values due to $\delta$-function-like density of states and negligible homogeneous broadening, even single quantum dot lasing may become possible.[32]

### 4.5. Vertical cavity lasers without Bragg reflectors and cavity

Highly reflective Bragg mirrors on both sides of the cavity are necessary for QW VCSELs, since relatively small maximum gain in these structures [about $10^3$ cm$^{-1}$ (Ref. 4)] requires low external losses of the device. However, if the maximum gain can be made high enough, there is no need to make Bragg mirrors highly reflective. For gain values exceeding $10^5$ cm$^{-1}$ and active layer thickness of 200 nm the facet (or mirror) reflectivity of the order of 30% is enough to achieve vertical lasing. Due to the low finesse of the cavity and the self-adjustment effect, strict fitting of the cavity and the gain spectrum is not required. This effect was demonstrated for 20-times stacked CdSe submonolayer QD insertions in a ZnMgSSe host[33] grown on a GaAs substrate.

### 4.6. Quantum dot composites

The gain of the array of QDs is not defined by a simple sum of gains of single QDs. Interaction of electromagnetic fields of anisotropic QDs or anisotropic QD lattices makes the splitting of the TE and TM modes for the same QD exciton transition inavoidable. This effect can result in splitting as large as several tens of meV, as was predicted theoretically and is proven experimentally[33]. Maximum gain of the QD ansemble is also a strong function of the relative arrangement of QDs.[33]

### 4.7. "quantum-well" lasers

Most of recent industrial QW lasers are based on thin layers of alloys used as active regions. It is clear now that these layers in most cases exhibit quasiperiodic nanoscale compositional modulations that create in many cases dense arrays of quantum wire- or QD-like structures.[12] By using the same average alloy composition the luminescence peak energy can be tuned at several hundred meV by tuning the growth conditions.[34] Careful evaluation of the impact of such

effects on lasing characteristics of modern lasers is necessary to clear up the role of self organized QDs or quantum wires in this case.

## 5. CONCLUSION

QDs (Ref. 35) modified all the basic commandments of the double heterostructure (DHS) laser.[36,37]

| DHS, DHS QW | QD |
|---|---|
| —lattice matching | undesirable |
| —material gain | orders of magnitude higher |
| —exciton screening | is not important |
| —homogeneous broadening at RT | is small |
| —cladding with low $n$ layers | is not necessary |
| —VCSEL: Bragg reflectors and cavity | are not necessary |
| —lasing in optical and near-IR range | and simultaneous FIR emission |
| —one family (III-V–III-V,...) | is not necessary (InAs/Si QDs) |
| limited wavelength range on GaAs | is extended to 1.8 $\mu$m |

It appears that the QD laser seems to be a completely new device with properties which can remarkably expand our possibilities in many applications, rather than simply a laser with some parameters improved with respect to the DHS or DHS QW laser.[3,4]

*E-mail: shura.lebe@pop.ioffe.rssi.ru

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

[2] Al. L. Efros and A. L. Efros, Fiz Tech. Poluprovodn, 16 1209 (1982) [Sov. Phys. Semicond. **16**, 772 (1982)].

[3] Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).

[4] M. Asada, M. Miyamoto, and Y. Suematsu, IEEE J. Quantum Electron. **QE-22**, 1915 (1986).

[5] N. N. Ledentsov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, M. V. Maximov, I. G. Tabatadze, and P. S. Kop'ev, Fiz. Tekh. Poluprovodn **28**, 1484 (1994) [Semiconductors **28**, 832 (1994)].

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

[7] V. I. Marchenko, JETP Lett. **33**, 381 (1981).

[8] V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, Phys. Rev. Lett. **75**, 2968 (1995).

[9] L. Goldstein, F. Glas, J. Y. Marzin, M. N. Charasse, and G. Le Roux, Appl. Phys. Lett. **47**, 1099 (1985).

[10] N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. Schmidt, R. Heitz, J. Böhrer, D. Bimberg, V. M. Ustinov, V. A. Shchukin, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, A. O. Kosogov, P. Werner, G. Richter, U. Gösele, and J. Heydenreich, Proc. MSS7 (1995) [Solid-State Electron. **40**, 785 (1996)].

[11] V. A. Shchukin, D. Bimberg, V. G. Malyshkin, and N. N. Ledentsov, Phys. Rev. B **57**, 12 262 (1998).

[12] N. N. Ledentsov, Prog. Cryst. Growth Charact. **35**, 289 (1997).

[13] N. N. Ledentsov, J. Böhrer, D. Bimberg, S. V. Zaitsev, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, A. O. Kosogov, U. Gösele, and S. S. Ruvimov, in *Mater. Res. Soc. Symp. Proc.*, edited by R. J. Shul, S. J. Pearton, F. Ren, and C.-S. Wu (Pittsburgh, 1996), V. 421, p. 133.

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

[15] K. Kammath, P. Bhattacharya, T. Sosnowski, and J. Phillips, Electron. Lett. **30**, 1374 (1996).

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

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

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

[19] N. N. Ledentsov, V. A. Shchukin, M. Grundmann, N. Kirstaedter, J. Böhrer, O. Schmidt, D. Bimberg, S. V. Zaitsev, V. M. Ustinov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, A. O. Kosogov, S. S. Ruvimov, P. Werner, U. Gösele, and J. Heydenreich, Phys. Rev. B **54**, 8743 (1996).

[20] M. V. Maximov, Yu. M. Shernyakov, A. F. Tsatsul'nikov, A. V. Lunev, A. V. Sakharov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, A. R. Kovsh, P. S. Kop'ev, L. V. Asryan, Zh. I. Alferov, A. O. Kosogov, and P. Werner, J. Appl. Phys. **83**, 5561 (1998).

[21] V. M. Ustinov, A. R. Kovsh, A. E. Zhukov, A. Yu. Egorov, N. N. Ledentsov, A. V. Lunev, Yu. M. Shernyakov, M. V. Maksimov, A. F. Tsatsul'nikov, B. V. Volovik, P. S. Kop'ev, and Zh. I. Alferov, Pis'ma Zh. Tekh. Fiz **24**, 17 (1999) [Tech. Phys. Lett. **24**(1), 49 (1999)].

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

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

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

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

[26] J. A. Lott, N. N. Ledentsov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, and D. Bimberg, Electron. Lett. **33**, 1150 (1997).

[27] D. L. Huffaker, G. Park, Z. Z. Zhou, O. B. Shchekin, and D. G. Deppe (unpublished).

[28] L. E. Vorob'ev, D. A. Firsov, V. A. Shalygin, V. N. Tulupenko, Yu. M. Shernyakov, N. N. Ledentsov, V. M. Ustinov, and Zh. I. Alferov, JETP Lett. **67**, 275 (1998).

[29] M. V. Maximov, N. N. Ledentsov, A. F. Tsatsul'nikov, B. V. Volovik, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, A. R. Kovsh, P. S. Kop'ev, Zh. I. Alferov, D. Bimberg, I. P. Soshnikov, and P. Werner, *Proc. ICPS24* (Jerusalem, 1989).

[30] G. N. Aliev, A. D. Andreev, R. M. Daitsev, S. V. Ivanov, S. V. Sorokin, A. B. Kapustina, I. L. Krestnikov, M. E. Sasin, and R. P. Seisyan, J. Cryst. Growth **184/185**, 315 (1989).

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

[32] N. N. Ledentsov, D. Bimberg, V. M. Ustinov, M. V. Maximov, Zh. I. Alferov, V. P. Kalosha, and J. A. Lott, Semicond. Sci. Technol. (1998) (in press); *Proc. ICPS 24* (Jerusalem, 1998).

[33] I. L. Krestnikov, P. S. Kop'ev, Zh. I. Alferov, M. Straßburg, N. N. Ledentsov, A. Hoffmann, D. Bimberg, I. P. Soshnikov, and P. Werner, *Proc. ICPS 24* (Jerusalem, 1998); V. P. Kalosha, G. Ya. Slepyan, S. A. Maksimenko, N. N. Ledentsov, I. L. Krestnikov, Zh. I. Alferov, and D. Bimberg, *Proc. ICPS 24* (Jerusalem, 1998).

[34] S. W. Jun, T.-Y. Seong, J. H. Lee, and B. Lee, Appl. Phys. Lett. **68**, 3443 (1996).

[35] D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Hetero-structures* (J. Wiley, 1998).

[36] Zh. I. Alferov and R. F. Kazarinov, Double Heterostructure Laser, Authors Certificate No 27448, Application No 950840 with a priority from March 30, 1963.

[37] Zh. I. Alferov, in *Proceedings of the Nobel Symposium 99* (Arild, Sweden, 1996); Phys. Scr. **68**, 32 (1996).

# Quantum dots: Paradigm changes in semiconductor physics

D. Bimberg

*Institute für Festkörperphysik, Technische Universität Berlin, 10623 Berlin, Germany*

Deposition of one or a few monolayers of a semiconductor having a lattice constant largely different from the underlying substrate leads to formation of coherent ''quantum dot arrays'' of densities beyond $10^{11}\,\mathrm{cm}^{-2}$ in a matter of seconds. Self-organization effects govern their massively parallel formation. Fundamental paradigms of semiconductor physics must be changed in describing such quantum dots or their ensembles.  © *1999 American Institute of Physics.* [S1063-7826(99)00509-8]

## 1. INTRODUCTION

Recent breakthroughs in solid state physics are related to the fast development of the new classes of structures—self-organized quantum dots (QDs). These unique objects combine advantages of a bulk semiconductor with those of single atoms. Their physical properties resemble an atom in a cage. It is possible to touch a semiconductor quantum dot with a tip of a scanning tunneling microscope, inject charged carriers and monitor emission from a single QD, and realize many other unique experiments. Interaction of carriers in QDs is defined completely by the laws of quantum mechanics and no ''motion'' of carriers is possible in a traditional sense. The main mechanism of the information transfer between QDs becomes single electron tunneling and Coulomb interaction of carriers in neighboring dots. These give rise to novel directions in microelectronics: single electronics and quantum dot cellular automata computing. QDs allow significant improvements of characteristics of many modern devices: e.g., infrared detectors, light-emitting diodes, lasers, and solar cells. Since the publication of the first experimental[1] and theoretical[2,3] papers on QDs, some fundamental understanding of their physical properties, their growth, and their device relevance has been obtained.[4–6]

## 2. ELECTRONIC SPECTRUM OF A QUANTUM DOT

To understand the reasons for the unique properties of QDs let us first refer to the electronic spectrum of a 3D solid. It is well known that a single atom has discrete energy levels separated by energy gaps, as shown in Fig. 1. When an atom is excited, the electron goes to the higher energy level, and when it relaxes back to the ground state, a photon with strictly defined energy is emitted. The width of the emission or absorption line ($\Delta E$) is defined by the fundamental relation with the lifetime of the electron in the upper state. The uncertainty in the emitted energy is

$$\tau \Delta E \geqslant \hbar, \tag{1}$$

where $\tau$ is the total lifetime of the electron in the excited state.

In contrast to the case of a diluted gas of atoms, the atoms in crystals are strongly bound to each other. Their high density in crystals plays a very important role in modern solid-state devices allowing high absorption or (in the case of population inversion) high gain, providing high conductivity and making possible high density flows of charged carries through the crystal. Accordingly, a modern semiconductor laser with length of 1 mm and cross section of $10^{-4}\,\mathrm{mm}^2$ can emit continuous light with power of several watts, while a corresponding gas laser is several meters long.

At the same time, small separation between atoms makes interaction of their electron levels unavoidable. This interaction results in the formation of wide bands of allowed states, in contrast with the discrete ($\delta$-function-like) energy spectrum of single atoms. In semiconductors, the last filled band of allowed states is called ''valence band'' and the next empty band is called ''conduction band.'' Due to the broad spectrum of allowed states in these bands, a wide range of transition energies between electrons from the filled valence band to empty states in the conduction band is possible. The absorption band is rather broad, on the order of several electron volts, in marked contrast with the sharp line absorption spectrum of a single atom. The excited electrons in the conduction band and the empty states in the valence band (''holes'') can move in the crystal via tunneling between sites of the crystal lattice. Since the atom potential profile in a crystal is periodic, electrons and holes can move freely through the crystal, as free carriers do in vacuum. However, the motion of charged carriers in crystals is described by a mass different from that of free electrons, defined by the crystal field. The carriers are therefore called ''quasiparticles.'' In the widely used optoelectronic III–V materials (gallium arsenide—GaAs; indium arsenide—InAs, etc.), the effective masses of the electrons are in the range of 0.01–0.1 of the electron mass in vacuum.

Wide bands of allowed states in the crystal provide many possibilities for the scattering of electrons and holes. Lattice vibrations easily stimulate transitions of charge carriers in the energy range defined by the lattice temperature and/or scatter their direction. The tails of the carrier distribution near the bottom of the conduction band and the top of the valence band increase markedly with temperature. Thus, the concentration of carriers per energy interval near the band edge drops. For the same concentration of injected car-
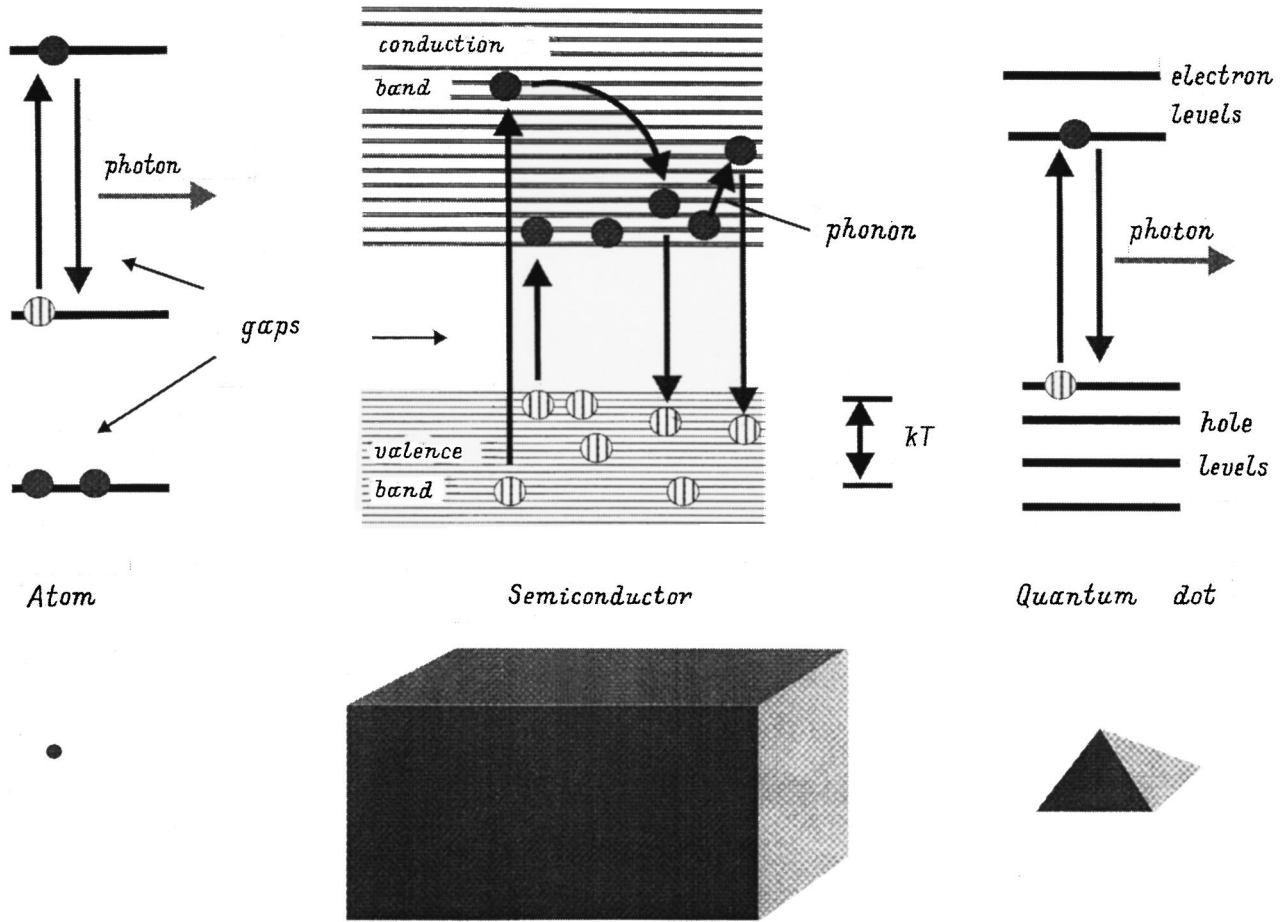
FIG. 1. Schematic representation of energy levels in a single atom, a bulk semiconductor and a quantum dot.

riers, broadening of their energy spectra, among other disadvantages, leads to a decrease in the maximum gain and to a degradation of laser performance.

The situation changes dramatically if the motion of the charged carrier in the crystal is limited to a very small volume, e.g., a three-dimensional rectangular box. Localization of carriers can be provided by a surrounding (host) material. In the case of a laser, it is important that the host material have a larger band gap than the box material and that the potential wells are attractive for electrons and for holes. Since an electron shows properties of both a particle and a wave if the size of the box is small, the electron energy spectrum is quantized similarly to the case of electron quantization in the attracting Coulomb potential of a nucleus. In a simplified case of infinite barriers at the box-host interface, the quantum-well energy is described by

$$E_{x,y,z} = \frac{\hbar^2}{2m_e^*} \left( \frac{\pi n}{L_{x,y,z}} \right)^2, \qquad (2)$$

where $m_e^*$ is the electron effective mass, $E_{x,y,z}$ is the quantum-well energy due to electron localization in the box with dimensions $L_x$, $L_y$, and $L_z$, respectively, and $n = 1, 2, 3,...$ .

Electrons in crystals usually have rather small effective mass, and an already large size of the box of about 10 nm can result in a large energy separation between electron sub-

levels (about 100 meV for GaAs QD). The latter value significantly exceeds the thermal energy at room temperature (26 meV), and thus the population of excited states can be avoided. In this sense, there will be no temperature dependence of the optical spectrum of such a box in a wide temperature range, and the realization of temperature-insensitive devices is possible.

## 3. SELF-ORGANIZED GROWTH OF QDs

Until recently, there were no ideas of how to realize QDs with size and uniformity required for applications and compatible with modern semiconductor technology. Traditional methods based on patterning of structures with ultrathin layers suffered either from insufficient lateral resolution or introduced heavy damage in the material upon processing. There were also predictions that, even if ideal QDs could be fabricated, they can be hardly used for real devices since ultralong relaxation times between electron sublevels were expected.

The breakthrough how to fabricate QDs with required properties came by using an effect which is traditionally considered as undesirable by crystal growers. It was found that if one deposits a layer of a material having a different lattice constant with the substrate, the growth of a strained material first proceeds in a planar mode, a ''wetting'' layer is formed.

However, at some critical thickness this planar growth is followed by a Stranski-Krastanow mode. Three-dimensional nanoscale islands are formed on surfaces, as was demonstrated first for the case of growth of indium arsenide on gallium arsenide in 1985 by a group at CNET, France.[5] When these islands are covered with GaAs, a GaAs ''pie'' with InAs ''raisins'' is formed. Since InAs has a much smaller band gap than GaAs, an array of InAs QDs is formed. Initially this technique did not attract much attention since dislocation-free QDs uniform in size and shape could not be produced.

The driving force for the formation of three-dimensional islands is related to elastic strain relaxation. The material on top of the pyramid can relax elastically expanding in vacuum, which gives rise to a significant energy advantage. For 45° facet angle of the pyramid, 60% of the elastic energy accumulated in the biaxially compressed flat layer is relaxed.[6,7] On the other hand, formation of pyramids results in an increase of the total surface area. If formation of islands results in an increase of the surface energy of the system, the initially formed islands will undergo ripening since the system will reduce the total surface area covered by QDs.

Quantum dots uniform in size and shape, stable with respect to ripening, can be fabricated only if the total surface energy of the island is smaller than the surface energy of the corresponding area of the wetting layer occupied by it. If one takes into account that the major surface properties, e.g., surface reconstruction, surface stress etc., are strongly affected by the strain state of the crystal, one can conclude that they can differ significantly for the strained wetting layer and the facet of the relaxed pyramid. Numerical estimates of the strain-induced renormalization of the surface energy, made by Shchukin et al.,[7] indicate that ''equilibrium,'' equisized, and equishaped islands which do not undergo ripening, probably can be formed.

Optimization of growth parameters to realize equilibrium arrays is a difficult task to be solved for each material system separately.[8] If islands uniform in size and shape are formed, one speaks of ''self-organized quantum dots'' since this system represents a clear example of spontaneous formation of macroscopic order from initially random distributions. In the case of dense arrays of QDs, their interaction via the strained substrate makes their lateral ordering favorable.[7] Growth on patterned surfaces can also lead ordered QDs.[9] For multistack QD deposition vertically correlated growth of QDs has been demonstrated[5,8] and, thus, quasicrystals composed of quantum dots either in two or in all three dimensions can be fabricated. For islands having a two-dimensional shape, either correlated or anticorrelated growth is possible, depending on the relative thickness of the spacer layer.[10]

## 4. PROOF OF ELECTRONIC QUANTUM DOT

An important breakthrough in understanding the electronic properties of semiconductor QDs occurred when it was demonstrated that ultranarrow luminescence lines from single InAs QDs (Refs. 11 and 12) exhibit no broadening with temperature,[13] a very unusual phenomenon for any
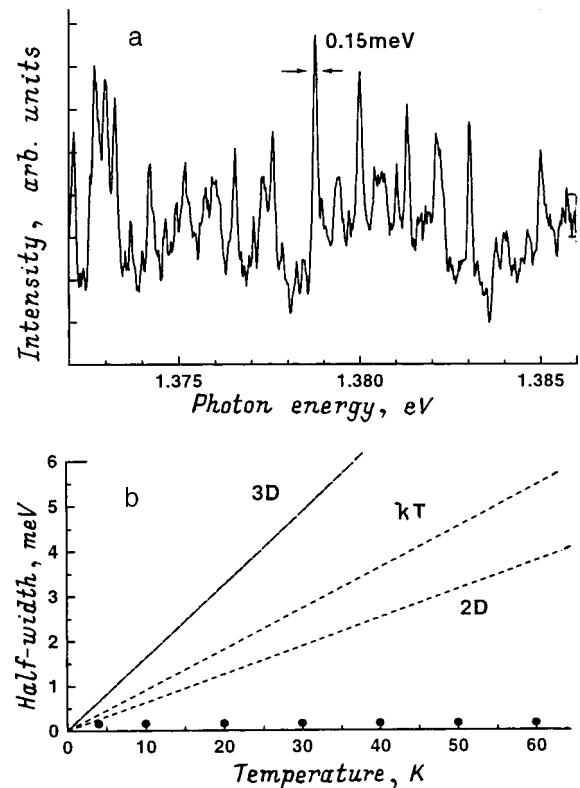


FIG. 2. Ultrasharp luminescence lines from single QDs at 20 K (a) which do not show broadening with temperature increase (b).

electron-hole emission in semiconductors, in complete agreement with theoretical predictions for electronic QDs (see Fig. 2).

A lot of progress has been made in the last few years, more than in decades of previous research. We have learned about QDs in different material systems, about the electron spectrum in type-II QDs, and about radiative recombination and relaxation processes. Numerous teams contributed to the development of this subject. Among them Ioffe Institute, CNET, UCSB, USC, Stanford, Lund, Sheffield, and Nottingham Universities, Max-Planck Institute in Stuttgart, and many other groups and institutions.

## 5. SINGLE ELECTRONICS

Single electronics is based on quantum effects related to the charging of QDs and tunneling of carriers through it. If the ground state of a QDs is not occupied, external electrons can tunnel to this level. If the QD ground state is occupied with one electron, an additional energy is necessary to put another electron in the QD because of the Coulomb repulsion with the existing electron. This ''Coulomb blokade'' regime makes it possible to realize principally new logic and operational principles for new generations of microelectronic circuits.

## 6. QUANTUM DOT MOLECULES. QUANTUM DOT CELLULAR AUTOMATA

If QDs are placed together forming a cell with well-defined symmetry and allowing ''switching of carriers'' be-
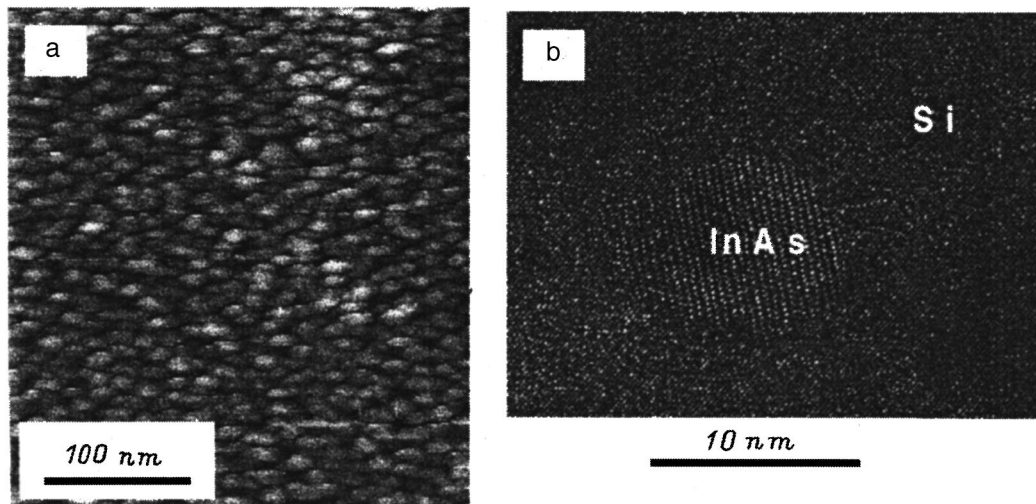
FIG. 3. STM image of free-standing InAs QDs of Si (100) surface (a) and high-resolution transmission electron microscopy image of Si-covered coherent InAs QDs in Si matrix (b).

tween two or more energetically equivalent positions in the cell, and if these cells are combined in a network, a new concept for construction of integrated circuits appears. In this case it is not the current which carries the information, but the polarization of the QD chain or the QD molecule. In principle, the operation of such a molecule can be made practically energy dissipation-free, and the information propagates at the speed of electromagnetic waves.[14]

## 7. DATA STORAGE

Ultrahigh area density of QDs and potentially very high localization energy of carriers in QDs make it possible to realize a new generation of memory devices. QDs can be also made of magnetic materials or serve as an active medium in photon-memory devices.

## 8. FAR-INFRARED DETECTORS AND EMITTERS

Size quantization of carriers in all three dimensions results in lifting of **k**-selection rules that prohibit intrasubband absorption in quantum wells (QWs) for the light propagating perpendicular to the surface of the wafer. This makes QD far-infrared detectors advantageous with respect to those based on QWs. Reduction of the depopulation rate of upper sublevels[15] in view of the necessity of multiphonon emission to fit the intrasublevel energy in properly designed QDs results in much better probability of attaining population inversion for intrasublevel transitions in QDs.

## 9. LIGHT EMITTERS

Electron and hole overlap is close to unity in a QD. The lifting of the **k**-selection rule for radiative annihilation of excitons and the lack of screening of Coulomb interaction between electrons and holes results in ultrahigh material gain in a QD. The material gain is also inversely proportional to the QD volume, and dense arrays of very small QDs can result in maximum gain on the order of $10^5$ cm$^{-1}$. Such a high gain affects the refractive index in the vicinity of the

absorption/gain resonance, resulting in a strong photon energy-dependent oscillatory variation of the refractive index, which leads to resonant waveguiding without a heterostructure.[16,17]

No Fermi function describes the statistical properties of a QD ensemble like carrier distribution. Basic equations of microstates describe static and dynamic properties like gain, capture, recombination, etc.[18] High-temperature stability of the threshold current in a QD laser is demonstrated,[19] in agreement with the theoretical predictions, and theoretical understanding of the QD laser has been achieved.[20,21] Competitive edge[22–24] and surface-emitting[25] lasers are created.

## 10. NEW CHALLENGES IN HETEROEPITAXY: InAs QUANTUM DOTS IN Si

Another field, in which application of QDs can play an extraordinary role, is related to silicon. Silicon, the major material for modern microelectronics, provides very low probability for injected electrons and holes to recombine radiatively, and cannot be used in lasers and light-emitting devices. If it is possible to insert narrow-gap QDs (e.g., made from InAs), which have a high probability of radiative recombination, in such a way that electrons and holes will be trapped in the QDs, a silicon-based device with extremely efficient radiative recombination can be created. The possibility of depositing such InAs quantum dots on Si surface was recently demonstrated.[26]

A scanning tunneling microscopy image of free-standing InAs QDs on a silicon (100) surface and a high-resolution transmission electron microscopy image of silicon-overgrown InAs QDs are shown in Fig. 3a and 3b, respectively. Coherent InAs QDs in a silicon host demonstrate broad photoluminescence band peaking at 1.3 $\mu$m at a temperature 77 K and at 1.6 $\mu$m at a temperature 300 K.

New generations of photonic and electronic devices based on quantum dots, which are revolutionizing semiconductor technology and industry, are emerging presently.

[1] A. I. Ekimov and A. A. Onushzhenko, JETP Lett. **34**, 2225 (1981).

[2] Al. L. Efros and A. L. Efros, Fiz. Tekh. Poluprovodn. (St. Peterburg) **16**, 1209 (1982) [Sov. Phys. Semicond. **16**, 772 (1982)].

[3] Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982); M. Asada, M. Miyamoto, and Y. Suematsu, IEEE J. Quantum Electron. **QE-22**, 1915 (1986).

[4] D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Hetero-structures* (J. Wiley, 1998).

[5] L. Goldstein, F. Glas, J. Y. Marzin, M. N. Charasse, and G. Le Roux, Appl. Phys. Lett. **47**, 1099 (1985).

[6] M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **51**, 11 969 (1995); O. Stier, M. Grundmann, and D. Bimberg, Phys. Rev. B (in press).

[7] V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, Phys. Rev. Lett. **75**, 2968 (1995); V. A. Shchukin and D. Bimberg, Rev. Mod. Phys. (1999) (in press).

[8] N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. Schmidt, R. Heitz, J. Böhrer, D. Bimberg, V. M. Ustinov, V. A. Shchukin, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, A. O. Kosogov, P. Werner, U. Richter, U. Gösele, and J. Heydenreich, *Proc. MSS7* (1995) [Solid-State Electron. **40**, 785 1996].

[9] T. T. Ngo, P. M. Petroff, H. Sakaki, and J. L. Merz, Phys. Rev. B **53**, 9618 (1996).

[10] V. A. Shchukin, D. Bimberg, V. G. Malyshkin, and N. N. Ledentsov, Phys. Rev. B **57**, 12 262 (1998).

[11] N. N. Ledentsov, M. Grundmann, N. Kirstaedter, J. Christen, R. Heitz, J. Böhrer, F. Heinrichsdorff, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, in *Proceedings of the 22nd International Conference on Physics of Semiconductors*, Vancouver (1994), edited by D. J. Lockwood (World Scientific, Singapore, 1995), V. 3, p. 1855.

[12] J.-Y. Marzin, J. M. Gerard, A. Izraël, D. Barrier, and G. Bastard, Phys. Rev. Lett. **73**, 716 (1994).

[13] M. Grundmann, J. Christen, N. N. Ledentsov, J. Böhrer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, and Zh. I. Alferov, Phys. Rev. Lett. **74**, 4043 (1995).

[14] C. S. Lent, P. D. Tuogaw, and W. Porod, Appl. Phys. Lett. **62**, 714 (1993).

[15] H. Benisty, C. M. Sotomayor Torres, and C. Weisbuch, Phys. Rev. B **44**, 10 945 (1991); T. Inoshita and H. Sakaki, Phys. Rev. B **46**, 7260 (1992).

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

[17] I. L. Krestnikov, M. V. Maximov, A. V. Sakharov, P. S. Kop'ev, Zh. I. Alferov, N. N. Ledentsov, D. Bimberg, and C. M. Sotomayor Torres, J. Cryst. Growth **184/185**, 545 (1998).

[18] M. Grundmann and D. Bimberg, Phys. Rev. B **55**, 4181 (1997).

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

[20] L. V. Asryan and R. A. Suris, IEEE J. Select. Topics Quant. Electron **3**, 148 (1997).

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

[22] N. N. Ledentsov, V. A. Shchukin, M. Grundmann, N. Kirstaedter, J. Bhrer, O. Schmidt, D. Bimberg, S. V. Zaitsev, V. M. Ustinov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, A. O. Kosogov, S. S. Ruvimov, P. Werner, U. Gösele, and J. Heydenreich, Phys. Rev. B **54**, 8743 (1996).

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

[24] M. V. Maximov, Yu. M. Shernyakov, A. F. Tsatsul'nikov, A. V. Lunev, A. V. Sakharov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, A. R. Kovsh, P. S. Kop'ev, L. V. Asryan, Zh. I. Alferov, A. O. Kosogov, and P. Werner, J. Appl. Phys. **83**, 5561 (1998).

[25] J. A. Lott, N. N. Ledentsov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, and D. Bimberg, Electron. Lett. **33**, 1150 (1997).

[26] G. E. Cirlin, V. N. Petrov, V. G. Dubrovsky, S. A. Masalov, A. O. Golubok, N. I. Komyak, N. N. Ledentsov, Zh. I. Alferov, and D. Bimberg, Pis'ma Zh. Tekh. Fiz. **24**(4), 290 (1998) [Tech. Phys. Lett. **24**(8), 10 (1998); A. F. Tsatsul'nikov, A. Yu. Egorov, P. S. Kop'ev, A. R. Kovsh, M. V. Maximov, N. A. Bert, V. M. Ustinov, B. V. Volovik, A. E. Zhukov, Zh. I. Alferov, G. E. Cirlin, A. O. Golubok, S. A. Masalov, V. N. Petrov, N. N. Ledentsov, R. Heitz, M. Grundmann, D. Bimberg, I. P. Soshnikov, P. Werner, and U. Gösele, *Proceedings ICPS 24*, Jerusalem, 1998 (World Scientific, 1998).

# Electrons and phonons in quantum wells

J. Požela,* A. Namajūnas, K. Požela, and V. Jucienė

*Semiconductor Physics Institute, 2600 Vilnius, Lithuania*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1049–1053 (September 1999)

The modulation of electron and polar optical phonon states in an AlGaAs/GaAs/AlGaAs quantum well (QW) with an inserted thin AlAs barrier is considered. The OW width dependence of electron–phonon scattering rates are estimated. The large contribution to the change of the electron subband population, the photovoltaic effect, and the electron mobility in the QW accounts for the resonant intersubband scattering of electrons by interface phonons. The decrease of electron mobility limited by polar optical phonon scattering with increasing carrier concentration in the QW is established. The conditions for the increase of mobility in the QW by inserting the AlAs barrier are found. © *1999 American Institute of Physics.*
[S1063-7826(99)00609-2]

## INTRODUCTION

The confinement of electrons and phonons in semiconductor quantum wells (QW's) improves their electrical and optical properties with respect to bulk materials. Semiconductor QW structures have received considerable attention because of their potential for use in lasers operating in the infrared region,[1–3] for nonlinear optical elements, and for infrared photoelectric detectors.[4] The enhancement of electron mobility in QW's has attracted increasing attention by many authors.[5–14]

In this paper we consider the role of modulation of electron and polar optical (PO) phonon states in AlGaAs/GaAs/AlGaAs QW's in the change of electron mobility, photoexcited electron subband population and photovoltaic effect. The conditions for electron mobility enhancement and for inversion of photoexcited electron subband population are determined. The case of AlGaAs/GaAs/AlGaAs structure with and without an inserted AlAs barrier is considered (Fig. 1).

This paper is organized as follows. In Sec. 1 the peculiarities of electron–phonon scattering rate in QW's are considered. The dependences of photoexcited electron subband population and photovoltaic effect on intersubband scattering rates are determined in Secs. 2 and 3. The contribution of intra- and intersubband electron-PO-phonon scattering to the electron mobility in the QW is calculated in Sec. 4.

## 1. ELECTRON–PO-PHONON SCATTERING RATE IN A QUANTUM WELL

According to the Fermi golden rule, the transition frequency of electrons, which are confined in the QW, from the initial state $\mathbf{k}_i$, $E_{si}$ to any final states $\mathbf{k}_f$, $E_{sf}$ in all $f$ electron subbands by emission (absorption) of all $\nu$-modes of PO phonons can be written as

$$W^{\pm}(k_i, E_{si}) = \sum_{\nu} \sum_{f} \sum_{\mathbf{k}_f} \frac{4\pi m e^2}{\hbar^3} |G_{\nu}(z)|^2 F_{q\nu}$$

$$\times \left( N_{q\nu} + \frac{1}{2} \pm \frac{1}{2} \right) \delta(k_f^2 - k_i^2 + \beta_{\pm}), \quad (1)$$

$$G_{\nu}(z) = \int_{L_e} \varphi_{e1}(z) \varphi_{e2}(z) \varphi_{q\nu}(z) dz, \quad (2)$$

where $z$ is the quantization direction, $L_e$ is the area where electron wave function $\varphi_{ei} \neq 0$, $m$, $e$, $\mathbf{k}_i$, $E_{si}$, and $\varphi_{ei}$ are the electron mass, charge, wave vector, subband energy, and normalized wave function, respectively, $\varphi_{q\nu}(z)$ is the $z$-component of the $\nu$-mode phonon potential function, $N_{q\nu}$ is the number of $\nu$-mode phonons, and

$$\beta_{\pm} = \frac{2m}{\hbar^2} (E_{sf} - E_{si} \pm \hbar \omega_{\nu}). \quad (3)$$

The upper (plus) sign is for phonon emission and the lower (minus) one is for phonon absorption.

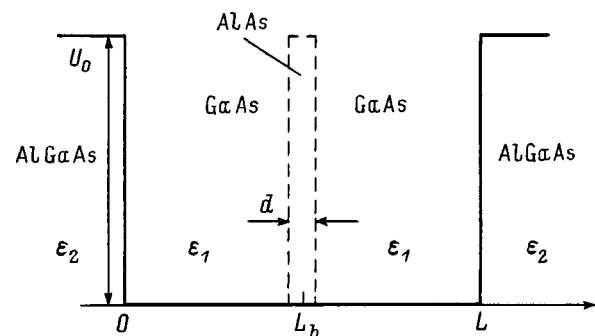In the dielectric continuum approach the square of phonon normalization coefficient can by written as[11]



FIG. 1. Schematic view of an $Al_{0.3}Ga_{0.7}As/GaAs/Al_{0.3}Ga_{0.7}As$ QW structure. $L$ is the QW width, $L_b$ is the position of a thin AlAs barrier with the thickness $d$. $U_0 = 0.3$ eV is the heterojunction potential.

$$F_{q\nu} = \frac{\hbar}{S} \left[ \left( \frac{d\varepsilon_1}{d\omega} \right)_\nu \int_1 f_\nu dz + \left( \frac{d\varepsilon_2}{d\omega} \right)_\nu \int_2 f_\nu dz \right]^{-1}, \quad (4)$$

where $S$ is the in-plane of the normalization area, $\varepsilon_1$ and $\varepsilon_2$ are the dielectric functions, and

$$f_\nu = q_{0\nu}^2 |\varphi_{q\nu}|^2 + \left| \frac{d\varphi_{q\nu}}{dz} \right|^2,$$

where $q_{0\nu}$ is the wave vector of emitted (absorbed) $\nu$-mode phonon in the plane of QW:

$$q_{0\nu}^2 = k_i^2 + k_f^2 - 2k_i k_f \cos\theta, \quad k_f^2 = k_i^2 + \beta_\pm, \quad (5)$$

where $\theta$ is the angle between the initial wave vector $\mathbf{k}_i$ and the final one $\mathbf{k}_f$. The subscripts 1 and 2 correspond to the QW and barrier regions, respectively.

The electron wave functions are obtained by solving the Schrödinger equation for the QW under consideration. The phonon potentials are obtained by solving the Laplace equation using dielectric continuum approach. The phonon potentials in the AlGaAs/GaAs/AlGaAs QW with a thin AlAs barrier can be divided into three separate branches: interface phonons localized at the AlAs barrier with the potential $\varphi_q^B$, interface phonons localized at the AlGaAs/GaAs heterojunction with the potential $\varphi_q^{HJ}$, and confined phonons with the potential $\varphi_q^C$.

Using Eq. (1) and the calculated expressions for the electron wave function and phonon potentials, we can obtain the electron–phonon scattering rates by various phonon branches. For the confined electron scattering rate by confined phonons

$$W_{if}^{C\pm} = F_C^\pm \sum_{n=1}^\infty |G_C(z)|^2 \frac{2}{L} \int_0^{2\pi} \frac{d\theta}{(q_\nu^2 + q_z^2)}, \quad (6)$$

with

$$F_C^\pm = \frac{me^2}{2\pi\hbar^3} \left( N_{qC} + \frac{1}{2} \pm \frac{1}{2} \right) \frac{\hbar\omega_0}{2} \left( \frac{1}{\varepsilon_{1\infty}} - \frac{1}{\varepsilon_{1s}} \right), \quad (7)$$

$$G_C(z) = \int_{L_e} \varphi_{ei} \varphi_{ef} \sin(q_z z) dz, \quad q_z = n\frac{\pi}{L}, \quad (8)$$

where $\hbar\omega_0$ is the energy of a bulk longitudinal phonon, $\varepsilon_{1\infty}$ and $\varepsilon_{1s}$ are the high-frequency and static dielectric constants, respectively, and $L$ is the QW width.

For the confined electron scattering by the heterojunction $\nu$-mode interface phonon at $qL>2$

$$W_{if\nu}^{HJ\pm} = F_{IF\nu}^\pm \int_0^{2\pi} |G_{HJ\nu}(z)|^2 \frac{d\theta}{2q_{0\nu}}, \quad (9)$$

where

$$F_{IF\nu}^\pm = \frac{me^2}{2\pi\hbar^3} \left( N_{q\nu} + \frac{1}{2} \pm \frac{1}{2} \right) \hbar \left[ \left. \frac{d\varepsilon_1}{d\omega} \right|_\nu + \left. \frac{d\varepsilon_2}{d\omega} \right|_\nu \right]^{-1}, \quad (10)$$

$$G_{HJ\nu}(z) = \int_{L_e} \varphi_{ei} \varphi_{ef} \varphi_{q\nu}^{HJ} dz. \quad (11)$$

For the confined electron scattering by the barrier $\nu$-mode interface phonons at $qL>2$, $qd \ll 1$
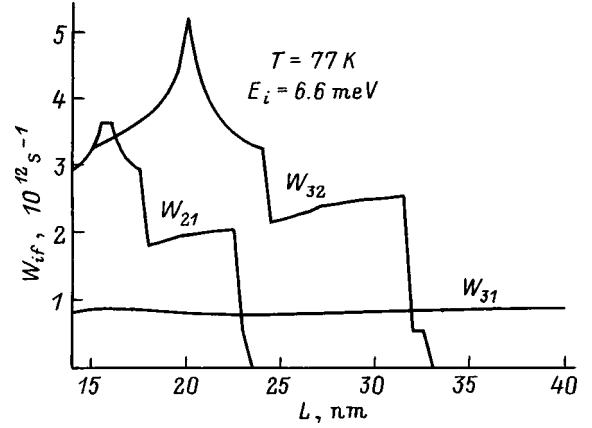


FIG. 2. The QW width dependence of the intersubband scattering rates $W_{if}$ between $i$- and $f$-subbands for electrons with kinetic energy $E_i = 6.6$ meV.

$$W_{if\nu}^B = F_{if\nu}^\pm \int_0^{2\pi} |G_{B\nu}(z)|^2 \frac{d\theta}{2q_{0\nu}}, \quad (12)$$

where

$$G_{B\nu}(z) = \int_{L_e} \varphi_{ei} \varphi_{ef} \varphi_{q\nu}^B dz. \quad (13)$$

The frequency of the confined phonons is equal to the longitudinal PO phonon frequency in a bulk GaAs $\hbar\omega_0 = 36.2$ meV. The frequencies of the heterojunction interface phonons at $qL>2$ are $\hbar\omega_\nu = 33.1$, 35.4, and 45.9 meV. The frequencies of the interface phonons at the AlAs barrier at $qd \ll 1$, (where $d$ is the barrier thickness) are $\hbar\omega_\nu = 36.2$ and 50.1 meV.

The electron–phonon scattering rates depend significantly on a QW width. The main peculiarity of this dependence is a very sharp change of the intersubband scattering rate at the QW width at which the electron intersubband energy is equal to the phonon energy. In this case the emitted (absorbed) phonon wave vector $\mathbf{q}_0 \to 0$ [see Eq. (5)] and the scattering rates proportional to $q_0^{-1}$ increase sharply. Figure 2 shows the QW width dependence of the intersubband scattering rate (which includes scattering by all phonon modes).

## 2. PHOTOEXCITED ELECTRON SUBBAND POPULATION IN A QUANTUM WELL

Let us calculate the intersubband distribution of electrons photoexcited from the valence band to the third subband of the conduction band. The balance equations for nonequilibrium photoelectron subband population $n_i$ in three subband levels are

$$\alpha_3 I_g - W_{31} n_3 - W_{32} n_3 + W_{13} n_1 + W_{23} n_2 = W_{03} n_3 \equiv I_3,$$

$$\alpha_2 I_g + W_{32} n_3 - W_{21} n_2 - W_{23} n_2 + W_{12} n_1 = W_{02} n_2 \equiv I_2,$$

$$\alpha_1 I_g + W_{31} n_3 + W_{21} n_2 - W_{13} n_1 - W_{12} n_1 = W_{01} n_1 \equiv I_1, \quad (14)$$

where $a_i I_g$ is the flow of photoexcited electrons down to the bottom of the $i$-subband, which is proportional to the inten-
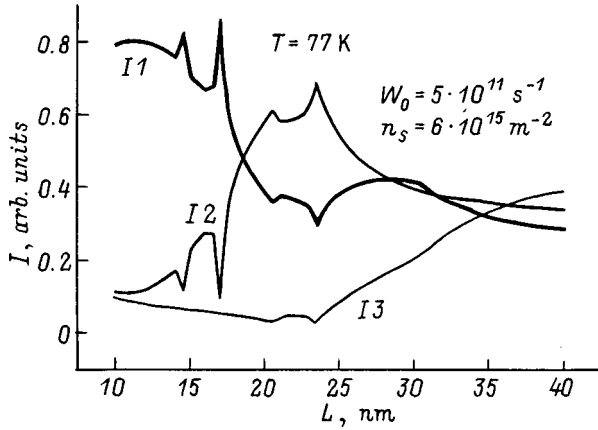
FIG. 3. The well width dependence of PL peaks $I_i$ corresponding to the subbands $i$ in the $Al_{0.3}Ga_{0.7}As/GaAs/Al_{0.3}Ga_{0.7}As$ quantum well.



FIG. 4. The dependence of electron wave function squared $|\varphi_{ei}|^2$ on the coordinate $z$ across the QW in the three lower subbands (1,2,3) in a GaAs QW of width $L = 20$ nm with 1-nm AlAs barrier inserted at $L_b = 7.5$ nm.

sity $I_g$ of photoluminescence (PL) excitation source and to the probability $\alpha_i \sim \langle [1 - f(E)] \rangle$ that the subband's bottom electron state is not occupied,

$$W_{if} = \langle W_{if}(E)[1 - f(E \pm \hbar\omega)] \rangle \qquad (15)$$

is the intersubband transition rate, $W_{0i}$ is the exciton formation rate (we assume that the radiative transitions through the electron-hole exciton states are dominant), and $I_i$ are the PL peak intensities which are proportional to $i$-subband photoexcited electron population. The brackets $\langle \rangle$ means the average value:

$$\langle A \rangle = \int A f(E) dE \bigg/ \int f(E) dE, \qquad (16)$$

where $f(E)$ is the Fermi-Dirac distribution function. We normalize the PL peak intensities: $I_1 + I_2 + I_3 = 1$, and assume $I_g = 1$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

Figure 3 shows the calculated QW width dependence of PL peak intensity in the $Al_{0.3}Ga_{0.7}As/GaAs/Al_{0.3}Ga_{0.7}As$ structure. We see that at $17 < L < 26$ nm the inversion of photoexcited electron population in the second and first subbands takes place. At these QW widths the subband energy separation between the third and the second subbands is larger than the confined and interface phonon energies, and between the second and the first subbands is less than these energies. As a result, $W_{32} \gg W_{21}$. These are the main criteria for achieving the population inversion in the QW. The inversion in PL peak intensities ($I_2 > I_1$) has been observed experimentally.[15]

## 3. ELECTRON IN ASYMMETRIC COUPLED QUANTUM WELLS. PHOTOVOLTAIC EFFECT

The insertion of a thin AlAs barrier into the GaAs quantum well (QW) deforms the electron wave function $\varphi_e$. Figure 4 shows the variation of $|\varphi_e|^2$ across the QW of width $L = 20$ nm for the three lower electron subbands when the barrier is inserted. We see that the distribution of the electron charge across the QW in the subbands is nonsymmetric. The deformation of the electron wave function causes the voltage across the QW to appear.
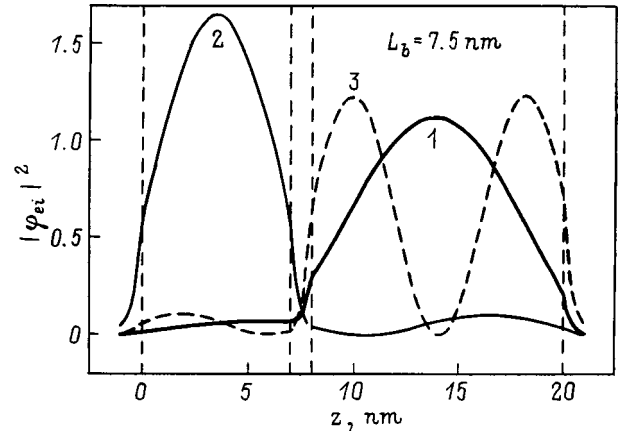
The photovoltaic potential arises across the asymmetric coupled QW when the electrons are photoexcited from the first subband to the second subband due to different electron charge deformations in these subbands

$$V_{p\nu} = \frac{e}{\varepsilon_1} \Delta n \int_{L_1}^{L_2} \left[ \int_{L_1}^{z} (|\varphi_{e1}(z')|^2 - |\varphi_{e2}(z')|^2) dz' \right] dz. \qquad (17)$$

$L_2 - L_1$ is the area, where $|\varphi_{ei}|^2 \neq 0$.

The calculated photovoltaic voltage for the case where the AlAs barrier is inserted in the QW at $L_b = 7.5$ nm, and where electrons with the concentration $\Delta n = 8 \times 10^{14}$ m$^{-2}$ are photoexcited from the first subband to the second one, is equal to 2.2 mV. The photovoltaic effect in asymmetric coupled QW's can be used for the detection of infrared radiation. The photovoltaic response to the radiation which excites the electrons from the first QW subband to the second one is

$$\frac{V_{p\nu}}{P} = \frac{e}{\varepsilon_1} \frac{W_{32}}{W_{21}(W_{31} + W_{32})h\nu_{13}}$$
$$\times \int_{L_1}^{L_2} \left[ \int_{L_1}^{z} (|\varphi_{e1}(z')|^2 - |\varphi_{e2}(z')|^2) dz' \right] dz, \qquad (18)$$

where $P$ and $h\nu_{13}$ are the power and the energy of the optical excitation signal. This signal is proportional to the $W_{21}^{-1}$ which can be large when $E_{s2} - E_{s1} < \hbar\omega_0$.

## 4. ELECTRON MOBILITY IN THE QUANTUM WELL

Let us estimate the contribution of PO phonon confinement in the QW to the electron mobility in the $Al_{0.3}Ga_{0.7}As/GaAs/Al_{0.3}Ga_{0.7}As$ QW with and without the inserted thin AlAs barrier. The electron mobility limited by PO phonon scattering we shall estimate in the relaxation time approximation. The relaxation time $\tau_r$ of the perturbed electron state $\mathbf{k}_i$ to the equilibrium state $\mathbf{k}_f$ (Ref. 16)

$$\frac{1}{\tau_r(\mathbf{k}_i)} = \sum_{\mathbf{k}_f} \{W_{if}[1-f(\mathbf{k}_f)] + W_{fi}f(\mathbf{k}_f)\}$$

$$= \sum_{\mathbf{k}_f} W_{if}\frac{1-f(\mathbf{k}_f)}{1-f(\mathbf{k}_i)} \tag{19}$$

is involved as momentum relaxation time. Here $f(\mathbf{k})$ is the Fermi-Dirac distribution function, and $W_{if}$ is the probability of electron transition from the state $i$ to the state $f$. In the parabolic subbands we have the following expression for electrons with energy $E$ in the subband $i$:

$$\frac{1}{\tau_{ri}(E)} = \sum_{\nu} \sum_{f} W_{if\nu}^{\pm} \frac{1-f(E \mp \hbar\omega_\nu)}{1-f(E)}, \tag{20}$$

where $\nu$ and $f$ are the numbers of phonon mode and electron subband, respectively. The upper (plus-minus) sign is for phonon emission and lower one is for phonon absorption. The $W_{if\nu}^{\pm}$ is determined for each phonon mode by Eq. (1). The mobility of the subband $i$ electrons is

$$\mu_i = \frac{e}{m}\left\langle \frac{1}{\tau_{ri}(E)}\right\rangle^{-1}. \tag{21}$$

The average mobility in the QW is

$$\mu = \sum_i \mu_i n_i / n_0, \tag{22}$$

where

$$n_i = D \int_{E_i}^{\infty} f(E)dE \tag{23}$$

is the electron concentration in the subband $i$, $D = m/\pi\hbar^2$, and $n_0 = \Sigma_i n_i$.

It is known that the relaxation time approximation gives only crude estimation of the mobility limited by PO phonon scattering, but this approximation is expected to be sufficient for the purpose in which we are interested: The relative difference between the mobilities in QW's with different widths and the relative contribution to electron mobility of various electron scattering mechanisms by various phonon modes.

Note that the values of the mobility calculated in the relaxation time approximation in bulk GaAs [$\mu = 50\,\mathrm{m^2/(V\cdot s)}$ at $T = 77\,\mathrm{K}$ and $\mu = 0.65\,\mathrm{m^2/(V\cdot s)}$ at $T = 293\,\mathrm{K}$] are close to the experimentally observed values.

Figure 5 shows the calculated well width dependences of the Fig. 5 electron mobility in the AlGaAs/GaAs/AlGaAs QW. The dependences are characterized by the superposition of the two types of minima: the smooth (sawtooth-like) minima and the sharp minima. The sharp minima correspond to the resonant increase of electron scattering rates by the interface phonons when the subband energy separation is equal to the heterojunction interface phonon energies. This situation takes place at $L = 14$ and $16$–$17\,\mathrm{nm}$. Disregard of scattering by the interface phonons with allowance for only confined phonon scattering shows that the smooth (sawtooth-like) dependences are determined by two competing effects. The first effect is the reduction in the strength of the electron-confined phonon interaction, which lead to the mobility enhancement in proportion to the increase of the QW
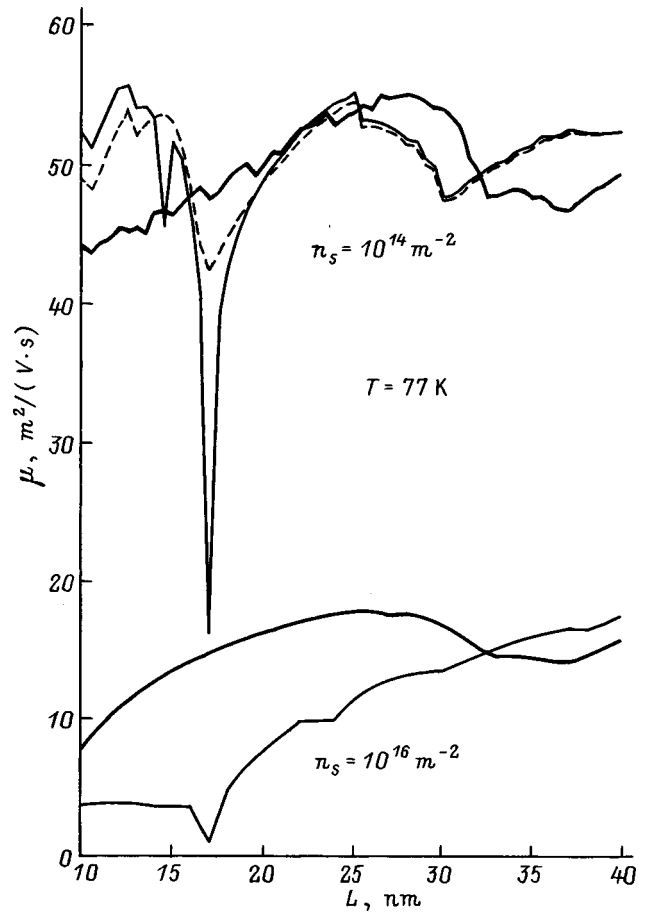


FIG. 5. The electron mobility as a function of QW width $L$ at $T = 77\,\mathrm{K}$ for $n_S = 10^{14}\,\mathrm{m^{-2}}$ (upper curves) and for $n_S = 10^{16}\,\mathrm{m^{-2}}$ (lower curves). The thick and thin solid lines correspond to electron–phonon scattering in the QW with and without the inserted barrier, respectively. The dashed lines correspond to the electron scattering by confined phonons only in the QW without the barrier.

width $L$ [see Eq. (6)]. The second one is the decrease of the electron mobility at QW widths when the intersubband scattering by absorption of the confined phonon from the lower subband to the second subband (at $L = 16\,\mathrm{nm}$) and to the third subband (at $L = 30\,\mathrm{nm}$) becomes energetically possible ($\Delta E_{if} \leqslant \hbar\omega_0$).

At low electron concentration ($n_S = 10^{14}\,\mathrm{m^{-2}}$) the mobility in the QW of width $L = 10$–$15\,\mathrm{nm}$, $L = 20$–$30\,\mathrm{nm}$, and $L > 35\,\mathrm{nm}$ is slightly larger than in a bulk material. The increase of electron concentration ($n_S = 10^{16}\,\mathrm{m^{-2}}$) in the QW decreases drastically the electron mobility which is limited by PO phonon scattering (see Fig. 5). The rise of the electron energy in the lower subband due to the electron degeneration is responsible for the strong enhancement of the intersubband scattering rate by the confined phonons, and, consequently, for the decrease of the electron mobility.

We see that the confinement of electrons and phonons in the AlGaAs/GaAs/AlGaAs HEMT channels does not enhance the electron mobility limited by PO phonon scattering. Moreover, the increase of the carrier concentration in the channel decreases significantly the electron mobility. Note that mobility measured experimentally in the AlGaAs/GaAs/AlGaAs QW's is lower than in a bulk GaAs. In the QW of

width $L=15-20$ nm at $T=80$ K $\mu=4$ m$^2$/(V·s) at $n_S$ $=10^{16}$ m$^{-2}$, (Ref. 7), and $\mu=6$ m$^2$/(V·s) at $n_S=10^{15}$ m$^{-2}$ (Ref. 6). These experimental data are close to the calculated mobilities (see Fig. 5). This mobility decrease can be compensated for by inserting a thin AlAs barrier into the QW, which is transparent to electrons but reflects PO phonons. We call this barrier a phonon wall.[17]

The insertion of the barrier into the center of the QW changes the electron subband energy spectra. The subband energies are grouped into pairs with increasing separation energy between the pairs. This separation energy is comparable with the phonon energy only at $L\geqslant30$ nm. As a result, the sharp resonant peaks of the mobility disappear and the positions of the wide mobility minimum shift to large QW widths. This is shown in Fig. 5. The particularly large increase of the mobility is observed in the case of high electron sheet concentration ($n_S=10^{16}$ m$^{-2}$). In the case of low electron sheet concentration ($n_S=10^{14}$ m$^{-2}$) the insertion of the barrier increases the electron mobility in the QW in the interval of widths $16<L<31$ nm. Note that in this interval the increase of electron mobility was observed experimentally.[6]

## CONCLUSIONS

It is shown that the polar optical phonon confinement in QW's strongly affects the electron–phonon scattering strength. The resonant increase of the intersubband electron–phonon scattering rate takes place when the energy of the interface phonons is equal to the subband energy separation in the QW. At QW widths which correspond to the resonance conditions the electron subband population and mobility change sharply.

The inversion of photoexcited electron subband population and electron mobility enhancement in the QW can be achieved at appropriate parameter values for which intersubband scattering is reduced.

The increase of electron–phonon scattering rate when the enhancement of electron concentration in the QW takes place is responsible for the large decrease of electron mobility in the lower subbands. This decrease can be partly compensated for inserting a thin AlAs barrier into the GaAs QW.

The insertion of a thin barrier into the QW deforms the electron wave function differently in various subbands, and the photovoltaic effect occurs when the intersubband electron photoexcitation takes place.

[*]E-mail: pozela@uj.pfi.lt

[1] F. H. Julien, A. Sa'ar, J. Wang, and J.-P. Leburton, Electron. Lett. **31**, 838 (1995).

[2] P. Boucaud, F. H. Julien, D. D. Yang, J. M. Lourtioz, E. Rosencher, P. Bois, and J. Nagle, Appl. Phys. Lett. **57**, 3 (1990).

[3] J. Faist, F. Capasso, D. Sivco, C. Sirtori, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho, Science **264**, 553 (1994).

[4] B. F. Levine, J. Appl. Phys. **74**, R1 (1993).

[5] T. Tsuchiya and T. Ando, Phys. Rev. B **48**, 4599 (1993).

[6] X. T. Zhu, H. Goronkin, G. N. Maracas, R. Droopad, and M. A. Stroscio, Appl. Phys. Lett. **60**, 2141 (1992).

[7] I. Inoue and T. Matsuno, Phys. Rev. B **47**, 3771 (1993).

[8] K. Yokoyama and K. Hess, Phys. Rev. B **33**, 5595 (1986).

[9] N. Mori and T. Ando, Phys. Rev. B **40**, 6175 (1989).

[10] H. Rücker, E. Molinary, and P. Lugli, Phys. Rev. B **45**, 6747 (1992).

[11] I. Lee, S. M. Goodnick, M. Gulia, E. Molinaty, and P. Lugli, Phys. Rev. B **51**, 7046 (1995).

[12] B. K. Ridley, Phys. Rev. B **39**, 5282 (1989).

[13] J. Požela, V. Juncienė, and K. Požela, Semicond. Sci. Technol. **10**, 1555 (1995).

[14] X. Zianni, C. D. Simserides, and G. P. Triberis, Phys. Rev. B **55**, 16 324 (1997).

[15] J. Požela, V. Jucienė, A. Namajūnas, K. Požela, V. G. Mokerov, Yu. V. Fedorov, V. E. Kaminskii, and A. V. Hook, J. Appl. Phys. **82**, 5564 (1997).

[16] V. F. Gantmakker and I. B. Levinson, *Scattering of current carriers in metals and semiconductors*, Nauka, Moscow (1984), p. 50.

[17] J. Požela, V. Jucienė, A. Namajūnas, and K. Požžela, J. Appl. Phys. **81**, 1775 (1997).

# Physics and technologies of superhigh-efficiency tandem solar cells

M. Yamaguchi*

*Toyota Technological Institute, 2-12-1 Hisakata, Tempaku, Nagoya 468-8511, Japan*

The present status of superhigh-efficiency tandem solar cells has been reviewed and the key issues for realizing superhigh-efficiency have been discussed. The mechanical, stacked, three-junction cells of monolithically grown InGaP/GaAs two-junction cells and InGaAs cells have reached the highest efficiency (attainable in Japan) of 33.3% at 1-sun AM 1.5. Future prospects for realizing superhigh-efficiency and low-cost tandem solar cells are discussed. © *1999 American Institute of Physics.* [S1063-7826(99)00709-7]

## 1. INTRODUCTION

Substantial increase in conversion efficiency (efficiency values over 30%) can be realized by multijunction (tandem) solar cells due to their wide-band photoresponse in comparison with single-junction cells. Tandem solar cells have been studied since 1960 (Ref. 1). Although AlGaAs/GaAs tandem cells, including tunnel junctions and metal connectors, were developed in the early years, high-efficiency values close to 20% were not obtained,[2] because of difficulties in making stable, high-performance tunnel junctions, and because of the defects associated with oxygen in AlGaAs.[3] A double heterostructure (DH structure) tunnel junction was found to be useful for preventing diffusion from the tunnel junction and improving the tunnel junction performance.[4] InGaP as a material for the top cell was proposed by NREL.[5] As a result of performance improvements in the tunnel junction and the top cell, over 30% efficiency has been obtained with InGaP/GaAs tandem cells.[6] InGaP/GaAs two-junction solar cells have recently drawn increased attention for space applications, because of the possibility of high conversion efficiency (over 30%). In fact, the first commercial satellite (HS 601HP) with a two-junction GaInP/GaAs on Ge solar arrays was launched in August 1997 (Ref. 7).

In this paper we review the present status and future prospects of superhigh-efficiency tandem solar cells and discuss the key issues for realizing superhigh efficiency.

## 2. KEY ISSUES FOR REALIZING HIGH-EFFICIENCY TANDEM CELLS

Key issues for realizing high-efficiency tandem cells are discussed on the basis of our results.

Selection of top cell materials is also important for high-efficiency tandem cells. As a top cell material lattice-matched to GaAs or Ge substrates, InGaP has some advantages such as lower interface recombination rate, minor oxygen problem, and good window layer material compared to AlGaAs, as shown in Table I.

The top cell characteristics depend on the minority carrier lifetime in the top cell layers. Figure 1 shows changes in photoluminescence intensity ($I_{PL}$) of the solar cell active layer as a function of minority carrier lifetime ($\tau$) of the

$p$-InGaP base layer grown by MOCVD and surface recombination rate ($S$). The lowest surface recombination rate was obtained by introducing an AlInP window layer and the highest minority carrier lifetime was obtained by introducing a buffer layer and optimizing the growth temperature. The best conversion efficiency of the InGaP single-junction cell was 18.5%. Another important goal for realizing high-efficiency monolithic-cascade-type tandem cells is achievement by establishing optically and electrically low-loss interconnection of two or more cells. A degenerately doped tunnel junction is attractive because it only involves one extra step in the growth process. To minimize optical absorption, formation of thin and wide-bandgap tunnel junctions is necessary, as shown in Fig. 2. However, the formation of a wide-bandgap tunnel junction is very difficult, because the tunnel peak current density ($J_t$) decreases exponentially with increasing bandgap energy ($E_g$).

In addition, impurity diffusion from a highly doped tunnel junction during overgrowth of the top cell increases the resistivity of the tunnel junction. A DH structure was found to be useful for preventing diffusion.[4] More recently, an InGaP tunnel junction has been tried for the first time for an InGaP/GaAs tandem cell.[6] As $p$-type and $n$-type dopants, Zn and Si, respectively, were used. Peak tunneling current of the InGaP tunnel junction is found to increase from 5 to 300 mA/cm$^2$ by making the DH structure with the AlInP barriers. Furthermore, higher tunneling current up to 2 A/cm$^2$ has been obtained by increasing the doping density in the junction. Therefore, the InGaP tunnel junction was found to be very effective for high tunneling current, and DH structure

TABLE I. Comparison of top cell materials (InGaP and AlGaAs).

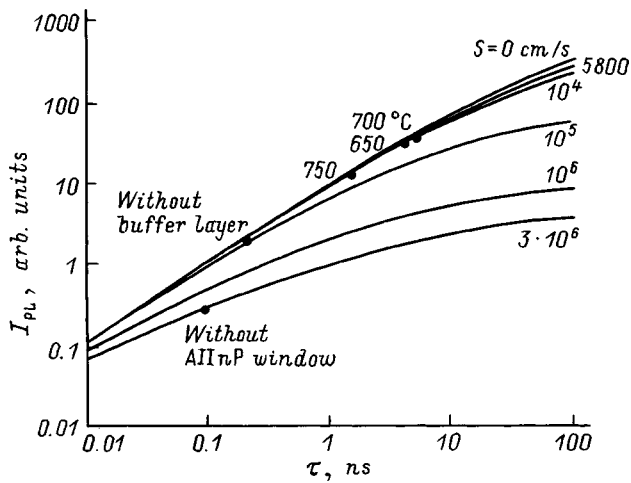| Items | InGaP | AlGaAs |
|---|---|---|
| Interface recombination rate | $<5\times10^3$ cm/s | $10^4$–$10^5$ cm/s |
| Oxygen-related defects | Minor problem | Major problem |
| Window layer | AlInP | AlGaAs |
| ($E_g$) | (2.5 eV) | (2.1 eV) |
| Other problems | High doping in $p$-AlInP | Lower efficiency (2.6% down) |

FIG. 1. Changes in photoluminescence intensity ($I_{PL}$) of the solar cell active layer as a function of the minority carrier lifetime ($\tau$) of the $p$-InGaP base layer and surface recombination rate ($S$). Values of the growth temperature and the surface recombination rate are shown.
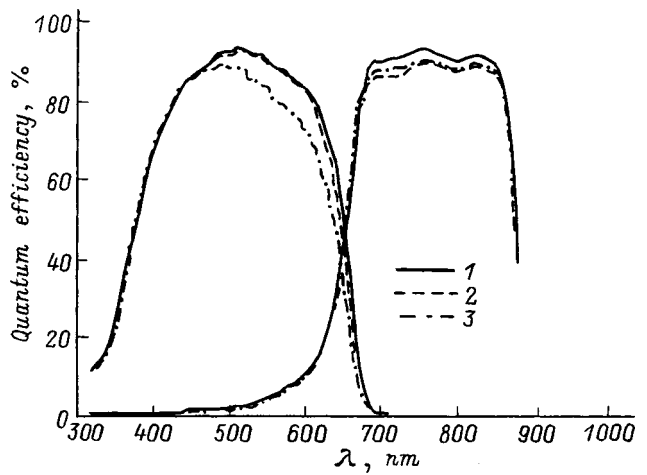


FIG. 3. Spectral response of the InGaP/GaAs tandem cells which depend on their tunnel junction structures: *1*—AlInP/InGaP, *2*—AlInP/GaAs, *3*—InGaP/GaAs. All the tunnel junctions have DH structure.

was also found to be useful for preventing diffusion, as was indicated elsewhere.[4]

The impurity diffusion from a highly doped tunnel junction also degrades the top cell performance. The dependence of the tandem cell characteristics on the tunnel junction structure, such as InGaP/GaAs, AlInP/GaAs, and AlInP/InGaP DH structures, has been compared by using spectral responses of the tandem cells, as shown in Fig. 3. The upper barrier layer of the tunnel junction also takes the part of the back surface field (BSF) of the top cell. A large reduction in quantum efficiency at wavelengths between $\lambda = 500$ and 650 nm due to the diffusion of Zn during epitaxial growth has been observed in a InGaP/GaAs tunnel junction. On the other hand, an increase in the quantum efficiency of the bottom GaAs cell due to the elimination of absorption losses in the GaAs tunnel junction has been confirmed in the cell with an InGaP tunnel junction. In the case of AlInP/InGaAs DH structures tunnel junction, the AlInP upper barrier layer of the tunnel junction was found not only to suppress the Zn diffusion, but also to produce the effective BSF due to the band discontinuity with InGaP.
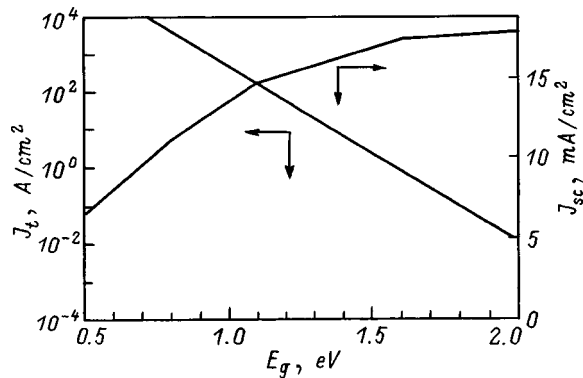
DH structure effect on suppression of impurity diffusion from the tunnel junction has been examined. Figure 4 shows formation and migration enthalpy of group III-vacancies versus bond-gap energy of the materials. Effective suppression of Zn diffusion from tunnel junction by the InGaP tunnel junction with the AlInP DH structures is thought to be attributable to the lower diffusion coefficient for Zn in the wider bandgap energy materials such as the AlInP barrier layer and InGaP tunnel junction layer.

## 3. SUPERHIGH-EFFICIENCY TANDEM CELLS

Figure 5 shows a schematic cross section of the high-efficiency InGaP/GaAs two-junction cell. The InGaP/GaAs cell layers were grown on a GaAs substrate by the MOCVD method. The top and bottom cells were connected by a InGaP tunnel junction. Figure 6 shows the light current–voltage ($I$–$V$) characteristic of the high-efficiency InGaP/



FIG. 2. Calculated tunnel peak current density ($J_t$) and short-circuit current density of GaAs bottom cell ($J_{sc}$) as a function of bandgap energy ($E_g$) of the tunnel junction.
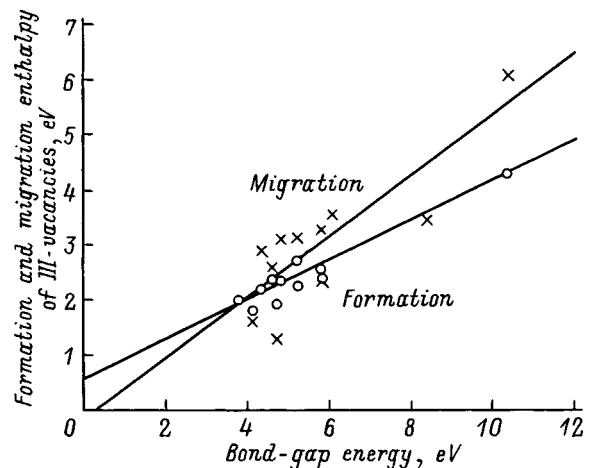


FIG. 4. Changes in formation and migration energies of III-group element vacancies as a function of band-gap energy of the III–V compound materials.
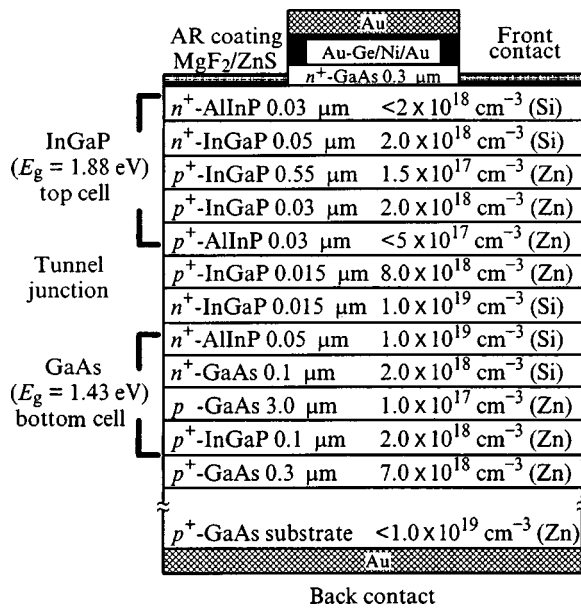
FIG. 5. A schematic cross section of the high-efficiency InGaP/GaAs two-junction cell.

FIG. 6. Light current–voltage and power–voltage curves of the high-efficiency InGaP/GaAs tandem cell measured at the Japan Quality Assurance Organization: cell area 4 cm$^2$, AM 1.5, 100 mW/cm$^2$, 25.3 °C, short-circuit current $I_{sc} = 56.88$ mA, open-circuit voltage $V_{oc} = 2.488$ V, fill-factor $FF = 85.6\%$, efficiency 30.3%.

GaAs tandem cell. The high efficiency was realized by introducing DH structures in the InGaP top cell, InGaP tunnel junction, and GaAs bottom cell.

More recently, monolithically grown $3 \times 3$-cm$^2$ InGaP/GaAs two-junction solar cells with 1-sun AM 1.5 efficiency of 30.6% have been successfully fabricated by improving InGaP top cell and GaAs bottom cell properties as a result of improvement of the epitaxial growth and introduction of the C-doped AlGaAs/Si-doped InGaP heterostructure tunnel junction with AlInP barrier by Japan Energy Co. This value is the highest ever reported for the two-junction cells under 1-sun illumination. The mechanically stacked three-junction cells of monolithically grown InGaP/GaAs two-junction cells and InGaAs bottom cells have reached highest effi-
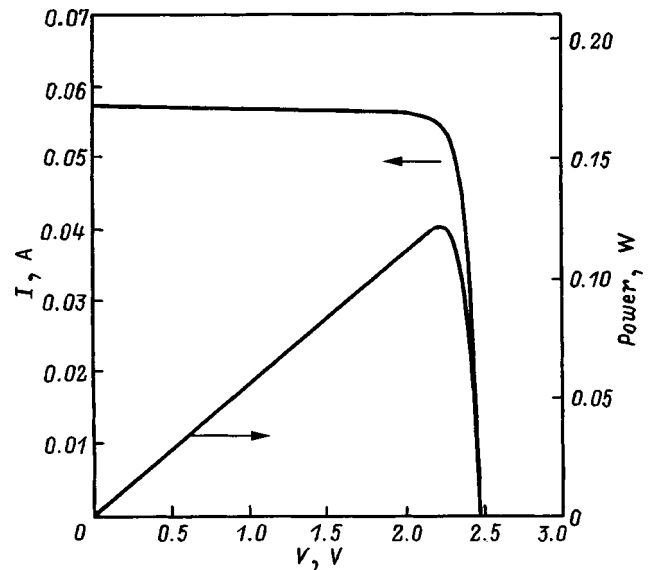
ciency of 33.3% (Ref. 8) at 1-sun AM 1.5 in the joint work by Japan Energy Co. and Sumitomo Electric Co.

Table II summarizes the research activities of III–V compound solar cells in Japan.

## 4. FUTURE PROSPECTS FOR OBTAINING 40% EFFICIENCY

Some effort has been made to put this type of cells into commercial production for space applications based on the multijunction solar cell manufacturing technology program.[9] In fact, the first commercial satellite with two-junction

TABLE II. Summary of research activities on the III–V compound solar cells in Japan.

| Solar cells | | Area, cm$^2$ | AM | Efficiency,% | Organization | Year |
|---|---|---|---|---|---|---|
| | GaAs bulk | 0.25 | AM 1.5 | 25.4 | Hitachi Cable | 1996 |
| Bulk | | 4 | AM 0 | 22.5 | Mitsubishi Electric | 1987 |
| | InP bulk | 0.25 | AM 1.5 | 21.4 | NTT | 1986 |
| Thin-film | GaAs on Ge | 1 | AM 1.5 | 23.2 | Hitachi Cable | 1996 |
| | GaAs on Si | 1 | AM 1.5 | 20.0 | NTT | 1989 |
| | InGaP/GaAs 2-junction | 4 | AM 1.5 | 30.3 | Japan Energy | 1996 |
| | | 9 | AM 1.5 | 30.6 | Japan Energy | 1998 |
| | | 4 | AM 0 | 26.9 | Japan Energy Toyota Tech. Inst. | 1997 |
| Tandem | AlGaAs/GaAs 2-junction | 0.25 | AM 1.5 | 20.2 | NTT | 1987 |
| | | 1 | AM 1.5 | 28.8 | Hitachi Cable | 1996 |
| | GaAs/InGaAs Mechanically stacked | 1 | AM 1.5 | 28.8 | Sumitomo Electric | 1996 |
| | AlGaAs/Si 2-junction | 0.25 | AM 0 | 21.2* | Nagoya Inst. Tech. | 1996 |
| | InGaP/GaAs/InGaAs MS 3-junction | 1 | AM 1.5 | 33.3 | Japan Energy Sumitomo Electric | 1997 |
| Concentrator tandem | InGaP/GaAs 2-junction | 1 | AM 1.5 (×5.1) | 31.2 | Toyota Tech. Inst. Japan Energy | 1998 |

*Active-area efficiency.
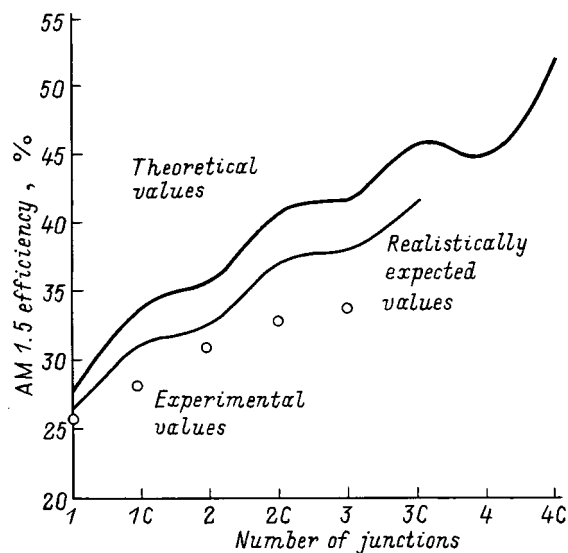
FIG. 7. Theoretical and realistically expected conversion efficiencies of single-junction and multijunction solar cells in comparison with experimentally realized efficiencies. *C* indicates concentration.

GaInP/GaAs on Ge solar arrays was launched in August 1997 (Ref. 7). Therefore, tandem solar cells will be widely used in space.

In order to apply superhigh-efficiency cells widely, it is necessary to improve their conversion efficiency and reduce their cost. Figure 7 shows theoretical and realistically expected conversion efficiencies of single-junction and multijunction solar cells in comparison with experimentally realized efficiencies. Therefore, concentrator three-junction and four-junction solar cells have great potential for realizing superhigh efficiency over 40%. As a three-junction combination, GaInP/GaAs/Ge cell on Ge substrate will be widely used because this system has been already developed. The four-junction combination of the AlGaInP ($E_g = 2.0$ eV) top cell, the GaAs second-layer cell, the third-layer cell made of a material with 1.05-eV bandgap, for example, GaInAsP or GaInAsN, and the Ge bottom cell is lattice-matched to Ge substrates and has theoretical 1-sun AM 0 efficiency of about 42%. This system also has a potential of over 45% under 500-suns AM 1.5 condition. Although this system is ideal for maximum theoretical efficiency, the selection of third-layer cell materials and improvement in the material quality are problems to overcome.

The use of three-junction cells on Si or metal substrates is a way of realizing low cost. One of future targets must be also three-junction cells on metal substrates. Concentrator thin-film multijunction solar cells fabricated on inexpensive substrates such as Si and polycrystalline materials have great potential for realizing high-efficiency (with efficiency more than 35%) and low-cost cells if the dislocation density can be reduced to less than $5 \times 10^5$ cm$^{-2}$ and if the grain size can be increased to more than 0.1 cm (Ref. 10).

## 5. SUMMARY

The present status of superhigh-efficiency tandem solar cells was reviewed. InGaP/GaAs tandem solar cells with newly recorded efficiency of 30.6% at AM 1.5 (1-sun) were obtained by improvements in InGaP top cells and AlGaAs/InGaP DH structure tunnel junctions with AlInP barriers. The mechanical stacked three-junction cells of monolithically grown InGaP/GaAs two-junction cells and InGaAs cells reached the highest efficiency (achievable in Japan) of 33.3% at 1-sun AM 1.5.

Key technologies and basic physics for realizing superhigh-efficiency and low-cost multijunction solar cells were also discussed.

*E-mail: masafumi@toyota-ti.ac.jp

[1] M. Wolf, Proc. IRE **48**, 1246 (1960).
[2] J. A. Hutchby, R. J. Markunas, and S. M. Bedair, *Photovoltaics*, edited by S. K. Ded [Proc. SPIE **543**, 543 (1985)].
[3] K. Ando, C. Amano, H. Sugiura, M. Yamaguchi, and A. Salates, Jpn. J. Appl. Phys. **26**, L266 (1987).
[4] C. Amano, H. Sugiura, A. Yamamoto, and M. Yamaguchi, Appl. Phys. Lett. **51**, 1998 (1987).
[5] J. M. Olson, S. R. Kurtz, and K. E. Kibbler, Appl. Phys. Lett. **56**, 623 (1990).
[6] T. Takamoto, E. Ikeda, H. Kurita, M. Ohmori, and M. Yamaguchi, Jpn. J. Appl. Phys. **36**, 6215 (1997).
[7] M. R. Brown, L. J. Goldhammer, G. S. Goodelle, C. U. Lortz, J. N. Perron, J. S. Powe, J. A. Schwartz, B. T. Cavicchi, M. S. Gillanders, and D. D. Krut, in *Proceedings of the 26th IEEE Photovoltaic Specialists Conference* (IEEE, N.Y., 1997) p. 805.
[8] T. Takamoto, E. Ikeda, T. Agui, H. Kurita, T. Tanabe, S. Tanaka, H. Matsubara, Y. Mine, S. Takagishi, and M. Yamaguchi, in *Proceedings of the 26th IEEE Photovoltaic Specialists Conference* (IEEE, N.Y., 1997), p. 1031.
[9] D. N. Keener, D. C. Marvin, D. J. Brinker, H. B. Curtis, and P. M. Price, in *Proceedings of the 26th IEEE Photovoltaic Specialists Conference* (IEEE, N.Y., 1997) p. 787.
[10] M. Yamaguchi and C. Amano, J. Appl. Phys. **58**, 3601 (1985).

# Growth of III–N materials and devices by metalorganic chemical vapor deposition

R. D. Dupuis,* P. A. Grudowski, C. J. Eiting, and J. Park

*The University of Texas at Austin, Microelectronics Research Center PRC/MER 1606D-R9900,
Austin, TX 78712-1100, USA*

C. J. Eiting

*The Air Force Research Lab., Materials & Manufacturing Directorate, WPAFB, OH 45433-7707, USA*

The characteristics of III–V nitride semiconductor epitaxial layers grown by metalorganic
chemical vapor deposition are of interest for the realization of many technologically important
devices. This paper will review heteroepitaxial growth on (0001) sapphire substrates as
well as the selective-area and subsequent lateral epitaxial overgrowth on masked substrate surfaces.
© *1999 American Institute of Physics.* [S1063-7826(99)00809-1]

## 1. INTRODUCTION TO SEMICONDUCTING NITRIDES

The growth of high-quality III–V nitride semiconductor epitaxial films is a recent development dating from about 1989 with the work of Akasaki *et al.*[1] Because of its large direct bandgap extending from ultraviolet (UV) (6 eV) through the entire visible spectrum to red (2 eV), the indium aluminum gallium nitride (InAlGaN) system has great potential for many different optoelectronic and electronic devices. Already, extensive work has been carried out on visible light emitting diodes (LED's),[2,3] injection laser diodes (LD's),[4–9] UV photodetectors,[10,11] and field-effect transistors.[12]

Compound semiconductors in the nitride material system can be grown in both the cubic (zinc blende) and hexagonal (wurtzite) crystalline forms, although the wurtzite form is the preferred metastable state of InAlGaN alloys. The source of many of the problems associated with the growth of high-quality nitride films is the lack of a lattice-matched substrate. While commercially viable substrates such as cubic Si and GaAs have been used with limited success, the substrates of choice have been hexagonal **c**-plane (0001) sapphire ($Al_2O_3$) and the hexagonal form of silicon carbide ($6H$-SiC). Recently, alternative approaches to heteroepitaxial growth using selective-area growth have been developed for high-quality nitride films. We will review recent progress in this field.

## 2. MOCVD GROWTH OF III–N FILMS

The largest advance in the research on nitride semiconductors was the development of the buffer layer growth prior to deposition of the single crystal GaN. Yoshida *et al.*, using molecular-beam epitaxy (MBE), first discovered this ''two-step'' growth process.[13] Akasaki *et al.*, employed this idea using metalorganic chemical vapor deposition (MOCVD) and extensively studied the effect of a thin (50–100 nm) low-temperature (600 °C) intermediate AlN layer when growing GaN on sapphire.[1] This group showed from X-ray diffraction studies that the mosaicity, and therefore the defect structure related to the degree of mosaicity, of the GaN epi-

taxial layers was reduced by using the AlN buffer layer and proposed a model in which the initial deposition results in small AlN crystallites that combine to form an amorphous layer. As GaN is deposited on this amorphous buffer layer, three-dimensional islands begin to form, but eventually coalesce to form a smooth, continuous layer. The coalescence is enhanced by a reduction of the interfacial free energy between GaN and AlN. The highly defective buffer layer tended to ''absorb'' a large fraction of the strain and reduced the number of dislocations extending to the surface of the epitaxial layer grown on top of it. Deposition without the buffer layer results in large three dimensional islands which eventually join and form heavily boundary regions throughout the layer.

Nakamura adapted this process to a two-step growth technique that does not require the use of AlN. Using atmospheric-pressure MOCVD, he was able to obtain the same beneficial effects of an AlN buffer layer by using GaN low-temperature layer. In fact, with 20-nm GaN buffer layer, he achieved higher electron mobilities, lower background electron concentrations, and lower x-ray FWHM values than for GaN grown on an AlN buffer layer.[1]

All of the materials discussed in this work were grown at the University of Texas at Austin by metalorganic chemical vapor deposition in the EMCORE D125 vertical geometry, rotating disk MOCVD reactor. The system is capable of using six hydride sources for group V and dopant precursors and ten metalorganic sources for group-III and dopant precursors. The hydrides are ammonia for the group-V nitrogen source and the mixture of 50 ppm silane ($SiH_4$) in hydrogen ($H_2$) for the silicon dopant source. The group-III metalorganics used are trimethylgallium (TMGa), triethylgallium (TEGa), trimethylaluminum (TMAl), and trimethylindium (TMIn). The dopant source is bis(cyclopentadienyl)magnesium ($Cp_2Mg$).

### 2.1. Si-doped and Mg-doped GaN

To date, the best doping success in GaN has been achieved with silicon (*n*-type) and magnesium (*p*-type) dop-
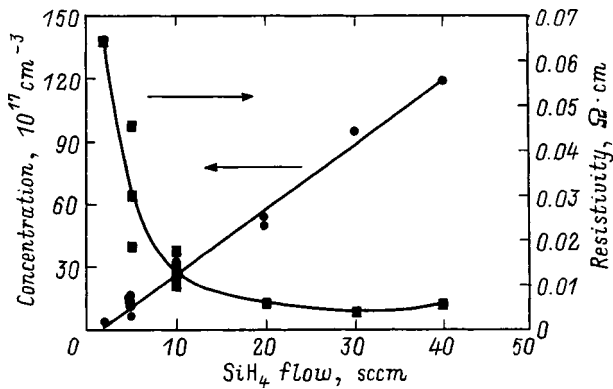
FIG. 1. Room-temperature Hall-effect data for Si-doped GaN films showing the relationship between electron carrier concentration and film resistivity versus the input SiH$_4$ mixture bulk flow.

ant atoms. We have achieved *n*-type silicon doping with electron carrier concentrations ranging from $n = 5 \times 10^{17}$ to $1 \times 10^{19}$ cm$^{-3}$ and maximum electron mobility $\sim 300$ cm$^2$/(V·s) for $n = 2 \times 10^{18}$ cm$^{-3}$. Figure 1 shows the relationship between the electron carrier concentration and resistivity, as determined by room-temperature Hall effect measurements, and the SiH$_4$ molar flow. The carrier concentration is very linear with dopant molar flow, indicative of constant dopant atom activation even at higher flows. The resistivity, however, obeys the nonlinear relationship at higher flows due to decrease in the electron mobility. The mobility decrease with higher carrier concentration is an expected effect due to the higher degree of impurity scattering.

*P*-type doping with magnesium has also been demonstrated; however, the 300-K free-hole carrier concentrations for GaN:Mg films range from $p \simeq 7 \times 10^{15}$ to $2 \times 10^{18}$ cm$^{-3}$ with maximum hole mobility of 13 cm$^2$/(V·s). These samples were obtained by optimizing the Cp$_2$Mg molar flow and thermally activating Mg atoms with a post growth *in situ* (growth chamber) or *ex situ* (rapid thermal annealing furnace) N$_2$ ambient activation anneal at 700–1000 °C for 10–30 min. Figure 2 shows the 300-K Hall effect results for the hole carrier concentration and resistivity versus Cp$_2$Mg molar flow.
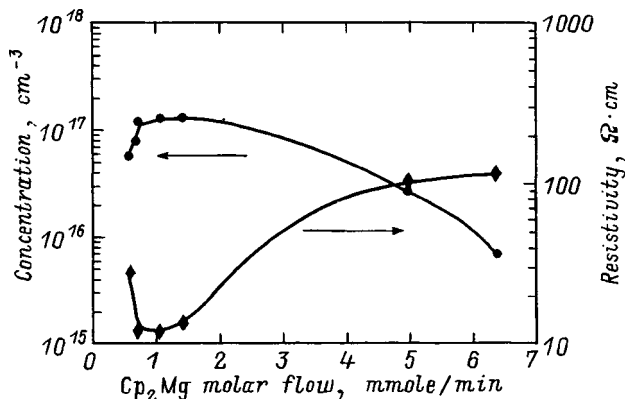
## 2.2. Growth of multiple quantum-well heterostructures

After achieving the growth of *n*-type and *p*-type materials, we studied the growth of the AlGaN and InGaN epitaxial layers and heterojunctions required for high-efficiency LED's and injection lasers. These structures were designed to become the light-confining and light-emitting component of the device structure. An important issue we investigated was the structure of the multiple quantum well (MQW) active region. Multiple quantum wells are chosen over a single quantum well to increase the luminescence intensity of the structures. Obviously, increase of the number of emitting layers will increase the total number of electron-hole pairs and, therefore, the number of emitted photons. The tradeoff in the nitride material system involves the excess strain induced by heteroepitaxial growth. The higher In content in InGaN and the thicker the overall active region, the more strain exists. Finally, if the critical layer thickness is approached, relaxation of the layers occurs and defect generation is enhanced. All of the MQW regions reported here have the total thickness, which should be much below the critical layer thickness of InGaN alloys. The effect of varying the number of periods in the MQW structure consisting of 35-Å In$_{0.13}$Ga$_{0.87}$N quantum wells and 70-Å In$_{0.03}$Ga$_{0.97}$N barrier was investigated. We determined that the optimum structure was a 5-period MQW for this well/barrier combination.

With this MQW optimized in terms of the Si doping level and the number of periods, AlGaN cladding layers were added to the photoluminescence (PL) structure to examine the effect of increased overall strain in the samples. Laser-diode structures require some type of AlGaN cladding layers for effective carrier and optical confinement, because of the band offsets and refractive index steps between GaN, InGaN, and AlGaN. However, adding AlGaN thick layers with low Al content will increase the amount of heteroepitaxial strain on the structure. The 300-K PL data in Fig. 3 show the effect of adding a 150-nm Si-doped Al$_{0.10}$Ga$_{0.90}$N lower cladding (below the MQW active region) and a 20-nm Al$_{0.10}$Ga$_{0.90}$N cap to the optimized MQW PL structure. The upper clad-



FIG. 2. Room-temperature Hall effect data for Mg-doped GaN films showing the relationship between electron carrier concentration and film resistivity versus the input Cp$_2$Mg molar flow.
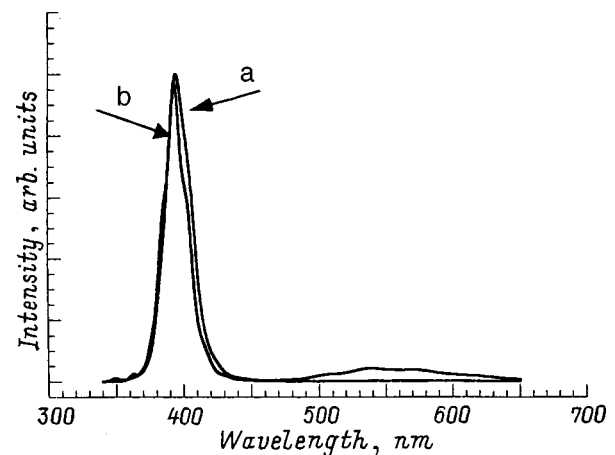


FIG. 3. Room-temperature PL spectra for MQW structures grown at 830 °C, consisting of 35 Å In$_{0.13}$Ga$_{0.87}$N quantum wells and 70 Å In$_{0.03}$Ga$_{0.97}$N barriers, doped with 20 sccm SiH$_4$, and (a) without or (b) with upper and lower Al$_{0.10}$Ga$_{0.90}$N:Si cladding layers. FWHM: a—174 meV, b—170 meV.

ding, or cap, had to be thin in PL structures. This data shows that there is absolutely no strain-induced degradation caused by the added AlGaN. The MQW luminescence peak intensity, and FWHM are nearly identical for both samples.

X-ray diffraction is used primarily on these samples to verify the period thickness of the MQW region and to confirm InGaN and AlGaN alloy compositions. By measuring the spacing of the superlattice satellite peaks, an accurate measurement of the period thickness can be made. Shown in Fig. 4 is the (00.6) reflection $\omega$-$2\theta$ curve for two different MQW active regions. From this data and the formula for X-ray diffraction from a period structure, the period thickness measured from the spacing between the zeroth-order and first-order satellite peaks for sample *a* was 101 Å and for sample *b* was 106 Å. The intended period thickness was 105 Å for both samples, so the agreement is quite good and very reproducible from sample to sample. This shows that growth rates obtained from the growth and scanning electron microscopy characterization of thick (~100 nm) InGaN layers are very accurate down to the short growth times required for well and barrier thicknesses. Also notice that the zeroth-order peak for curve *b* is slightly further separated from the GaN peak than for curve *a*. This peak corresponds to the average indium content of the MQW. The average indium content of *b* is larger because the indium content of the well and barrier is slightly higher (15 and 5% compared to 13 and 3%) and the fraction of the well thickness to the overall thickness is higher (43/105 compared to 30/105). So, again, the X-ray diffraction data are in good agreement with the structure that we intended to grow.

The material quality of the MQW active region has been qualified by performing 300-K high-intensity photopumping experiments. Photopumping provides the high power densities needed for population inversion in the active region without the need for electrical injection through doped layers. Therefore, a qualitative analysis of the active region's potential for achieving stimulated emission is possible without having to worry about internal heating in the structure due to resistive current paths or about defect induced degradation due to complicated epitaxial layer structures. Figure 5
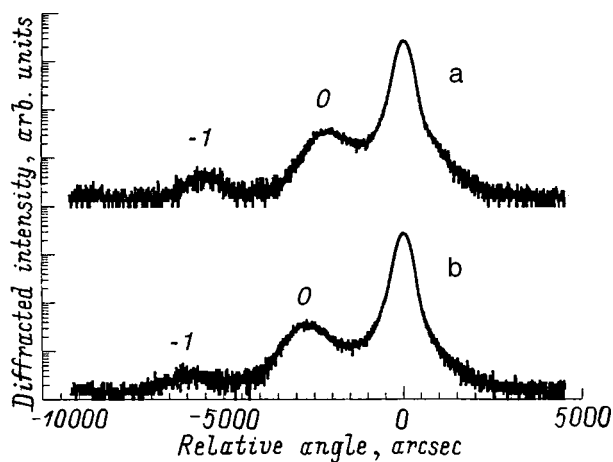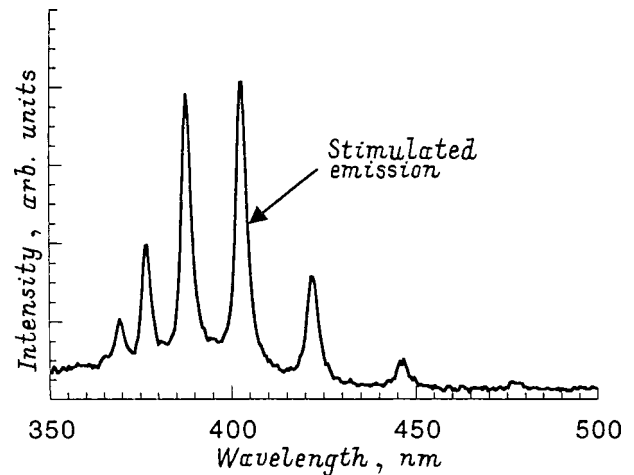


FIG. 5. Stimulated emission spectrum for photopumped (pulsed $N_2$ layer) MQW structure consisting of five periods of 42 Å $In_{0.15}Ga_{0.85}N$ quantum well and 63 Å $In_{0.05}Ga_{0.95}N$ barrier.

shows the results of the photopumping experiment on the MQW PL structure which corresponds to an optimized MQW active region. This experiment was performed with a pulsed (frequency $f < 100$ Hz) nitrogen laser capable of providing several MW/cm$^2$ peak power density. Attenuators were placed in the beam to reduce the optical energy incident on the sample. The recorded luminescence emitted from the surface of the sample (perpendicular to the growth plane) clearly shows significant narrowing to the linewidth at the characteristic wavelength $\lambda = 400$ nm, which is a good indication of stimulated emission. Also, Fabry–Perot cavity modes, whose spacing corresponds roughly to the vertical thickness of the entire epitaxial structure (~1.5 $\mu$m) are evident. The air/semiconductor interface and the semiconductor/substrate interface create the vertical cavity. We also conclude from the change in slope of the sample output power versus, relative pump input power data that this material is operating under stimulated emission conditions.

### 2.3. Selective-area and lateral epitaxial overgrowth

Joyce and Baldrey first demonstrated selective-area epitaxy (SAE) of semiconductors in 1962. They demonstrated SAE of Si using $SiO_2$ masks and $SiCl_4$ as a precursor.[14] An extension of this work, SAE and subsequent lateral epitaxial overgrowth (LEO) of a semiconductor film over the mask has also been reported for a variety of materials.[15] Recently, Davis *et al.*[16–18] have reported using MOCVD for the selective-area lateral epitaxial overgrowth (SALEO) of GaN on 6$H$-SiC substrates and have reported that the portion of the GaN film grown over the oxide mask has a reduced density of threading dislocations. More recently, Nakamura *et al.*[19] have reported employing the SALEO process to grow complete AlGaN/InGaN/GaN injection laser device structures having a greatly improved reliability performance due to the reduction of the density of threading dislocations in the active region, which is located in the SAE portion of the layer grown directly over the mask openings. The dislocation structure of these LEO materials is of great interest and it is reported to be lower by over two orders of magnitude than



FIG. 4. X-ray $\omega$-$2\theta$ (00.6) curves for MQW active regions consisting of (a) 35 Å $In_{0.13}Ga_{0.87}N$ quantum wells, 70 Å $In_{0.03}Ga_{0.97}N$ barriers and (b) 42 Å $In_{0.15}Ga_{0.85}N$ quantum wells, 63 Å $In_{0.05}Ga_{0.95}N$ barriers.

the materials grown directly on sapphire substrates.[20] While these demonstrations of high-quality III–N materials have resulted in an increased interest in the SALEO process, not many reports of the dependence of the details of the GaN SALEO upon growth conditions have been published.

The SALEO films in this study are grown at a pressure $\sim$100 Torr and at temperatures in the range $1030 \leqslant T_g \leqslant 1100\,°C$. The GaN heteroepitaxial ''substrates'' are prepared by growing $\sim$2 $\mu$m of GaN on (0001) sapphire by MOCVD at $T_g \simeq 1050\,°C$. Next, an $\sim$100-nm-thick $SiO_2$ mask is deposited by plasma-enhanced chemical vapor deposition (PECVD) using $SiH_4$ and $N_2O$ precursors. Conventional optical lithography and wet chemical etching are used to produce various stripe patterns in the $SiO_2$ mask layer.

We have achieved smooth vertical sidewalls and rectangular cross section GaN SALEO mesas for mask stripes oriented along the $\langle 1\bar{1}00 \rangle$ direction, while triangular cross section mesas are obtained for stripes parallel to the $\langle 11\bar{2}0 \rangle$ direction. The mesas with a rectangular cross section have surfaces bounded by a (0001) top facet and $\{11\bar{2}0\}$ side facets, while $\{1\bar{1}01\}$ side facets bound the triangular cross section stripes. We have observed that flat surfaces can be obtained for homoepitaxial films grown with mask stripes oriented along the $\langle 1\bar{1}00 \rangle$ and that the lateral-to-vertical growth ratio (LTVGR) is as large as $\sim$2.0 for this orientation. For stripes oriented along $\langle 11\bar{2}0 \rangle$, the growth in the lateral direction is limited by the slow-growing $\{1\bar{1}01\}$ facets, leading to the narrow triangular cross section typical of growth bounded by slow-growing planes. Using atom-force microscopy, we have measured a RMS roughness as low as $\sim$0.1 nm on the top of the flat mesas for stripes parallel to $\langle 1\bar{1}00 \rangle$. For comparison, the GaN film grown in the large-area ''open'' region of the mask at the wagon-wheel hub also has an RMS roughness of $\sim$0.1 nm. The facet sidewalls are found to be more vertical and the mesas are smoother for higher GaN growth temperatures $\sim$1100 °C.

The high-resolution scanning cathodoluminescence (CL) images (300 K) for undoped and Si-doped GaN SALEO mesas indicate that the intensity of the near-bandedge emission is more uniform and about a factor of 2 or 3 times higher for the GaN:Si LEO region grown over the $SiO_2$ mask compared to the GaN homoepitaxial SAE region grown between the mask stripes. The CL linewidths and peak wavelengths are essentially the same for all areas of the wafer. Also, the yellow-band emission is reduced relative to the bandedge emission for the LEO region. This may indicate that a higher density of nonradiative traps exists in the SAE region compared to the LEO material.[21] Initial plan-view and cross-section transmission electron microscope characterization of these materials supports this interpretation and has shown that there are $\sim 5 \times 10^8\,cm^{-2}$ threading dislocations in the SAE portion of the film; however, no threading dislocations are observed in the LEO portion of the film.[22] The CL emission spectra for the InGaN film grown on the GaN SALEO film are shown in Fig. 6. The peak intensity is brighter from the region corresponding to growth over the mask while the wavelength and FWHM are somewhat reduced for the SAE
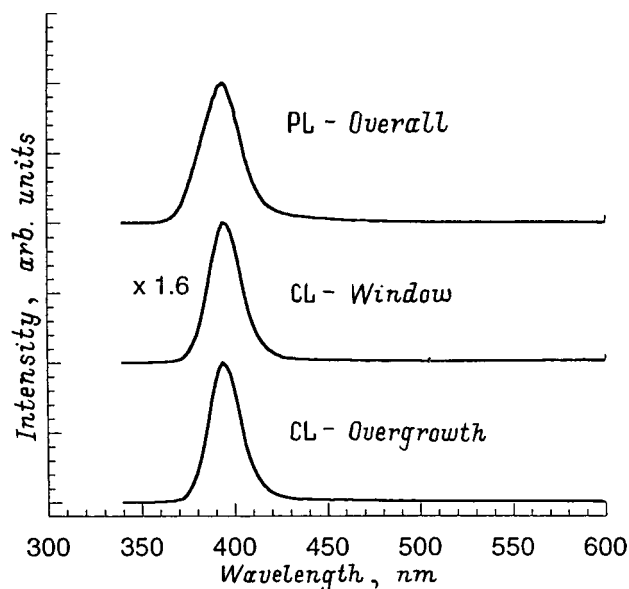


FIG. 6. Room-temperature cathodoluminescence spectra for InGaN film grown on GaN SALEO ''substrate.'' The CL from the LEO region is 1.6 times the peak intensity of the SAE region.

compared to the LEO regions. These results show promise for the application of SALEO to improved nitride devices.

## 3. SUMMARY AND CONCLUSIONS

We have shown that MOCVD process can produce high-quality AlGaN/InGaN/GaN heterostructure films and we have optimized the properties of the MQW structures. These structures have operated as pulsed lasers at 300 K. In addition, we have employed the SALEO process to produce GaN films on sapphire substrates with no apparent threading dislocations. The SALEO process can be repeated two or more times to produce a high-quality epitaxial film which is free of threading dislocations from the GaN underlying heteroepitaxial film.

*E-mail: dupuis@mail.utexas.edu

[1] I. Akasaki, H. Amano, Y. Koide, K. Hiramatsu, and N. Sawaki, J. Cryst. Growth **98**, 209 (1989).

[2] I. Akasaki and H. Amano, J. Electrochem. Soc. **141**, 2266 (1994).

[3] S. Nakamura, M. Senoh, N. Iwasa, and S. Nagahama, Jpn. J. Appl. Phys. **34**, L797 (1995).

[4] S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, Jpn. J. Appl. Phys. **35**, L74 (1996).

[5] S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimito, Jpn. J. Appl. Phys. **35**, L217 (1996).

[6] S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, Y. Sugimoto, and H. Kiyoku, Appl. Phys. Lett. **70**, 1417 (1997).

[7] I. Akasaki, H. Amano, S. Sota, H. Sakai, T. Tanaka, and M. Koike, Jpn. J. Appl. Phys. **34**, L1517 (1995).

[8] K. Itaya, M. Onomura, J. Nishio, L. Sugiura, S. Saito, M. Suzuki, J. Rennie, S. Nunoue, M. Yamamoto, H. Fujimoto, Y. Kokubun, Y. Ohba, G. Hatakoshi, and M. Ishikawa, Jpn. J. Appl. Phys. **35**, L1315 (1996).

[9] G. E. Bulman, K. Doverspike, S. T. Sheppard, T. W. Weeks, M. Leonard, H. S. Kong, H. Dieringer, C. Carter, J. Edmond, J. D. Brown, J. T. Swindell, J. F. Schetzina, Y.-K. Song, M. Kuball, and A. Nurmikko, *Device Research Conference and Electronic Materials Conference* (Ft. Collins, Co., 1997).

[10] Q. Chen, M. A. Khan, C. J. Sun, and J. W. Yang, Appl. Phys. Lett. **31**, 1781 (1995).

[11] J. C. Carrano, P. A. Grudowski, C. J. Eiting, R. D. Dupuis, and J. C. Campbell, Appl. Phys. Lett. **70**, 1992 (1997).

[12] Z. Fan, S. N. Mohammad, O. Aktas, A. E. Botchkarev, A. Salvador, and H. Morkoc, Appl. Phys. Lett. **69**, 1229 (1996).

[13] S. Yoshida, S. Misawa, and S. Gonda, J. Appl. Phys. **63**, 6844 (1982).

[14] B. D. Joyce and J. A. Bradkey, Nature (London) **195**, 458 (1962).

[15] See A. D. Morrison, T. Daud, U.S. Patent No. 04522661, for a discussion of the application of SALEO to the growth of high-purity semiconductors.

[16] R. F. Davis, *Gordon Conf. Electronic Materials: Chemistry, Excitations, and Processing* (Henniker, New Hampshire, 1997) (invited paper, unpublished).

[17] O. H. Nam, M. D. Bremser, T. S. Zheleva, and R. F. Davis, Appl. Phys. Lett. **71**, 2638 (1997).

[18] T. S. Zheleva, O. H. Nam, M. D. Bremser, and R. F. Davis, Appl. Phys. Lett. **71**, 2472 (1997).

[19] S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, Y. Sugimoto, T. Kozaki, H. Umemoto, M. Sano, and K. Chocho, Appl. Phys. Lett. **72**, 211 (1998).

[20] S. Nakamura, *The 24th International Symposium on Compound Semiconductors* (San Diego, CA, 1997) paper plenary 1.

[21] J. Park, P. A. Grudowski, C. J. Eiting, and R. D. Dupuis, Appl. Phys. Lett. **73**, 333 (1998).

[22] J. Park, P. A. Grudowski, C. J. Eiting, R. D. Dupuis, and Z. Liliental-Weber (unpublished).

# High density 2DEG in III–V semiconductor heterostructures and high-electron-mobility transistors based on them

V. G. Mokerov,* Yu. V. Fedorov, and A. V. Hook

*Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1064–1065 (September 1999)

Situation in high-electron-mobility transistor (HEMT) technology is discussed. The $N$-AlGaAs/InGaAs/GaAs pseudomorphic HEMT's are now considered as most advanced for mm-wave monolithic circuits, but metamorphic $N$-In$_x$Al$_{1-x}$As/In$_y$Ga$_{1-y}$As/In$_x$Al$_{1-x}$As HEMT's grown on GaAs substrates are very promising for the future high-frequency devices. High density 2DEG in HEMT's is analyzed by means of the Hall effect and photoluminescence measurements. Processing technology of the sub-0.25-$\mu$m pseudomorphic HEMT's, metamorphic HEMT's and their characteristics are also described. © *1999 American Institute of Physics.* [S1063-7826(99)00909-6]

High-electron-mobility transistor (HEMT) technology is widely used for high-frequency application due to the enhanced electron mobility ($\mu_{2D}$) and velocity ($v_{st}$) at high electron density ($n_{2D}$) in a device channel. This provides high values of the device current $I_D$, transconductance $g_m = dI_D/dV_G$, cutoff frequency $f_\tau = v_{st}/2\pi L_g$, output gain $G$, and the low value of noise. Main tendencies of evolution of this technology correspond to investigation of heterostructures with higher values of the conduction band offset $\Delta E_c$ and, respectively, $n_{2D}$, with higher values of the energy separation $\Delta E_{\Gamma L}$ between $\Gamma$ and $L$ valleys, with lower electron effective mass $m_e$ and, respectively, with higher values of $\mu_{2D}$ and $v_{st}$, combined with the reduction of the HEMT gate length $L_g$.

Since 1980 HEMT technology has passed the following stages:

—lattice matched (LM) $N$-AlGaAs/InGaAs/GaAs HEMT's with $\Delta E_c \simeq 250$ meV and maximum electron density $n_{2D}^{max} < 1 \times 10^{12}$ cm$^{-2}$;

—pseudomorphic (PM) $N$-Al$_x$Ga$_{1-x}$As/In$_y$Ga$_{1-y}$As/GaAs HEMT's,[1] which provide higher $\Delta E_c$ ($\Delta E_c \simeq 300$ meV) and $n_{2D} \simeq 2 \times 10^{12}$ cm$^{-2}$; however, the effect of strain relaxation restricts the maximum values of In content $y$ and thickness $d_{ch}$ of In$_y$Ga$_{1-y}$As channel layer by critical values, and, accordingly, restricts the values of $n_{2D}$, $\mu_{2D}$, and $v_{st}$;

—the highest frequency is realized in InP-based LM-In$_{0.52}$Al$_{0.48}$As/In$_{0.53}$Ga$_{0.47}$As HEMT's (Ref. 2): $\Delta E_c > 500$ meV, $n_{2D} > 3.5 \times 10^{12}$ cm$^{-2}$, and $\mu_{2D} > 10\,000$ cm$^2$/(V·s) at 300 K; there are some disadvantages associated with a rather low value of the Schottky barrier height for In$_{0.52}$Al$_{0.48}$As and with impact ionization in In$_{0.53}$Ga$_{0.47}$As channel due to rather small energy band gap, $E_g \simeq 750$ meV;

—recent advantages of HEMT technology are associated with the metamorphic (MM) $N$-In$_x$Al$_{1-x}$As/In$_y$Ga$_{1-y}$As/In$_x$Al$_{1-x}$As HEMT-structures with desirable In contents $y$ and $x$, ranging from 0 to 0.60, grown on the lattice-mismatched GaAs substrate;[3] they provide $\Delta E_c > 600$ meV,

$n_{2D} > 4 \times 10^{12}$ cm$^{-2}$, and $\mu_{2D} > 10\,000$ cm$^2$/(V·s) at 300 K.

In this paper we investigate the high-density 2DEG in $N$-AlGaAs/InGaAs/GaAs and $N$-InAlAs/InGaAs/InAlAs MM-structures and HEMT's based on them, because the former is now the most advanced industrial technology for millimeter-wave monolithic circuits and the latest is very promising for the future high-frequency electronics.

## MOLECULAR BEAM EPITAXY OF HEMT-STRUCTURES

HEMT-structures studied here were grown by molecular beam epitaxy (MBE) on (001) GaAs substrates. PM-HEMT-structure consists of an 0.3-$\mu$m GaAs buffer layer, an undoped InGaAs channel layer ($d_{ch} < 25$ nm), an undoped AlGaAs spacer layer (thickness $d_s < 10$ nm), a $\delta$ (Si)-doped layer, undoped 20-nm AlGaAs Schottky layer and a Si-doped $n^+$-GaAs ($3 \times 10^{18}$ cm$^{-3}$) cap layer. The quality of the AlGaAs/InGaAs interface and maximum values $\mu_{2D}^{max}$ and $n_{2D}^{max}$ for 2DEG are essentially limited by the surface segregation of In atoms during MBE growth. Removing these atoms by means of the growth interruption and subsequent surface heating can achieve $\mu_{2D}$ as high as 53\,000 cm$^2$/(V·s) at 77 K.

Typical MM-HEMT-structure consists of an undoped 0.3-$\mu$m GaAs buffer layer, an undoped 1.3-$\mu$m AlInGaAs relaxed buffer layer (RBL) with changeable composition from GaAs to In$_x$Al$_{1-x}$As, an undoped 0.3-$\mu$m In$_x$Al$_{1-x}$As layer, a short-period superlattice, an undoped In$_y$Ga$_{1-y}$As channel layer, an undoped In$_x$Al$_{1-x}$As spacer layer ($d_s < 10$ nm), a $\delta$ (Si)-doped layer, an undoped 20-nm In$_x$Al$_{1-x}$As Schottky layer, and a 40-nm $n^+$-In$_y$Ga$_{1-y}$As ($5 \times 10^8$ cm$^{-3}$) cap layer. RBL is of main importance in MM-HEMT's. It is inserted between the substrate and the active layers to accommodate large lattice mismatch by formation of dislocations, to trap them, and to prevent their propagation in HEMT-structure. Samples with different RBL composition and depth profile have been grown. The highest electron mobility was achieved in the samples with the RBL depth profile corresponding to the nearly constant value of

the gradient of the lattice mismatch. They have $\mu_{2D} = 32\,000\,\mathrm{cm^2/(V \cdot s)}$ and $n_{2D} = 2.2 \times 10^{12}\,\mathrm{cm^{-2}}$ at $T = 77\,\mathrm{K}$, a good surface morphology (associated with the cross-match pattern) and a large photoluminescence (PL) signal.

## TRANSPORT PROPERTIES AND PHOTOLUMINESCENCE SPECTRA OF HEMT STRUCTURES

In order to understand, how the variations of InAs mole fraction $y$ and thickness of the channel layer $d_{ch}$ influence the characteristics of 2DEG in $N$-$Al_xGa_{1-x}As/In_yGa_{1-y}As/GaAs$ heterostructures, including the influence of the backside heterobarrier, and also to determine how 2DEG is redistributed as a result of varying the thickness $d_{ch}$, the Hall measurements of $\mu_{2D}$ and $n_{2D}$ and the photoluminescence studies of PM-HEMT-structures have been performed. It was shown that $\mu_{2D}$ is virtually independent of $d_{ch}$ in the range 19–3 nm. According to the PL studies, the regular spatial redistribution of 2DEG between InGaAs and GaAs layers occurs when $d_{ch}$ is reduced. At $d_{ch} > 10\,\mathrm{nm}$ 2DEG is mainly located in InGaAs. However, at $d_{ch} < 7-8\,\mathrm{nm}$ more electrons transfer from InGaAs to GaAs, and at $d_{ch} < 3.5-4\,\mathrm{nm}$ nearly all the electrons transfer to GaAs. Because $\mu_{2D}$ is approximately as high in GaAs as in InGaAs, the electron mobility may be independent of $d_{ch}$, as observed experimentally. At the same time, these results demonstrate that electron scattering associated with the backside heterobarrier (its effect should increase with decreasing $d_{ch}$) has no visible impact on $\mu_{2D}$. This finding is in contrast with $\mu_{2D}(d_{ch})$ dependence for AlGaAs/GaAs/AlGaAs system, where the reduction of $d_{ch}$ below 30 nm results in a strong degradation of $\mu_{2D}$ due to the enhanced influence of the backside interface. To determine the role of the backside heterobarrier in the electron scattering, PM-HEMT-structures with AlGaAs layer as a backside heterobarrier also have been studied. It was found that replacement of GaAs by AlGaAs leads to a significant reduction of $\mu_{2D}$ (about four times) and $n_{2D}$, and to even stronger reduction with decreasing $d_{ch}$. The cause of this behavior, along with the higher interface electron scattering, can be attributed to progressively higher penetration of electron wave functions into AlGaAs barrier (due to quantum-mechanical tunneling). The electrons will therefore spend progressively more time in AlGaAs, where their mobility is extremely low.

PL spectra of MM-$In_xAl_{1-x}As/In_yGa_{1-y}As/InGaAs$ heterostructures with $x, y = 0.32$ and 0.52 also have been studied. At $d_{ch} > 40\,\mathrm{nm}$ these spectra correspond to the bulk material and the two PL lines correspond to respective band gaps $E_g$. The band-gap discontinuities $\Delta E_g$ at the interfaces are 0.70 and 0.72 eV for $x, y$ of 32 and 52%, respec-

tively. For $d_{ch} < 30\,\mathrm{nm}$ the InGaAs ''bulk,'' PL line in each sample is split up into two components, which are associated with the optical transitions between the two lowest 2DEG subbands and the hole subband. It should be noted that the PL lines in MM-HEMT-structures are weaker and broader than those in PM-HEMT-structures, evidently due to the higher dislocation density.

## SUBQUARTER-MICROMETER PM-HEMT's AND MM-HEMT's

The realization of HEMT's with the gate length $L_g < 0.25\,\mu\mathrm{m}$ was performed by using the electron lithography in combination with a special mask consisting of a $SiO_2$ layer, a metallic layer, and an electron resist layer. After the exposure and the development of the electron resist, the very narrow opening was formed in the metallic layer by $Ar^+$-ion beam etching. Such a metallic mask makes it possible to realize $L_g < 0.2\,\mu\mathrm{m}$ and to avoid the problem of the low stability of the electron resist to the plasma etching.

In PM-HEMT the typical extrinsic transconductance $g_m^{ex}$ was in the range 300–550 mS/mm, and the drain current density $I_D$ was in the range 300–500 mA/mm. In MM-HEMT's the best results were obtained for In content of 0.32. Their main advantages over PM-HEMT's are associated with the higher values of $g_m^{ex}$ and $I_D$: $g_m^{ex}$ was in the range from 500 to 800 mS/mm, and $I_D$ was in the range from 500 to 900 mA/mm, which are explained by larger $n_{2D}$ and $v_{st}$ in MM-HEMT's. In the case of PM-HEMT's rf characteristics in the frequency range $f = 12-37\,\mathrm{GHz}$ have also been studied. The presence of the strong correlation between the output gain $G$ and $g_m^{ex}$ was confirmed at 12 GHz. It was shown that the dependence of $G$ on the gate width $W_g$ at 37 GHz has a maximum at $W_g = 120\,\mu\mathrm{m}$, but for 12 GHz $G$ increases with $W_g$ to 150 $\mu\mathrm{m}$ without any maximum. The increase of $G$ in the range 40–100 $\mu\mathrm{m}$ at $f = 37\,\mathrm{GHz}$, and to 150 $\mu\mathrm{m}$ at $f = 12\,\mathrm{GHz}$ are explained in terms of the increase of the absolute transconductance. The reduction of $G$ at $W_g > 120\,\mu\mathrm{m}$ at $f = 37\,\mathrm{GHz}$ is a result of a mismatch of the PM-HEMT parameters with the input of the measurement waveguide line in the mm-wave range. At $f = 12\,\mathrm{GHz}$ such a mismatch is expected at $W_g > 150\,\mu\mathrm{m}$.

[*]E-mail: mok@mail.cplire.ru

[1] L. D. Nguyen, P. J. Tasker, D. C. Radulescu, and L. F. Eastman, IEEE Trans. Electron Devices **36**, 2243 (1989).

[2] L. D. Nguyen, L. E. Larson, and U. K. Mishra, Proc. IEEE **80**, 491 (1992).

[3] P. Win, Y. Druelle, and A. Cappy, Appl. Phys. Lett. **61**, 922 (1992).

# Heteroepitaxial growth of InAs on Si: a new type of quantum dot

G. E. Cyrlin,[*] V. N. Petrov, V. G. Dubrovskii, Yu. B. Samsonenko, N. K. Polyakov,
A. O. Golubok, S. A. Masalov, and N. I. Komyak

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

V. M. Ustinov, A. Yu. Egorov, A. R. Kovsh, M. V. Maximov, A. F. Tsatsul'nikov,
B. V. Volovik, A. E. Zhukov, P. S. Kop'ev, N. N. Ledentsov, and Zh. I. Alferov

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

D. Bimberg

*Institut für Festkörperperphysik, Technische Universität, Berlin, Germany*
(Submitted March 1, 1999; accepted for publication March 2, 1999)
Fiz. Tekh. Poluprovodn. **33**, 1066–1069 (September 1999)

The mechanism for heteroepitaxial growth in the InAs/Si system is studied by reflection high-energy electron diffraction, scanning tunnelling microscopy, and photoluminescence. For certain growth conditions, InAs nanostructures are found to develop on the Si surface immediately during the growth process in the course of molecular beam epitaxy. The range of substrate temperatures that lead to formation of nanosized islands is determined. InAs quantum dots grown on a buffer Si layer with a silicon layer of thickness 50 nm grown on the top produced photoluminescence lines at a wavelength of 1.3 $\mu$m at 77 K and 1.6 $\mu$m at 300 K.  © *1999 American Institute of Physics.* [S1063-7826(99)01009-1]

## INTRODUCTION

Silicon is currently the main material for microelectronics, accounting for roughly 90% of the market for semiconductor devices. Its properties, such as high thermal conductivity and mechanical durability, the presence of a stable oxide on its surface, and the fact that it can be made into large diameter, dislocation-free substrates make this material irreplaceable for most applications.

In the meantime, silicon has not gained wide acceptance for use in optoelectronic devices (lasers, light emitting diodes, etc.), because it has an indirect-gap band structure and because radiative recombination cannot occur without an interaction with an additional particle (for example, a phonon). This situation is typical of all indirect-gap semiconductors (AlAs, GaP, etc.). It is known that among the III–V compounds, introducing thin quantum-well layers, quantum dots, or quantum wires of a direct-gap semiconductor (GaAs, InAs) into an indirect-gap host (AlAs) leads to a dramatic increase in the efficiency of radiative recombination and therefore makes it possible to fabricate lasers and light-emitting diodes.

By analogy with the III–V compounds, we may assume that introducing quantum dots of a narrow direct-gap semiconductor (for example, InAs) into a silicon host will also make it possible to create efficient light-emitting devices.[1] We have demonstrated experimentally the possibility, in principle, of producing InAs quantum dots on a silicon surface in an earlier paper.[2]

In this paper we study the heteroepitaxial growth of InAs on an Si(100) surface by reflection high-energy electron diffraction (RHEED), scanning tunnelling microscopy (STM), and phototluminescence techniques.

## 1. EXPERIMENT

The growth experiments were done on an ÉP 1203 molecular beam epitaxy (MBE) system using Si(100) substrates which had undergone initial preparation by the method described in Ref. 3. The samples were glued with indium onto a standard molybdenum holder. The flux of indium was calibrated in advance from the oscillations in the RHEED mirror reflex and corresponded to 0.1 monolayer (ML) per second for InAs during growth of an InGaAs solid solution on GaAs with a molecular fraction of ~15% of indium. The oxide layer was removed from the silicon by heating the substrate in the growth chamber to a temperature in the range 820–870° for a 15-min exposure, after which a $(2\times1)$ surface reconstruction typical of an Si(100) surface or a mixed domain structure of type $(2\times1)$ and $(1\times2)$ was observed. Subsequently, the sample temperature was lowered and InAs was deposited by traditional MBE. After deposition of a certain amount of InAs, the growth process was interrupted and the sample was rapidly cooled to room temperature. The processes of removing the oxide layer and growth were monitored *in situ* using systems for recording and processing of events RHEED.[4] The morphology of the surface of the samples was studied *ex situ* using STM in accordance with the methods[5,6] for studying quantum-well structures of InAs on GaAs surfaces. For extended STM measurements, the samples were placed in vacuum pump oil to protect the surface from oxidation in the air. For the photoluminescence excitation studies, a 50 to 100-nm-thick layer of Si was
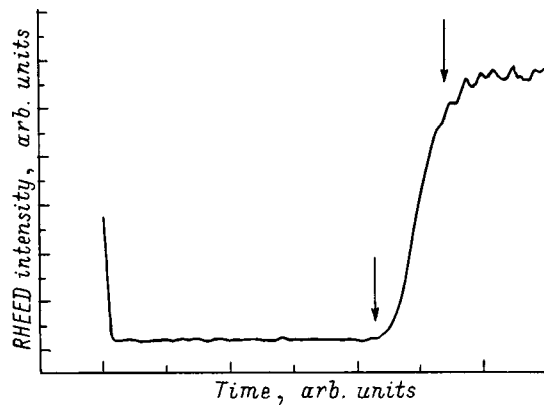
FIG. 1. Typical dynamic variation in the intensity of the RHEED reflection on going from two-dimensional to three-dimensional heteroepitaxial growth of InAs/Si.



FIG. 2. Critical thickness for formation of nanostructures, $d_{cr}$, as a function of the substrate temperature $T_s$ for flux ratios $[As_4]/[In]=2$ (*1*), 5 (*2*), 8 (*3*), and 12 (*4*).

grown over the quantum-well InAs structures by traditional MBE. Photoluminescence was excited using an $Ar^+$ laser with an output wavelength of 514.5 nm and the detector was a Ge photodiode. The excitation power density was 0.2–200 $W/cm^2$.

## 2. EXPERIMENTAL RESULTS

The characteristic feature of InAs/Si(100) is a strong dependence of the growth mechanism on the growth conditions (substrate temperature, flux ratios, etc.), which shows up in the morphology of the surface. Thus, an observation of the RHEED dynamics shows that, for growth at a sufficiently high substrate temperature ($T_s = 470$– $500\,°C$), the diffraction patterns retain a line structure right up to thicknesses of 50 ML or more, while the surface reconstruction changes from ($2\times1$) to ($3\times1$) or ($1\times1$). Deposition of InAs at a lower substrate temperature ($T_s < 450\,°C$) causes a transformation of the RHEED line spectrum into a pattern typical of three-dimensional growth (point reflections) at thicknesses of $\sim 0.7 - 5$ ML (depending on the ratio of the fluxes). STM studies of the surface shows that two types of morphological features form during growth at high temperatures $T_s > 450\,°C$: a corrugated structure with a modulation depth of $\sim 20$ nm and clusters with sizes of $\sim 400$ nm. Another situation is observed when the temperature is lowered to $T_s \lesssim 450\,°C$. In this case, a spatially uniform array of islands in the nanometer range develops.[2]

The difference in the surface morphologies that develops at different sample temperatures can be explained, first of all, by a change in the surface free energy of the system as the temperature rises and, therefore, by a transition from the formation of coherent three-dimensional islands of nanometer size to mesoscopic three-dimensional clusters which appear on the surface of the corrugated wetting surface and, second, by different attachment coefficients for InAs on an Si(100) surface at different temperatures.

Figure 1 shows the typical dynamic dependence of the zeroth order reflection in a RHEED pattern, taken with an angle of incidence of 1° and an angle of reflection of 1.5° under growth conditions such that a transition from a two- to a three-dimensional growth mechanism takes place. We as-
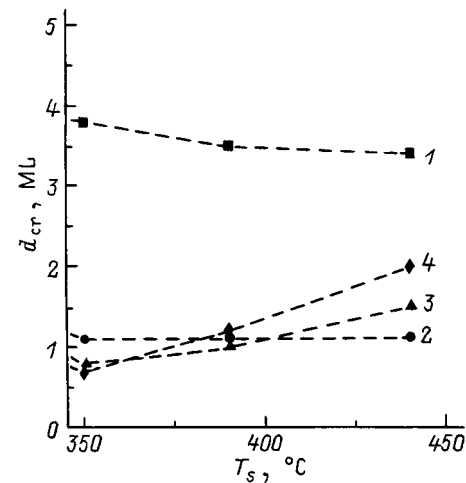
sumed that this transition begins at the time denoted by the arrow on the left, i.e., at the beginning of the conversion of the line structure of the RHEED pattern into point reflections. By analogy with the InAs/GaAs heteroepitaxial system, we call this the critical thickness for formation of nanostructures on a surface ($d_{crit}$). The time the volume pattern stabilizes is indicated by the arrow on the right. The characteristic time for formation of an array of InAs islands (the time interval between the arrows) is $\sim 1$ s (for a growth rate of 0.1 ML/s). The formation of quantum dots in the InAs/GaAs system is characterized by roughly the same time.

Figure 2 shows data obtained by RHEED on the substrate temperature dependence of $d_{crit}$ for $T_s = 350$– $450°$ and flux ratios $[As_4]/[In] = 2 - 12$. The lower limit in this temperature range was chosen for two reasons: first, because of the difficulty in determining an exact value of $T_s$ at such low values and, second, because the covering Si layer grown at low $T_s$ is of low crystal quality, so it is impossible to obtain device structures. This is confirmed by our data,[7] according to which no photoluminescence signal is observed in structures where InAs quantum dots have been overgrown by a layer of Si at substrate temperatures below 350 °C.

The data in Fig. 2 imply that transitions from two- to three-dimensional growth in an InAs/Si system (under these growth conditions) occur within 0.7–4.0 ML. This indicates that, in these heteroepitaxial systems, both purely island (Volmer–Weber) growth and mixed (Stranski–Krastanov) growth with formation of a wetting layer and a system of three- dimensional islands can take place. This phenomenon will be studied in detail in a separate article.

As an example, Fig. 3 shows an STM image of part of an Si (100) surface after deposition of 1.2 ML InAs with a flux ratio $[As_4]/[In] = 8$ and a deposition temperature of 380 °C. (Growth was halted at the time indicated by the rightmost arrow in Fig. 1.) It should be noted that, in this case, the islands in the base are predominantly triangular in shape, in contrast with the case of deposition at lower temperatures (250 °C), where the islands were rhomboidal with a substantially higher surface density.[2]
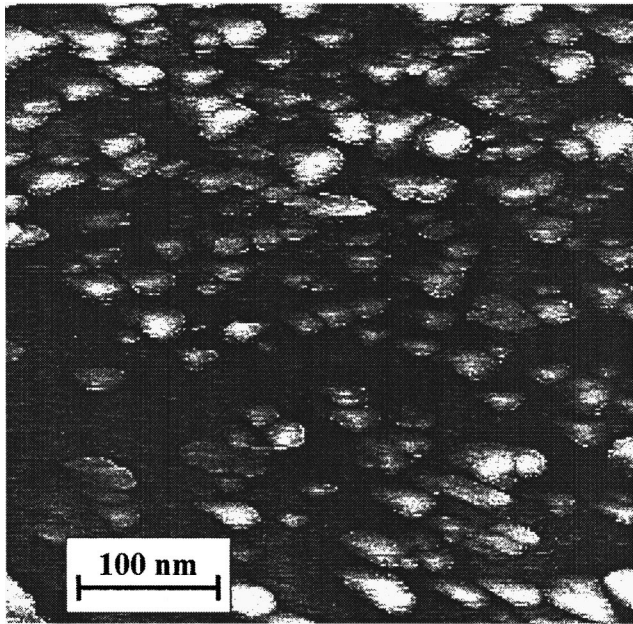
FIG. 3. STM image of a surface after deposition of 1.2 ML InAs on Si (100) at $T_s = 380$ °C. The sides of the image are parallel to the $[0\bar{1}1]$ and $[011]$ crystallographic directions.



FIG. 5. Photoluminescence spectra of InAs/Si quantum dots for excitation power densities $P$ (W/cm$^2$): *1* — 0.2, *2* — 2, *3* — 20, *4* — 200. $T = 10$ K.

A study of the optical properties of InAs quantum dots set in a silicon matrix by growing a 50-nm-thick layer of silicon on top of grown nanosized islands following deposition of 7 ML InAs at a substrate temperature $T_s = 400$ °C with an arsenic pressure of $1 \times 10^{-8}$ Torr in the vessel showed that the photoluminescence spectrum contains a broad line ($\sim 100$ meV) near a wavelength of $\lambda = 1.3$ $\mu$m for an observation temp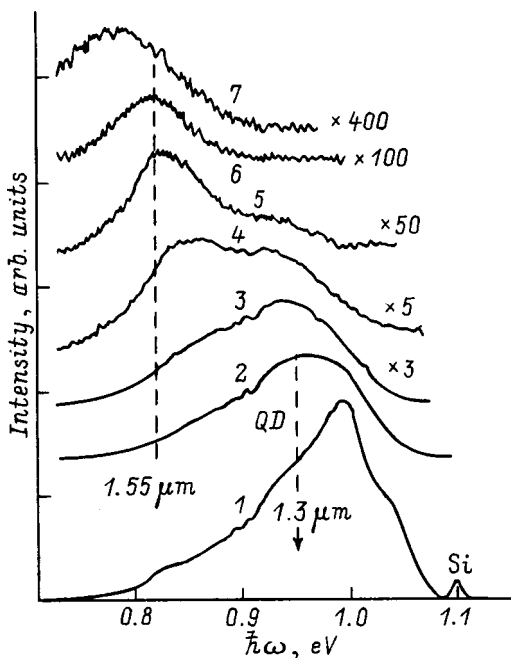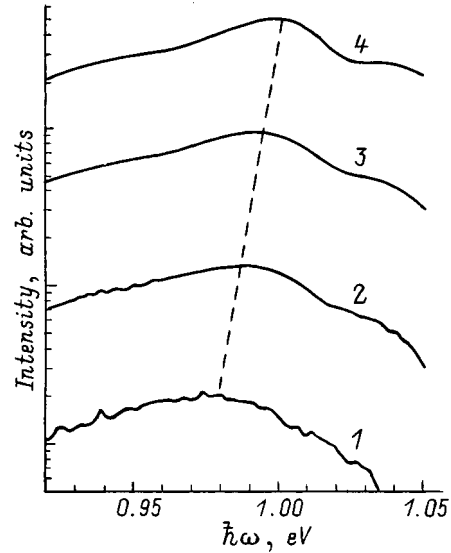erature $T = 77$ K. Raising the observation temperature to room temperature leads to a monotonic shift in the photoluminescence line toward longer wavelengths up to $\lambda = 1.6$ $\mu$m at $T = 290$ K (Fig. 4). This line was not seen in the spectra of the Si (100) substrates or in the spectra of samples in which the thickness of the deposited InAs was less than critical. An increase in the excitation intensity causes a shift in the peak of the photoluminescence line toward higher energies (Fig. 5). The nature of this shift is not entirely clear. A similar effect has been observed, for example, in type-II quantum dots;[8] it may be related to electrical charge that has accumulated in the quantum dots.[9]

We have, therefore, demonstrated the possibility, in principle, of directly forming InAs quantum dots on Si(100) during MBE.

The range of substrate temperatures appropriate for formation of nanosized islands has been determined. The photoluminescence spectra of InAs quantum dots with a silicon layer grown over them include a 1.3-$\mu$m line at 77 K and a 1.6-$\mu$m line at 300 K.

*)E-mail: cirlin@iai.rssi.ru; Fax: (812) 251 70 38



FIG. 4. Photoluminescence spectra of InAs/Si quantum dots at $T$ (K): *1* — 10, *2* — 70, *3* — 110, *4* — 150, *5* — 190, *6* — 230, *7* — 290. The excitation power density if $P = 100$ W/cm$^2$.

[1] N. N. Ledentsov, in M. Scheffler and R. Zimmermann (Eds.), *Prog. 23rd International Conference Physics Semiconductors*, Berlin (1996), World Scientific, Singapore (1996), Vol. 1, p. 19.

[2] G. E. Cyrlin, V. N. Petrov, V. G. Dubrovskii, S. A. Masalov, A. O. Golubok, N. I. Komyak, N. N. Ledentsov, Zh. I. Alfërov, and D. Bimberg, Pis'ma Zh. Tekh. Fiz., **24** (8), 10 (1998) [Tech. Phys. Lett. **24**, 290 (1998)].

[3] A. Ishisaka and Y. Shiraki, J. Electrochem. Soc. **133**, 666 (1986).

[4] G. M. Gur'yanov, V. N. Demidov, N. P. Korneeva, V. N. Petrov, Yu. B. Samsonenko, and G. E. Cyrlin, Zh. Tekh. Fiz. **42** (8), 111 (1997) [Tech. Phys. **42**, 882 (1997)].

[5] G. E. Cirlin, G. M. Guryanov, A. O. Golubok, S. Ya. Tipissev, N. N. Ledentsov, P. S. Kop'ev, M. Grundmann, and D. Bimberg, Appl. Phys. Lett. **67**, 97 (1995).

[6] A. O. Golubok, S. A. Masalov, N. B. Ponomareva, V. N. Petrov, S. Ya. Tipisev, and G. E. Cyrlin, *Surfaces. X-ray, Synchrotron, and Neutron Studies* [in Russian], No. 2, 70 (1998).

[7] A. Yu. Egorov, A. R. Kovsh, V. M. Ustinov, A. E. Zhukov, M. V. Maksimov, G. E. Cirlin, N. N. Ledentsov, D. Bimberg, P. Werner, and Zh. I. Alferov, J. Cryst. Growth (in press).

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

[9] A. F. Tsatsul'nikov, A. Yu. Egorov, P. S. Kop'ev, A. R. Kovsh, M. V. Maximov, V. M. Ustinov, B. V. Volivik, A. E. Zhukov, Zh. I. Alferov, G. E. Cirlin, A. O. Golubok, S. A. Masalov, V. N. Pertov, N. N. Ledentsov, R. Heitz, M. Grundmann, D. Bimberg, I. P. Soshnikov, P. Werner, U. Gösele, *et al.*, in *Proceedings of the 24th International Conference Physics Semiconductors*, Jerusalem, (1998), World Scientific, Singapore (in press).

Translated by D. H. McNeill

# High-efficiency AlGaAs/GaAs concentrator (2500 suns) solar cells

V. M. Andreev,* V. P. Khvostikov, V. R. Larionov, V. D. Rumyantsev, E. V. Paleeva, and M. Z. Shvarts

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

A study is made AlGaAs/GaAs heterostructure solar cells that have been optimized for ultrahigh concentrations of sunlight (1000–2500 suns). The maximum efficiencies were 25.1% at 500 suns, 25% at 1000 suns, and 22.8% at 2000 suns for sunlight passing through an air mass of AM1.5. Cells of this type open up the possibility of reducing the surface area of solar cells by more than three orders of magnitude and, as a result, substantially lowering the cost of electrical energy in power equipment with sunlight concentrators. © *1999 American Institute of Physics.* [S1063-7826(99)01109-6]

## 1. INTRODUCTION

Solar cells operating at ultrahigh concentrations of sunlight (more than 1000 suns) can have much smaller surface areas and greatly reduce the cost of electrical energy.

At present, the maximum confirmed efficiencies of concentrator solar cells are 24.6% for solar radiation in outer space[1–3] (AM0, 100 suns) and 27.6% for solar radiation near the earth's surface[3] (AM1.5, 255 suns). AlGaAs/GaAs heterostructure solar cells can be used to obtain high efficiencies, even for concentrations of sunlight exceeding 1000 suns. For example, an efficiency of 23% has recently been obtained[4] with a concentration of 1300 suns (AM1.5).

In order to obtain higher efficiencies at high concentrations of sunlight, resistive losses must be reduced while maintaining a high photocurrent and working voltage. In the work reported here, these goals have been met by creating an optimum AlGaAs/GaAs heterostructure by low-temperature (600–400 °C) liquid-phase epitaxy[2,5–7].

## 2. STRUCTURE OF THE SOLAR CELLS

Photocells were created on the basis of the heterostructures shown in Fig. 1. The following sequence of layers was grown on an $n$-GaAs substrate: $n$-AlGaAs (3 $\mu$m), $n$-GaAs (3 $\mu$m), $p$-GaAs (1–1.5 $\mu$m), and $p$-Al$_{0.85}$Ga$_{0.15}$As (0.05 $\mu$m). In order to reduce the resistive losses in this structure, the doping level was substantially increased (to $10^{19}$ cm$^{-3}$) in the $p$-GaAs layer next to the $p$-GaAs–$p$Al$_{0.85}$Ga$_{0.15}$As interface. In order to ensure a long diffusion path length for the electrons in this layer, an embedded electric field was created in this layer by smoothly reducing the acceptor concentration from the heterogeneous interface to the $p-n$-junction.

A strip front ohmic contact was fabricated into the highly doped part of the $p$-GaAs layer, which ensured a low contact resistance. A ''prismatic'' coating (Fig. 1) was cemented to the photoelement to reduce the optical losses due to shadowing by the contact strips; this coating deflected light rays onto a photoactive part of the element, free from contact strips.

## 3. RESULTS AND DISCUSSION

Because of the high efficiency of charge carrier collection over the spectral range 0.4–0.85 $\mu$m (Fig. 2) in photoelements based on these structures, near maximum photocurrents were obtained: 28.7 mA/cm$^2$ for ''earthbound'' solar radiation (AM1.5) and 35 mA/cm$^2$ in outer space (AM0).

In solar cells intended for operation at ultrahigh concentrations of sunlight, it is extremely important to achieve an optimum balance between the width and step size of the contact strips. Figure 3 shows the filling factor of the load characteristics as a function of concentration for three solar cells of different designs. A high value of $FF = 0.85$ can evidently be maintained in cells with a separation of 0.04 mm between contact strips with a width of 6-8 mm wide at sunlight concentrations $K_C = 4000$ suns. The resistive losses (spread resistance and contact resistance) in these cells were lowered by increasing the thickness of the $p$-GaAs layer to 1.2–1.5 $\mu$m, as well as by increasing the doping of this layer near the heterojunction to $(2-3) \times 10^{19}$ cm$^{-3}$. The gradient
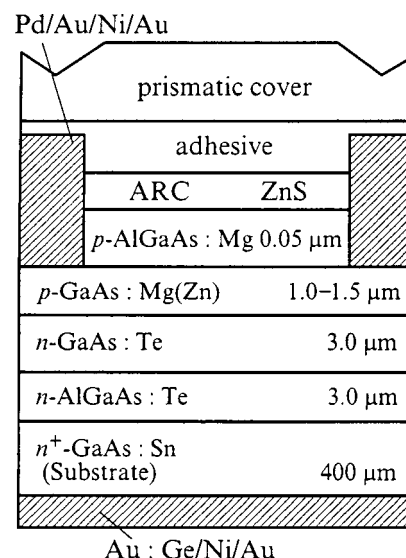


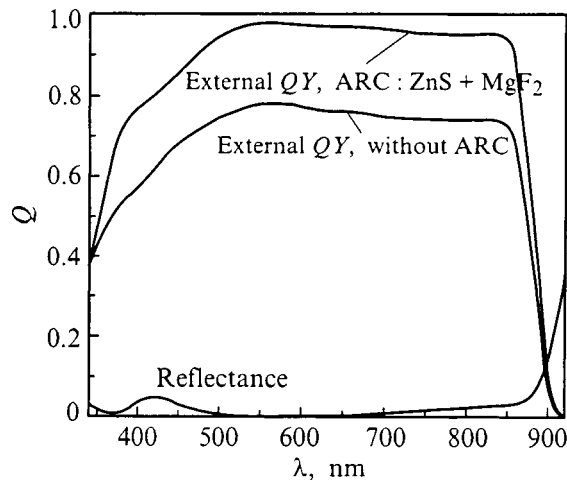FIG. 1. The cross section of an AlGaAs/GaAs heterostructure solar cell.

FIG. 2. Photoresponse spectra of a concentrator solar cell without an anti-reflecting coating and with a coating made of $ZnS + MgF_2$, as well as the reflection spectrum of the coated cell.



FIG. 4. The efficiency of one of our solar cells as a function of the degree of concentration of sunlight in outer space (AM0) and on the ground (AM1.5D).

in the hole concentration required to create a built-in ''pulling'' electric field was created in two ways: either by successive crystallization of several layers of *p*-GaAs with an increasing concentration of Mg in the melts, or by additional post-growth diffusion of zinc to a depth of 0.2–0.3 $\mu$m into a heterostructure with an epitaxial layer of *p*-GaAs of thickness 1.0–1.2 $\mu$m.

The contact strips shadowed about 15% of the cells without prismatic coatings. In order to reduce these losses, prismatic coatings made of transparent polymer (silicone) were cemented onto the solar cells by negative profiling of the matrix and lining this matrix with the silicone. With these coatings the losses due to shadowing were reduced to 3-4%.

Because of the reduced resistive losses in these cells, it was possible to attain record efficiencies at ultrahigh ($K_C \geqslant 1000$ suns) concentrations of sunlight: 25% for $K_C = 1000$, 24.8% for $K_C = 1300$, 22.8% for $K_C = 2000$, and 21.3% for $K_C = 2500$ suns. These studies showed that at the maximum attainable photocurrents of order 28 mA/cm$^2$ (AM1.5) in a single structure and at a load characteristic

filling factor of 0.85–0.87, the efficiency could be increased further to 27% at a sunlight concentration of 2000–3000 suns (Fig. 4).

The solar cells for operation at ultrahigh sunlight concentrations which we have developed can be used extensively in ground-based and space-based solar power systems: in ground-based concentrator solar power production equipment with low cost solar cells; in space based hybrid modules with concentrators and fiber optics and batteries for spacecraft aimed at the sun; and in systems for energy transmission by laser beams.

These devices permit a substantial buildup in the power of concentrator solar power systems without significant increases in production facilities for semiconductor heterostructures. Thus, at the Ioffe Institute, up to $10^5$ cm$^2$ heterostructures per year can be produced using the existing epitaxial equipment, which is sufficient to produce solar cells with a combined output of up to $2 \times 10^6$ W/g in modules operating at $K_C = 1000$ suns.

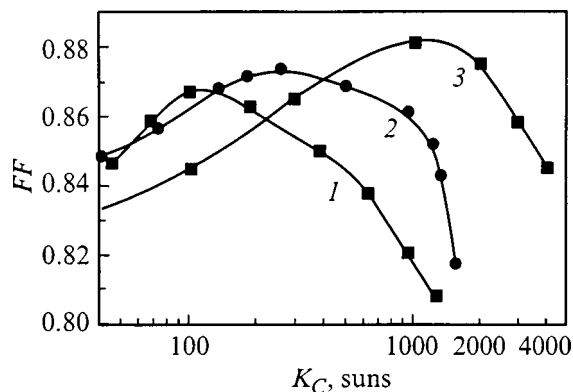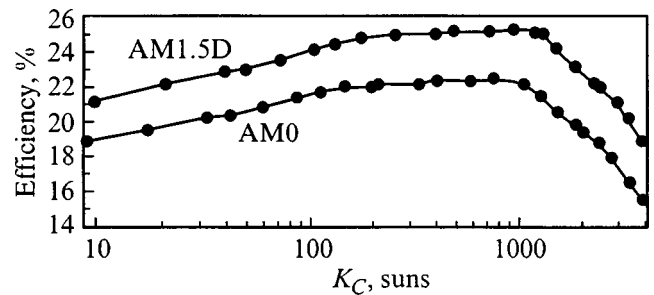*E-mail: andreev@scell.ioffe.rssi.ru



FIG. 3. The filling factor *FF* of the load characteristic as a function of illumination intensity for solar cells with different distances between contact strips: 0.1 $\mu$m (*1, 2*) and 0.04 $\mu$m (*3*) and with different areas of the active illuminated surface: 6 mm$^2$ (*1*) and 2 mm$^2$ (*2, 3*).

[1] S. M. Vernon, S. P. Tohin, V. E. Haven, L. M. Geoffroy, and M. M. Sanfacon, in *Proceedings of the 22nd IEEE Photovoltaic Specialists Conference*, Las Vegas (1991), p. 353.
[2] V. M. Andreev, A. B. Kazantsev, V. P. Khvostikov, E. V. Paleeva, V. D. Rumyantsev, and M. Z. Shvarts, in *Proceedings of the First World Conference on Photovoltaic Energy Conversion* (Hawaii, 1994) p. 2096.
[3] S. C. Bailey and D. J. Flood, Prog. Photovoltaics **6**, 1 (1998).
[4] J. C. Maroto, A. Marti, C. Algora, and C. L. Araujo, in *Proceedings of the 13th EPSEC*, Nice, France (1995), p. 343.
[5] V. M. Andreev, A. B. Kazantsev, V. P. Khvostikov, E. V. Paleeva, S. V. Sorokina, and V. D. Rumyantsev, Mater. Chem. Phys. **45**, 130 (1996).
[6] V.M. Andreev and V.D. Rumyantsev, Sol. Energy Mater. Sol. Cells **44**, 319 (1996).
[7] V. M. Andreev, V. A. Grilikhes, and V. D. Rumyantsev, *Photovoltaic Conversion of Concentrated Sunlight*, John Wiley & Sons Ltd. (1997), Ch. 3.

Translated by D. H. McNeill

# Parabolic negative magnetoresistance in $p$-Ge/Ge$_{1-x}$Si$_x$ heterostructures

Yu. G. Arapov,* G. I. Harus, V. N. Neverov, and N. G. Shelushinina

*Institute of Metal Physics, Ural Branch of RAS, 620219 Ekaterinburg, GSP-170, Russia*

O. A. Kuznetsov

*Scientific-Research Physicotechnical Institute of State University, 603600 N Novgorod, Russia*

Quantum corrections for the conductivity due to the weak localization (WL) and the disorder-modified electron-electron interaction (EEI) are investigated for the high-mobility multilayer $p$-Ge/Ge$_{1-x}$Se$_x$ heterostructures at $T=(0.1-20)$ K in magnetic field $B$ up to 1.5 T. Negative magnetoresistance with logarithmic dependence on $T$ and linear in $B^2$ is observed for $B \geqslant 0.1$ T. Such a behavior is attributed to the connection between the classical cyclotron motion and the EEI effect. The Hartree part of the interaction constant is estimated ($F_\sigma = 0.44$) and the WL and EEI contributions to the total quantum correction $\Delta\sigma$ at $B=0$ are separated ($\Delta\sigma_{\mathrm{WL}} \approx 0.3\Delta\sigma$; $\Delta\sigma^{ee} \approx 0.7\Delta\sigma$). © *1999 American Institute of Physics.*
[S1063-7826(99)01209-0]

## 1. INTRODUCTION

The diffusive nature of electron motion in disordered conductors results in quantum corrections for the effects with nontrivial dependences on temperature $T$ and magnetic field $B$.[1,2] These corrections are of the order of $(k_F l)^{-1}$, where $k_F$ is the Fermi wave vector, and $l$ is the mean free path. The total quantum correction for the Drude conductivity consists of the single-particle weak localization part and the part due to the disorder-modified electron-electron (e–e) interaction between particles with similar momenta and energies (in the diffusion channel) and between particles with small total momentum (in the Cooper channel). For two-dimensional (2D) system all three quantum corrections, i.e., localization, e–e interaction in the diffusion channel, and e–e interaction in the Cooper channel lead to the logarithmic low-temperature dependence of the conductivity at $B=0$:

$$\Delta\sigma(T) = \frac{e^2}{2\pi^2\hbar} \left\{ \left[ p + \left(1 - \frac{3}{4}F\right) - (p-1)\beta(T) \right] \times \ln\frac{kT\tau}{\hbar} - \ln\frac{\ln(kT_c\tau/\hbar)}{\ln(T_c/T)} \right\}. \quad (1)$$

The first term in square brackets in Eq. (1) is associated with the weak localization, the second term is a quantum correction due to the e–e interactions (EEI) in the diffusion channel, and the third term is the Maki–Thomson correction. The second term in braces is a quantum corrections due to EEI in the Cooper channel.

The different quantum corrections may be separated by application of an external magnetic field since each quantum effect has its own range of characteristic magnetic fields.[3] In the absence of spin scattering the magnetoresistance associated with the weak localization is negative. For this effect there exist two characteristic fields: the field $B_\varphi$ of crossover from parabolic to logarithmic $B$-dependence of magnetore-

sistivity ($B_\varphi = \hbar c/4eL_\varphi^2$, and $L_\varphi$ is the inelastic scattering length) and the field $B_{\mathrm{tr}} = \hbar c/2el^2$, where the magnetic length is less than the elastic scattering length. For the effect in the Cooper channel the characteristic field $B_{\mathrm{int}}$ is the field where the magnetic length is less than the coherence length $L_T$.

The localization effect is totally suppressed for field $B \gg B_{\mathrm{tr}} = \hbar c/2el^2$, where the magnetic length is less than the elastic mean free path $l$.[4] In this range, the only quantum correction for the conductivity is from EEI in the diffusion channel. In contrast with the $B$ sensitivity of the WL effect, the calculation for the EEI in the absence of spin effects[5–7] demonstrates that

$$\Delta\sigma_{xx} \equiv \Delta\sigma^{ee} - (e^2/2\pi^2\hbar)g\ln(kT\tau/\hbar), \quad \Delta\sigma_{xy} = 0, \quad (2)$$

irrespective of the strength of the applied magnetic field. Here $\tau$ is the elastic relaxation time and the interaction constant $g = (1 - F_\sigma)$, where the first universal term is due to the exchange (Fock) part and the second term ($F_\sigma$) is related to the direct (Hartree) part of the Coulomb repulsion.

Inverting the conductivity tensor[8] in the presence of EEI corrections, we have the following expression for the magnetoresistivity:

$$\rho_{xx}^{ee}(B) = 1/\sigma_0 + [1 - (\omega_c\tau)^2]\Delta\sigma^{ee}/\sigma_0^2, \quad (3)$$

where $\sigma_0$ is the Drude conductivity, and $\omega_c$ is the cyclotron frequency. The consequence of Eq. (3) is twofold: irrespective of temperature $\rho_{xx}(B) = 1/\sigma_0$ for $\omega_c\tau = 1$ and the link between the classical cyclotron motion and the EEI effect in the diffusion channel accounts for the parabolic negative magnetoresistance with logarithmic temperature dependence:

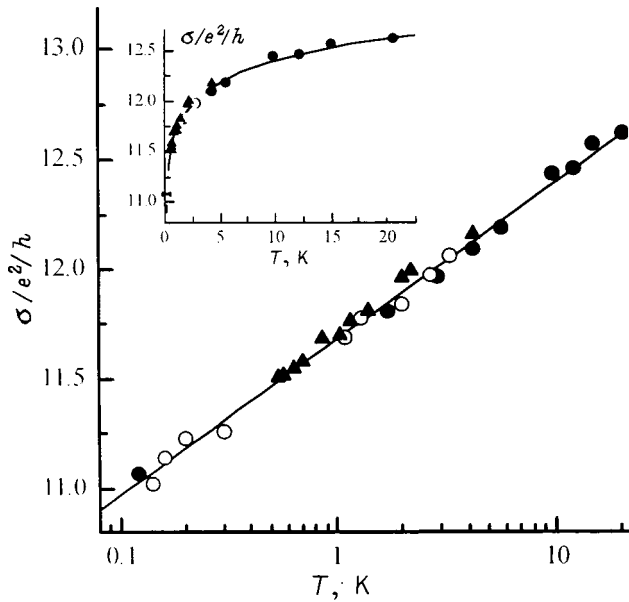$$\rho_{xx}^{ee}(B) - \rho^{ee}(0) = -(\omega_c\tau)^2\Delta\sigma^{ee}/\sigma_0^2 \approx -B^2\ln T.$$

FIG. 1. Temperature dependence of the conductivity at $B = 0$.

## 2. EXPERIMENTAL RESULT AND DISCUSSION

We have investigated the conductivity and magnetoresistance of strained multilayer p-type $Ge/Ge_{1-x}Si_x$ ($x = 0.03$) heterostructures with the hole densities $p = (2.4-2.6) \times 10^{11}\,cm^{-2}$ and mobilities $\mu = (1.0-1.7) \times 10^4\,cm^2/(V \cdot s)$, ($k_F l \geqslant 10$), on Ge layers at $T \geqslant 0.1$ K in magnetic fields up to 1.5 T. The conductivity at $B = 0$ varies as the logarithm of $T$ in a wide temperature range $(0.1-20.0)$ K (Fig. 1). For $B$ perpendicular to Ge layers the negative magnetoresistance is observed in a whole range of magnetic fields up to $\omega_c \tau = 1$ (Fig. 2). Because of the high mobility of holes, only a small
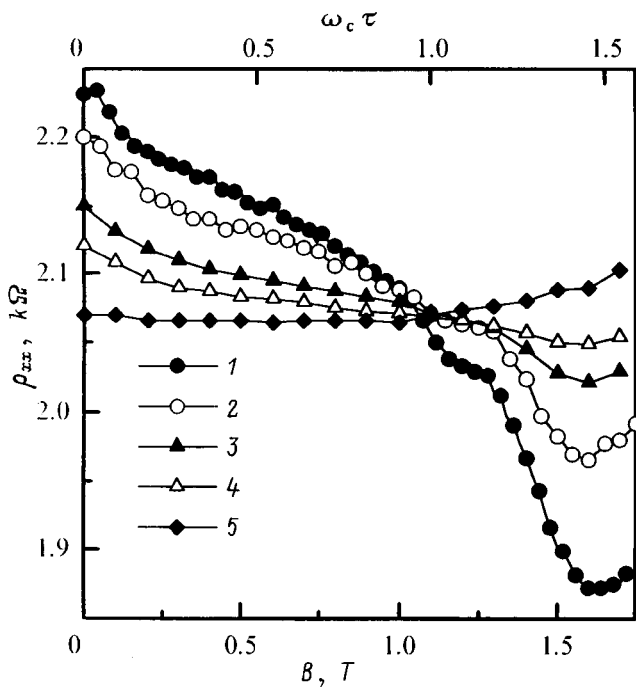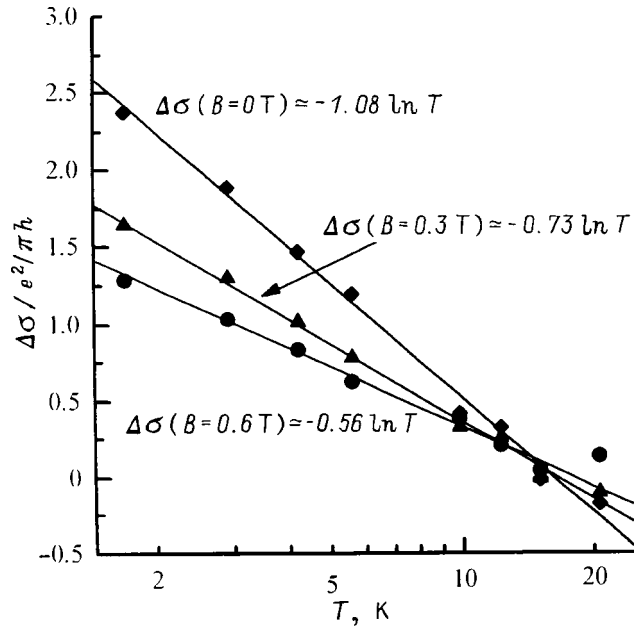


FIG. 3. Plot of the quantum correction for the conductivity $\Delta\sigma$ vs $\ln T$ for different magnetic fields.

magnetic field $B_{tr} = 0.03$ T is needed to suppress the effect of weak localization. The logarithmic dependence of $\Delta\sigma$ on $T$ at $B \gg B_{tr}$ (Fig. 3) unambiguously is evidence of the EEI quantum corrections. Figure 4 demonstrates that at $B \geqslant 3B_{tr}$ the magnetoresistance is parabolic. The intersection point of curves for different $T$ at $\omega_c \tau \cong 1$ is also observed (see Fig. 2).

The extrapolation of $B^2$ dependences to $B = 0$ according to Eq. (2) gives the values of $\rho^{ee}(0) = 1/\sigma_0 + \Delta\sigma^{ee}/\sigma_0^2$ for each $T$. From the universal value of $\rho_{xx}$ at $\omega_c \tau = 1$ we have



FIG. 2. The resistivity $\rho_{xx}$ as a function of the magnetic field at $T$, K: *1*—0.3, *2*—0.5, *3*—2.2, *4*—4.2, *5*—12.2.
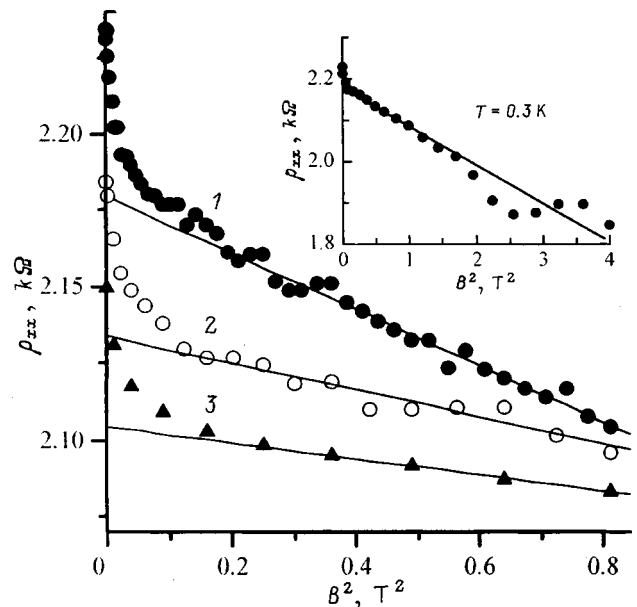


FIG. 4. The resistivity $\rho_{xx}$ against $B^2$ from $B = 0$ to $B = 0.9\,T$ for $T$, K: *1*—0.3, *2*—1.4, *3*—2.2. The solid line is the extrapolation of the $B^2$ dependence to zero field. The inset shows one of the curves (for $T = 0.3$ K) at higher magnetic fields ($\omega_c \tau > 1$).

the Drude conductivity $\sigma_0 = 12.4e^2/h$. Then, in accordance with Eq. (2), the Hartree interaction constant is estimated to be $F_\sigma = 0.44$. Because of the transparent parabolic $\rho_{xx}(B)$ dependence in a wide range of magnetic fields, where the EEI contribution is dominant, the separation of WL and EEI parts of the total quantum correction $\Delta\sigma$ at $B=0$ is possible. The result for our structures is that $\Delta\sigma^{ee} \approx 0.7\Delta\sigma$ and $\Delta\sigma_{WL} \approx 0.3\Delta\sigma$.

## 3. CONCLUSION

We have observed a large negative magnetoresistance of a high-mobility 2D hole gas in $p$-$Ge_{1-x}Si_x$/$Ge$/$Ge_{1-x}Si_x$ quantum wells. We find that a negative magnetoresistance is proportional to $B^2$ and has a logarithmic temperature dependence. We attribute this behavior to the link between the classical cyclotron motion and the EEI corrections for the conductivity in the diffusion channel (exchange and Hartree contributions). Our sample parameters indicate that the weak localization and the EEI in the Cooper channel effects are totally suppressed in this field and temperature regime ($B_\varphi < 3 \times 10^{-4}$ T, $B_{int} < 0.03$ T and $T < 2.2$ K) but the Zeeman splitting is not effective, $B_s > 1$ T, above 2 K. We find good agreement between theory and experiment on the value of the Hartree interaction constant.

\*E-mail: arapov@imp.uran.ru

[1] P. A. Lee and T. V. Ramakrishman, Rev. Mod. Phys. **57**, 287 (1985).
[2] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disorder Systems* (Amsterdam, 1985), p. 1.
[3] Yu. G. Arapov, V. N. Neverov, G. I. Harus, and N. G. Shelushinina, in *Materials of All-Russian Workshop*, 10–13 March, (N. Novgorod, 1998), p. 28 [Isvestiya RAN, ser. Fizicheskaya, 1999 (in press)].
[4] M. I. Dyakonov, Solid State Commun. **92**, 711 (1994).
[5] A. Houghton, J. R. Senna, and S. C. Ying, Phys. Rev. B **25**, 6468 (1982).
[6] S. M. Girvin, Phys. Rev. B **26**, 165 (1982).
[7] K. K. Choi, D. C. Tsui, and S. C. Palmaateer, Phys. Rev. B **33**, 8216 (1986).
[8] W. Poirier, D. Mailly, and M. Sanquer, Cond-mat/9706287.

# Role of thermal ejection of carriers in the burning of spatial holes in quantum dot lasers

L. V. Asryan*) and R. A. Suris**)

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

The role of thermal carrier ejection from quantum dots and free carrier diffusion in the burning of spatial holes in semiconductor quantum dot lasers is analyzed. The balance of the spatially inhomogeneous population inversion in the longitudinal direction of the cavity is shown to be controlled by thermal ejection from quantum dots. Because of this circumstance, hole burning in quantum dot lasers can show up more strongly and the threshold for multimode lasing can be lower than in semiconductor lasers with three-dimensional active regions or quantum-well lasers. The threshold for multimode lasing is determined as a function of the dispersion in the quantum dot size, cavity length, and temperature for structures that have been optimized to minimize the threshold current density of the fundamental mode. © *1999 American Institute of Physics.* [S1063-7826(99)01309-5]

Quantum dot (QD) semiconductor lasers have attracted increasing interest because of their expected advantages over ordinary semiconductor lasers with three-dimensional active regions and quantum-well (QW) lasers.[1] A theory has been developed for the threshold currents and characteristic temperature of QD lasers that takes into account inhomogeneous line broadening due to the spread in the parameters of the QD (e.g., its size).[2–5] The optimum parameters of the laser structure, with minimization of the threshold current density, were calculated as functions of the surface QD concentration, the dispersion of the QD size, losses (cavity length), and temperature.

In this paper we study the effect of spatial hole burning and multimode lasing in QD lasers. As in conventional semiconductor lasers with a three-dimensional active region and QW lasers (as well as in solid state lasers),[6,7] spatial hole burning in QD lasers is caused by nonuniform stimulated carrier recombination in the longitudinal direction in the cavity. Because light emitted beyond the lasing threshold is a standing wave in the cavity, stimulated carrier recombination will be most intense in QDs lying near the intensity antinodes of the emitted light and least intense in those lying near the nodes (Fig. 1). This causes an emptying of the quantum dots lying near the antinodes and overfilling of the quantum dots that lie near the nodes. This process can cause lasing in high-order longitudinal cavity modes, along with the fundamental mode.

The question of multimoding in semiconductor lasers is pressing, in both fundamental and practical terms. Studies of the physical processes controlling the multimode lasing threshold are needed in order to determine ways of suppressing additional modes and creating single mode lasers.

In lasers with three-dimensional active regions and QW lasers, diffusion smoothes out spatial nonuniformities in the distributions of the carriers and population inversion along the cavity, completely or partly suppressing the burning of holes in the spatial distribution of the carriers.[6,7]

The situation is fundamentally different in QD lasers. In

them, carrier diffusion plays an analogous, but minor role. In fact, only carriers localized in the QD contribute to stimulated emission. Free carriers also exist in the optical clipping region of the laser and they contribute to spontaneous emission, thereby increasing the threshold current density.[2–5] Thus, balancing a spatially nonuniform population inversion requires thermal ejection of carriers from the QD, as well as longitudinal diffusion of free carriers. The slower of the two processes (thermal ejection and diffusion) will control the spatial distribution of carriers in QD lasers.

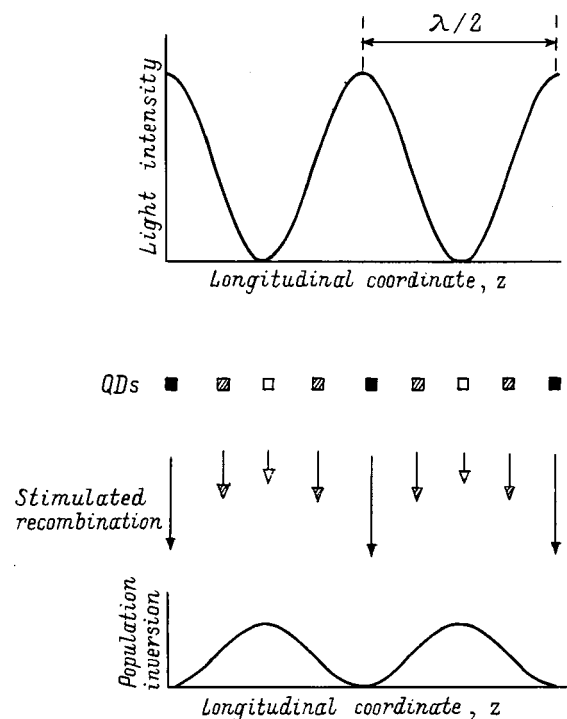Since the characteristic times for thermal ejection of car-



FIG. 1. Distributions of the stimulated emission intensity and population inversion along the cavity.

riers from a QD (see the formulas below) are usually much longer than those for diffusion over half the wavelength of the light in the medium (the characteristic scale length of the nonuniformity), smoothing of a spatially nonuniform population inversion can be strongly suppressed in a QD laser. Thus, spatial hole burning may show up more strongly.

Thus, thermal ejection of carriers from QDs, in contrast with the diffusion, controls the smoothing out of the spatial nonuniformities in the population inversion in QD lasers. The finite characteristic times for thermal ejection determine the threshold for multimoding in QD lasers. A similar situation has been examined for ''band–impurity center'' laser transitions.[8,9]

The multimode lasing threshold $\delta j$ is defined as the excess in the pump current density (injection) over the threshold current for the fundamental (nearest to the maximum in the gain spectrum) mode required for generation of the next longitudinal cavity mode. An analysis of this question yields the following expression for $\delta j$:

$$\delta j = j_2 - j_{\text{th}} = 2 \frac{|\delta g|}{g^{\text{max}}} \frac{e N_s^{\text{min}}}{\tau_n(1-f_n) + \tau_p(1-f_p)}, \qquad (1)$$

where $j_{\text{th}}$ and $j_2$ are the threshold current densities for the fundamental and next longitudinal modes, respectively, of the cavity, $g^{\text{max}}$ is the maximum in the gain spectrum, $f_{n,p}$ are the average (over the ensemble of QDs) degree of filling of the electron and hole quantum-well levels in the QD required for lasing in the fundamental mode,[2–5] $N_s^{\text{min}} = (4/\xi)$ $\times (\sqrt{\epsilon}/\lambda_0)^2 \tau_{\text{QD}} [(\Delta \varepsilon)_{\text{inhom}}/\hbar] \beta(a/\Gamma)$ is the minimally permissible surface concentration in the QD required to achieve lasing in the fundamental mode for given losses $\beta$ (a given cavity length $L$) and inhomogeneous line broadening $(\Delta \varepsilon)_{\text{inhom}}$,[2,3,5] $\xi$ is a numerical constant in the QD size distribution function ($\xi = 1/\sqrt{2\pi}$ and $\xi = 1/\pi$ for Gaussian and Lorentzian distributions, respectively), $\lambda_0$ is the vacuum wavelength of the fundamental mode, $a$ is the average QD size, and $\Gamma$ is the optical clipping coefficient in the layer with the QD (in the transverse direction in the cavity).[2]

The characteristic times for thermal ejection of electrons and holes in Eq. (1) are given by[2]

$$\tau_n = \frac{1}{\sigma_n v_n n_1}, \qquad \tau_p = \frac{1}{\sigma_p v_p p_1}, \qquad (2)$$

where $\sigma_{n,p}$ are the cross sections for capture of electrons and holes in a QD and $v_{n,p}$ are the thermal speeds. In Eq. (2), $n_1 = N_c^{\text{OCL}} \exp[-(\Delta E_c - \varepsilon_n)/T]$ and $p_1 = N_v^{\text{OCL}} \exp[-(\Delta E_v - \varepsilon_p)/T]$, where $N_{c,v}^{\text{OCL}} = 2(m_{c,v}^{\text{OCL}} T/2\pi\hbar^2)^{3/2}$ are the effective densities of states in the conduction and valence bands of the material in the optical clipping region, $\Delta E_c$ and $\Delta E_v$ are the gap widths in the conduction and valence bands at the heteroboundary between the QD and the optical clipping region, $\varepsilon_{n,p}$ are the quantum-well energies of electrons and holes in a QD of average size $a$, and $T$ is the temperature (in energy units).

The absolute magnitude of the difference in the gain coefficients for the fundamental and next cavity modes in Eq. (1) is given by

$$|\delta g| = \frac{1}{2} \left| \frac{\partial^2 g}{\partial E^2} \right| (\delta E)^2 = \frac{1}{2} \left| \frac{\partial^2 g}{\partial E^2} \right| \left( \hbar \frac{c}{\sqrt{\epsilon}} \frac{\pi}{L} \right)^2. \qquad (3)$$

The derivative in Eq. (1) is taken at $E = E_0$, where $E_0$ is the photon energy in the fundamental mode. The difference between the photon energies of neighboring longitudinal modes ($\Delta m = \pm 1$) of a Fabry–Perot cavity is $\delta E = \hbar(c/\sqrt{\epsilon})(\pi/L)$, where $c$ is the speed of light in vacuum, $\epsilon$ is the electric permittivity of the optical clipping region, and $L$ is the cavity length.

For a Gaussian distribution of the relative fluctuations in the QD size,

$$\frac{|\delta g|}{g^{\text{max}}} = \frac{1}{2} \left( \frac{\hbar(c/\sqrt{\epsilon})(\pi/L)}{(\Delta \varepsilon)_{\text{inhom}}} \right)^2, \qquad (4)$$

where $(\Delta \varepsilon)_{\text{inhom}} = (q_n \varepsilon_n + q_p \varepsilon_p) \delta$ is the inhomogeneous line broadening due to the spread in the QD sizes, $q_{n,p} = -(\partial \ln \varepsilon_{n,p}/\partial \ln a)$, and $\delta$ is the mean-square deviation of the relative fluctuations in the QD size.[2]

The threshold current density for the fundamental mode is given by[2–5]

$$j_{\text{th}} = \frac{e N_s}{\tau_{\text{QD}}} f_n f_p + e b B n_1 p_1 \frac{f_n f_p}{(1-f_n)(1-f_p)}, \qquad (5)$$

where $N_s$ is the surface concentration of QDs, $\tau_{\text{QD}}$ is the radiative recombination time in a QD,[2,3,5] $b$ is the thickness of the optical clipping region, and $B$ is the radiative recombination constant in the optical clipping region.

The relative threshold for multimode lasing is

$$\frac{\delta j}{j_{\text{th}}} = 2 \frac{|\delta g|}{g^{\text{max}}} \frac{N_s^{\text{min}}}{N_s} \frac{\tau_{\text{QD}}}{\tau_n(1-f_n) + \tau_p(1-f_p)}$$

$$\times \frac{1}{f_n f_p + (\tau_{\text{QD}}/N_s) b B n_1 p_1 [f_n f_p/(1-f_n)(1-f_p)]}. \qquad (6)$$

In the following we shall limit the discussion to the case of charge neutrality in the QD, where the degree of filling of the quantum-well electron and hole levels is given by[2,3,5]

$$f_n = f_p = \frac{1}{2} \left( 1 + \frac{N_s^{\text{min}}}{N_s} \right). \qquad (7)$$

It is evident from Eqs. (1), (2), and (6) that $\delta j$ and $\delta j/j_{th}$ increase with increasing $\sigma_{n,p}$. Naturally, the larger $\sigma_{n,p}$ is, the more intense the capture processes in the QD will be and the more weakly the spatial hole burning will manifest itself.

The results of these calculations will now be illustrated for an GaInAsP/InP laser heterostructure.[2–5] The thickness of the optical clipping region is $b = 0.28 \, \mu\text{m}$. The (power) reflectivities of the mirrors are $R_1 = R_2 \approx 0.33$. A Gaussian distribution of the relative fluctuations in the QD sizes is assumed. The average size of the cubic QDs is $a = 150$Å. The surface concentration of QDs, the mean-square deviation in the relative fluctuations in the QD size, the cavity length, and temperature are $N_s = 6.1 \times 10^{10} \, \text{cm}^{-2}$, $\delta = 0.025$ (5%), $L = 500 \, \mu\text{m}$, and $T = 300$ K, respectively, unless otherwise indicated. The corresponding minimum permissible surface
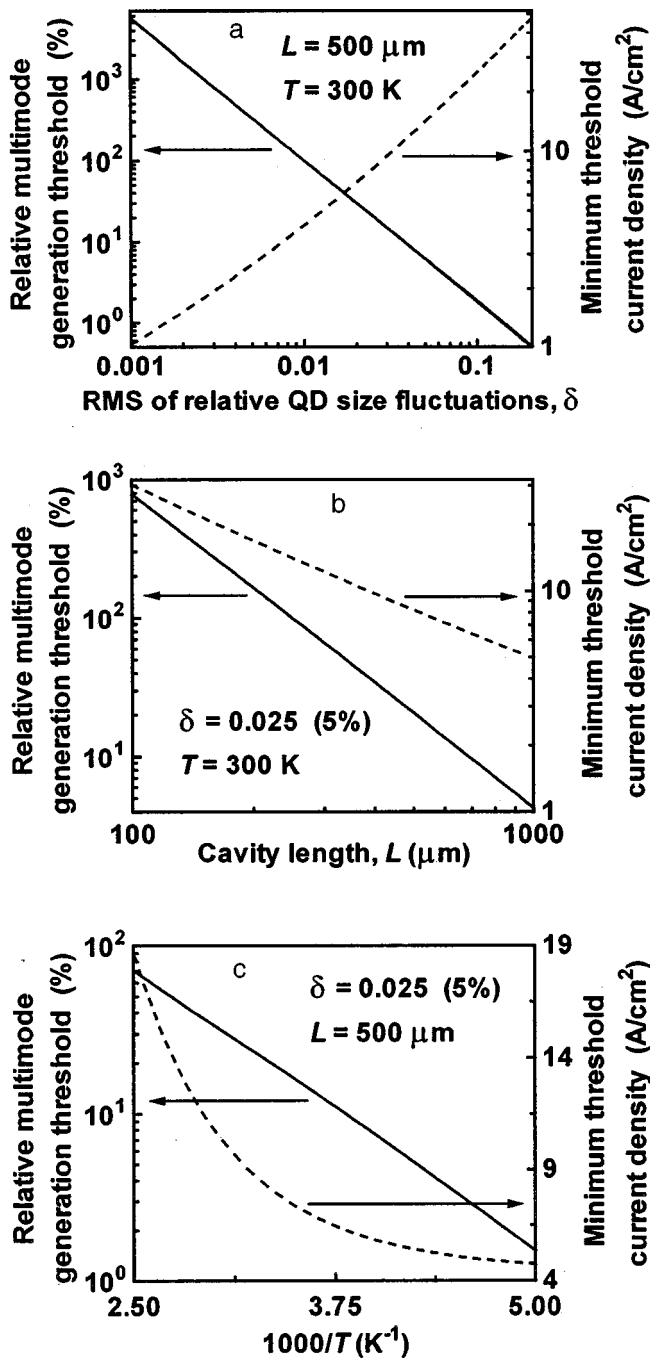
FIG. 2. Relative threshold for single mode lasing $\delta j / j_{th}$ (smooth curves) and the minimum threshold current density for the fundamental mode (dashed curves) as functions of (a) the average root-mean-square (RMS) deviation of the relative fluctuations in the QD size, (b) cavity length, and (c) temperature. Each point on the curves corresponds to a specific structure optimized for the given value of $\delta$ (a), $L$ (b) or $T$ (c), respectively.

It has been shown[2,5] that the optimum parameters for a laser structure (surface concentration of quantum dots and thickness of the optical clipping layer), such that the threshold current density for the fundamental mode is minimal, depend on the mean-square deviation in the relative fluctuations in the QD size, the cavity losses (i.e., cavity length), and temperature. Some calculations for structures that have been optimized to minimize the threshold current density for the fundamental mode are shown below.[2,3,5]

Figure 2 shows plots of the relative threshold for multimode laser generation (smooth curves) and the minimum threshold current density for the fundamental mode (dashed curves) as functions of (a) the mean-square deviation in the relative fluctuations in the QD size, (b) the cavity length, and (c) temperature. Each point on the curves corresponds to a concrete structure optimized for the given value of $\delta$ (Fig. 2a), $L$ (Fig. 2b), or $T$ (Fig. 2c).

Figure 2a shows that reducing the spread in the QD size leads both to a drop in the minimum threshold current density and to a substantial rise in the threshold for multimode lasing at a QD. For structures optimized at $\delta = 0.025$ (spread 5%) and 0.1 (spread 20%), $\delta j / j_{th} \approx 21$ and 2%, respectively. The corresponding minimum threshold current densities are 8 and 25 A/cm$^2$.

Increasing the cavity length leads to a desirable reduction in the minimum threshold current, but at the same time, to an undesirable reduction in the threshold for multimoding (Fig. 2b).

The reduction in the minimum threshold current density that occurs as the temperature is lowered is accompanied by a simultaneous drop in the multimoding threshold (Fig. 2c). Naturally, thermal ejection of carriers from a QD and, therefore, balancing of the spatially nonuniform population inversion are suppressed at low temperatures. Thus, the multimoding threshold is also low at low temperatures.

In this paper we have calculated the threshold for multimode lasing in a QD laser. It has been shown that the multimoding threshold is determined by the finite characteristic times for thermal ejection of carriers from a QD. A reduction in the spread in the QD sizes leads both to a drop in the minimum threshold current density and to a substantial rise in the multimoding threshold. The reduction in the minimum threshold current density that occurs as the temperature is lowered is accompanied by a simultaneous drop in the multimoding threshold.

*)E-mail: asryan@theory.ioffe.rssi.ru
**)E-mail: suris@theory.ioffe.rssi.ru

concentration of QDs[2,3,5] (controlled by the values of $\delta$, $L$, and $b$) required to obtain lasing in the fundamental mode is $N_s^{min} = 2.1 \times 10^{10}$ cm$^{-2}$.

Calculating $\sigma_{n,p}$ lies beyond the scope of this article. Here we have taken $\sigma_{n,p} = 10^{-13}$ cm$^2$ as an estimate for the multimode threshold (considerably smaller than the geometric cross section of a QD). It should be noted that the way the multimoding threshold depends on the structure parameters and temperature does not change as $\sigma_{n,p}$ is varied.

[1] Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).
[2] L. V. Asryan and R. A. Suris, Semicond. Sci. Technol. **11**, 554 (1996).
[3] L. V. Asryan and R. A. Suris, IEEE J. Sel. Top. Quantum Electron. **3**, 148 (1997).
[4] L. V. Asryan and R. A. Suris, Electron. Lett. **33**, 1871 (1997).
[5] L. V. Asryan and R. A. Suris, IEEE J. Quantum Electron. **34**, 841 (1998).

[6] H. Statz, C. L. Tang, and J. M. Lavine, J. Appl. Phys. **35**, 2581 (1964).

[7] C. L. Tang, H. Statz, and G. deMars, J. Appl. Phys. **34**, 2289 (1963).

[8] R. A. Suris, and S. V. Shtofich, Fiz. Tekhn. Poluprovodn. **16**, 551 (1982) [Sov. Phys. Semicond. **16**, 851 (1982)].

[9] R. A. Suris and S. V. Shtofich, Fiz. Tekhn. Poluprovodn. **17**, 538 (1983) [Sov. Phys. Semicond. **17**, 859 (1983).

Translated by D. H. McNeill

# Polarization anisotropy in optical reflection spectra of structures with open nanowires

N. S. Averkiev, S. O. Kognovitskii,*) and V. V. Travnikov

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

The optical reflection spectra of structures with open nanowires created from ZnCdSe/ZnSe heterostructures are investigated. A significant polarization anisotropy of the reflection is observed over a wide spectral range, including the exciton resonance regions of the wire, the barrier, and the substrate. The effect of additional laser pumping on the reflection spectrum is examined. The observed effects are interpreted in terms of a model of an effective anisotropic layer. A numerical simulation of the reflection spectra based on this model is in good qualitative agreement with experiment. © *1999 American Institute of Physics.* [S1063-7826(99)01409-X]

Semiconducting structures with surface lattices formed by a system of open (without growth over them) wires is characterized by a large difference between the dielectric permittivities of the wire material and the air in the gaps between them. This causes a spatial redistribution of the amplitudes of incident electromagnetic waves along the cross section of the lattice. This sort of redistribution is substantially different for waves polarized along (TE-waves) and perpendicular to (TM-waves) the direction of the wires.[1–3] To a great degree, this determines the strong polarization anisotropy in the optical properties of structures with open wires. This mechanism for formation of a polarization anisotropy is substantially more important than one-dimensional clipping when effective exciton quantization has still not been achieved. We call such wires nanowires below. A strong polarization anisotropy has been observed in the luminescence and Raman scattering spectra of these structures before.[4–6] In this paper we study the polarization anisotropy in the optical reflection spectra of structures with open ZnCdSe/ZnSe nanowires.

The test samples were produced by interference lithography followed by reactive ion etching[7] from undoped structures containing a single $Zn_{1-x}Cd_xSe$ ($x = 16\%$) quantum well of thickness 5 nm enclosed between 20- and 25-nm ZnSe barriers. The initial structures were grown on a substrate of GaAs with (100) orientation. In the resulting samples, the ZnSe/ZnCdSe/ZnSe wires, with a width $a = 70$ nm and height $b = 60$ nm, were positioned on the substrate with a period of $L = 250$ nm.

The experiments were done with sample temperatures $T = 2$, 77, and 300 K. The reflection spectra were measured with strictly normal incidence of a collimated (0.5°) beam of white light linearly polarized parallel (TE) or perpendicular (TM) to the direction of the wires. The effect of additional irradiation on the reflection spectra was studied using various laser lines with similar excitation energy fluxes ($\sim 10$ W/cm²).

The reflection spectra $R_\perp(h\nu)$ and $R_\parallel(h\nu)$ obtained for TM- and TE-waves, respectively, are shown in Fig. 1. The figure shows that $R_\perp$ is significantly greater than $R_\parallel$ over a wide spectral range. The ratio $R_\perp/R_\parallel$ is as high as 2.5 in the interval $2.48 - 2.75$ eV.

Additional photogeneration of carriers within the volume of the wires by 2.807-eV laser light causes $R_\perp$ to increase by
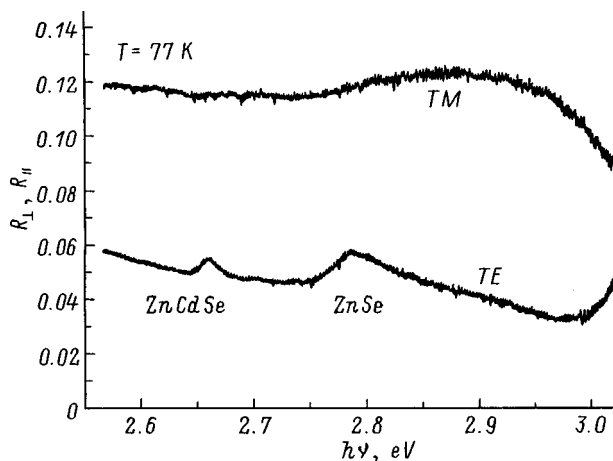


FIG. 1. Reflection spectra of a structure with open nanowires for TM- and TE-waves in the neighborhood of the exciton resonances of the wires.
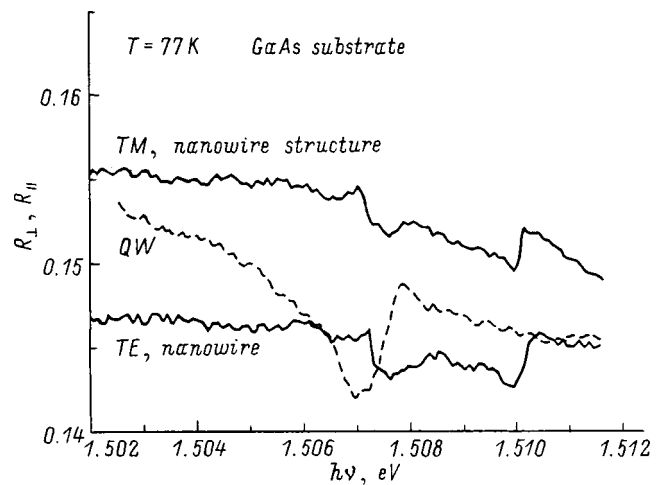


FIG. 2. Reflection spectra of a structure with open nanowires for TM- and TE-waves in the neighborhood of the exciton resonances of the substrate. The dashed curve is the reflection spectrum of an initial quantum well structure.

$\sim 15\%$ over the entire spectral range studied here but does not change the $R_{\parallel}$ spectrum (except for a narrow region near the exciton resonances of the wire). Additional irradiation by photons with energies less than the gap of the ZnCdSe/ZnSe structure has no effect on the reflection spectra. It should be noted that, in the case of an initial structure with a quantum well, additional photocarriers do not change the reflectivity within a wide spectral range. The different effect of laser irradiation on the reflectivities $R_{\perp}$ and $R_{\parallel}$ may be a consequence of different contributions of the boundary conditions at the side walls of the ZnSe-wires for TE- and TM-polarizations of the incident light. For TM-polarization, these boundary conditions determine the optical response of the wires to a significant degree. Thus, the change in the boundary conditions when the free carrier concentration is raised by screening the surface embedded electric fields will lead to a substantial change in the ''background'' reflectivity for this polarization. For TE-polarization, these boundary conditions have a significantly smaller effect.

It was possible to detect features associated with exciton resonances of the GaAs substrate in the reflection spectra. As can be seen in Fig. 2, these resonances are similar in the spectra for both TM- and TE-waves. Compared to the reflection spectrum from the initial quantum well structure, the spectra of the nanowire structure had an additional resonance feature at higher energies. This behavior probably originates in a periodic modulation of the elastic deformation along the substrate surface caused by a mismatch of the lattice parameters of the substrate and wire materials.

The resonance features corresponding to excitons of the ZnCdSe layer and the ZnSe barriers mostly show up only in the $R_{\parallel}$ spectrum. The amplitude of these resonances increases with additional laser pumping at photon energies exceeding the gap of the corresponding material. This is confirmed by Fig. 3, in which one of the spectra was measured while the specimen was being irradiated with 2.807-eV laser light. (Here the luminescence contribution to the detected signal was negligible.) The effect of additional photocarriers
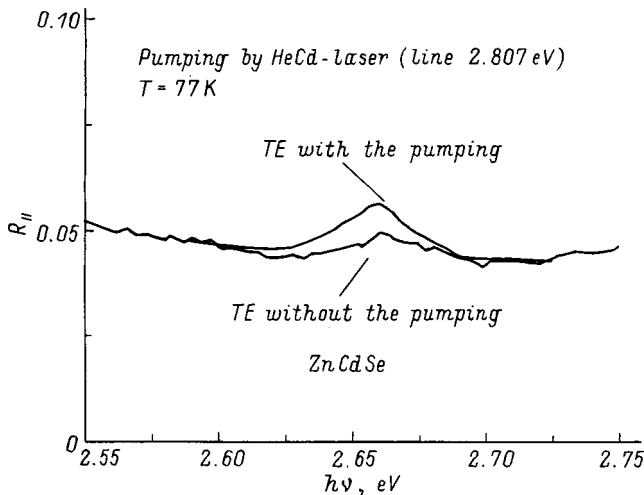
on the amplitude of the resonances is related to a reduction in exciton damping, which, in turn, is caused by a reduction in the embedded electric fields in the wires due to screening of surface charged states by photogenerated carriers or due to their neutralization as carriers are captured.

The polarization anisotropy of the reflection spectra can be described by a model that includes an effective anisotropic layer in structures with open nanowires. A short-period grating of semiconductor strips separated by air gaps can be regarded as a uniform anisotropic uniaxial layer that is characterized by effective refractive indices $n_0$ and $n_e$ for the ordinary (TE) and extraordinary (TM) waves, respectively,[8]

$$n_0^2 = an^2/L + (L-a)n_{\text{air}}^2/L,$$

$$1/n_e^2 = a/(Ln^2) + (L-a)/(Ln_{\text{air}}^2),$$

where $n$ is the refractive index of the semiconductor layer in the wire. The dielectric permittivity corresponding to a given $n$ is chosen to have a form that takes the exciton resonances into account, i.e.,

$$\varepsilon(E) = \varepsilon_0 + C(E_x)/(E_x^2 - E^2 - iE\Gamma_x),$$

where $\varepsilon_0$ is the background dielectric constant due to the high-energy bands, $E$ is the photon energy, $E_x$ is the energy of the exciton ground state, and $\Gamma_x$ is the homogeneous broadening of the exciton resonance. Inhomogeneous broadening was taken into account by averaging over the position of the exciton resonance. The Gaussian averaged profile for the ZnCdSe layer had a half-width of about 25 meV and that for a ZnSe barrier was about 47 meV. The coefficient $C$ is proportional to the square of the matrix element for the momentum $\langle p \rangle$ for interband transitions:

$$C(E_x) = (256\hbar^2 \varepsilon(0)^3 R^3 P^2)/(m_0 e^4 E_x),$$

where $\varepsilon(0)$ is the static dielectric constant, $R$ is the binding energy of the exciton ground state, and $P^2 = (2|\langle p \rangle|^2)/(3m_0)$.



FIG. 3. Reflection spectra of a structure with open nanowires for TE- waves in the neighborhood of the exciton resonance of the ZnCdSe. These spectra were obtained with and without additional illumination by 2.807-eV laser light.
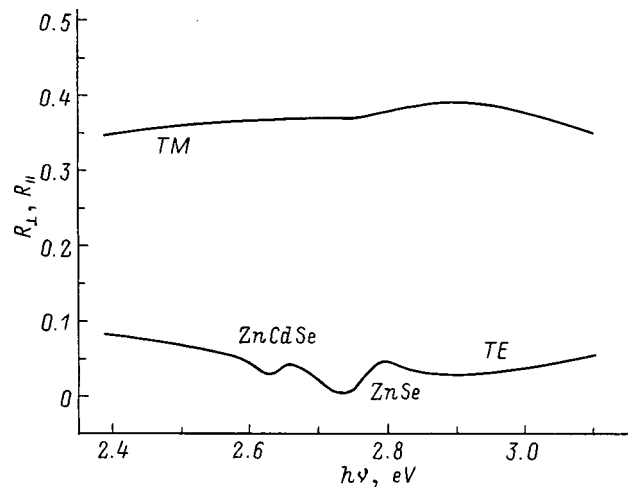


FIG. 4. Calculated reflection spectra of a structure with open nanowires for TM- and TE-waves in the neighborhood of the exciton resonances of the wires.

Expressions for the reflection coefficient $r$ for normally incident TE- and TM-waves were derived including reflection from three effective anisotropic layers containing strips of ZnSe, ZnCdSe, and ZnSe, respectively, and lying above a thick GaAs layer:

$$r = [r_{01} + t_{01}t_{10}/r_{01} + B \exp(-i4\pi n_1 h_1/\lambda)]/(1+B),$$

where

$$B = -r_{23} + t_{23}t_{32}/[-r_{23} + A \exp(i4\pi n_2 h_2/\lambda)],$$

and

$$A = r_{23} + t_{23}t_{32}/[r_{23} + \exp(-i4\pi n_3 h_3/\lambda)/r_{34}],$$

where $r_{jk}$ and $t_{jk}$ are amplitude reflection and transmission coefficients for the corresponding effective layer of thickness $h_j$, and $n_j$ is the effective refractive index of this layer for TM- or TE-waves.

The reflection spectra for the two polarizations were calculated using the actual structure parameters and tabulated data (for the ''background'' refraction and extinction coefficients of ZnSe and GaAs[9,10]). The quantities $C$, $\Gamma_x$, and $E_x$ were chosen for optimum matching of the calculated and experimental spectra. This made it possible to estimate the homogeneous broadening of the exciton resonances, which was about 20 meV for the ZnCdSe layer, as well as the parameter $P$ for bulk ZnSe, at $P^2 \approx 28.5$ eV. This value was close to published estimates.[11]

The fragment of a spectrum shown in Fig. 4 has all the characteristic features of the experimental spectra. The exciton resonance features corresponding to the ZnSe and ZnCdSe strips show up mainly in the $R_\parallel$ spectrum. Reducing the damping $\Gamma_x$ increases the amplitude of the resonance. The exciton resonances corresponding to GaAs substrate have the same amplitudes in the spectra for TM- and TE-waves.

$R_\perp$ is substantially greater than $R_\parallel$ over a wide range. The shape and amplitude of the $R_\perp$ spectrum is determined by reflection from the GaAs substrate. This is because the intensity of an electromagnetic wave polarized perpendicular to the direction of the wires is concentrated largely in the gap between the wires, so that there is an effective interaction with the substrate.

The slight quantitative difference between the theoretical and experimental spectra may be the result of the disregard, in our model, of the scattering of light on the surface roughness of the wires, diffraction of light into the substrate, and deviations of the wire cross sections from square cross section. In addition, we have ignored the anisotropy of the matrix elements of the optical transitions due to the different distributions of the amplitude of the electromagnetic field along the cross sections of the wires for TM- and TE-waves.

In conclusion, it should be noted that the successful interpretation of the polarization properties of the reflection spectra of structures with open nanowires in terms of an effective anisotropic layer model confirms the importance of taking into account the effect of the spatial distribution of the elements of a substantially inhomogeneous medium on the response of this medium to optical perturbations.

*)E-mail: kogn@spectr.ioffe.rssi.ru

[1] N. A. Gippius, S. G. Tikhodeev, and L. V. Keldysh, Superlattices Microstruct. **15**(4), 479 (1994).

[2] N. A. Gippius, S. G. Tikhodeev, A. Forchel, and V. D. Kulakovskii, Superlattices Microstruct. **16**(2), 165 (1994).

[3] U. Bockelmann, Europhys. Lett. **16**(6), 601 (1991).

[4] P. Ils, Ch. Greus, A. Forchel, V. D. Kulakovskii, N. A. Gippius, and S. G. Tikhodeev, Phys. Rev. B **51**(7), 4272 (1995).

[5] J. Rubio, Z. H. van Der Meulen, N. Mestres, J. M. Calleja, K. H. Wang, P. Ils, A. Forchel, N. A. Gippius, and S. G. Tikhodeev, Solid-State Electron. **40**, 707 (1996).

[6] N. V. Lomasov, V. V. Travnikov, S. O. Kognovitskiĭ, S. A. Gurevich, S. I. Nesterov, V. I. Skopina, M. Rabe, and F. Henneberger, Fiz. Tverd. Tela (St. Petersburg) **40**(8), 1559 (1998) [Phys. Solid State **40**, 890 (1998)]

[7] S. A. Gurevich, A. V. Kolobov, V. M. Lyubin, S. I. Nesterov, M. M. Kulagina, F. N. Timofeev, and S. I. Troshkov, Pis'ma Zh. Tekh. Fiz. **18**(17), 85 (1992) [Sov. Tech. Phys. Lett. **18**, 581 (1992)].

[8] P. Yeh, Opt. Commun. **26**, 289 (1978).

[9] Landolt–Börnstein. *Physics of II–VI and I–VII compounds, semimagnetic semiconductors*, III / 17b.

[10] Landolt–Börnstein. *Semiconductors: Intrinsic Properties of Group IV Elements and III–V, II–VI, and I–VII Compounds*, III / 22a.

[11] G. N. Aliev, A. D. Andreev, O. Coschug-Toates, R .M. Datsiev, S. V. Ivanov, S. V. Sorokin, and R. P. Seisyan, J. Cryst. Growth **184/185**, 857 (1998).

Translated by D. H. McNeill

# Photoluminescence of InAs quantum dots grown on disoriented GaAs substrates

G. V. Astakhov,*) V. P. Kochereshko, D. G. Vasil'ev, V. P. Evtikhiev, V. E. Tokranov,
I. V. Kudryashov, and G. V. Mikhaĭlov

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

The photoluminescence spectra in an external magnetic field of an ensemble of InAs quantum dots grown by molecular beam epitaxy on a (001) GaAs substrate with a disorientation in the [010] direction are studied. A redistribution of the photoexcited carriers among different groups of dots under the influence of the magnetic field is observed. The concentration of quantum dots is determined by analyzing the data. © *1999 American Institute of Physics.*
[S1063-7826(99)01509-4]

## 1. INTRODUCTION

Increasing interest of researchers has recently been concentrated on studies of structures with quantum dots and quantum filaments. Quantum dots (QDs), as the limiting case of carrier quantum wells, are extremely interesting to basic researchers. Since ideal QDs have a line emission spectrum, they also offer promise as low-threshold semiconductor lasers. One of the major problems in the technology of QDs is producing a uniform array of dots with a high density.[1] This problem is solved by using self-organization effects.[2] However, as the concentration of QDs increases, coalescence effects in which neighboring QDs merge into one are observed. This reduces the photoluminescence efficiency, since dislocations develop in these QD systems and become centers of nonradiative recombination.

A technique of growing InAs QDs using substrates of (001) GaAs that are disoriented in the [010] direction in order to raise the density and homogeneity of an array of QDs.[3] On surfaces of this sort, steps develop in the [110] and [$\bar{1}$10] directions. The intersection of these steps creates a network of terraces, with alternating wide and narrow regions that are bounded by steps on all sides. Quantum dots grown on surfaces of this sort lie predominantly on the wide regions of the terraces (Fig. 1).

## 2. EXPERIMENT

We have studied samples with InAs QDs grown on disoriented GaAs substrates (disorientation angle 4°) using self-organization effects. The photoluminescence spectrum of the QDs is shown in Fig. 2. The experiment was done at a temperature of 1.6 K and the photoluminescence was excited by He-Ne laser light with an intensity of 1 W/cm$^2$. In the spectrum two inhomogeneously broadened photoluminescence profiles, each with a Gaussian shape, can be distinguished. The peaks in these profiles are separated by $\sim 100$ meV. A detailed analysis of the photoluminescence line profiles is given elsewhere.[4] Two photoluminescence peaks show up because there are two QD arrays. The average size of the QDs in these arrays differs by $\sim 15\%$. One array, QD-1, is associated with quantum dots that grow in the cor-

ners of the terraces (position 1 in Fig. 1). The other array, QD-2, is associated with points that developed spontaneously in plane sections of the terraces (position 2 in Fig. 1). Because of the easier growth conditions and partial relaxation of the stress at the points on the boundary with another terrace,[3] the average size of QD-1 is greater than that of QD-2. Therefore, the photoluminescence maximum of the QD-1 array lies at a lower energy than that of QD-2. A comparison of the integrated intensities of the QD-1 and QD-2 arrays yields an estimate of the relative number of quantum dots grown at the corners and centers of the terraces: in these samples $n_1/n_2 \approx 1$.

The intensity of the QD photoluminescence is found to increase significantly in an external magnetic field (see the inset in Fig. 2). This behavior stems from the presence of another type of quantum dots, QD-3 (position 3 in Fig. 1),
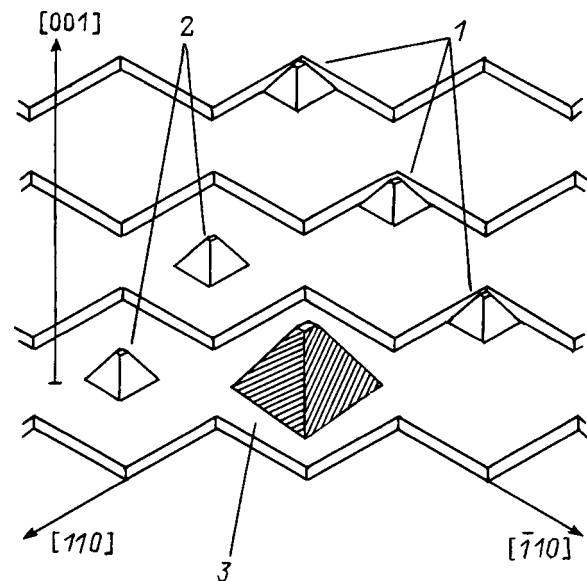


FIG. 1. A sketch of a structure with InAs quantum dots (pyramids), grown on a (001) GaAs substrate, disoriented in the [010] plane. Three possible types of quantum dots are shown. *1* and *2* correspond to quantum dots located in the corners and center of the terraces, respectively, and *3* to quantum dots formed as a result of coalescence.
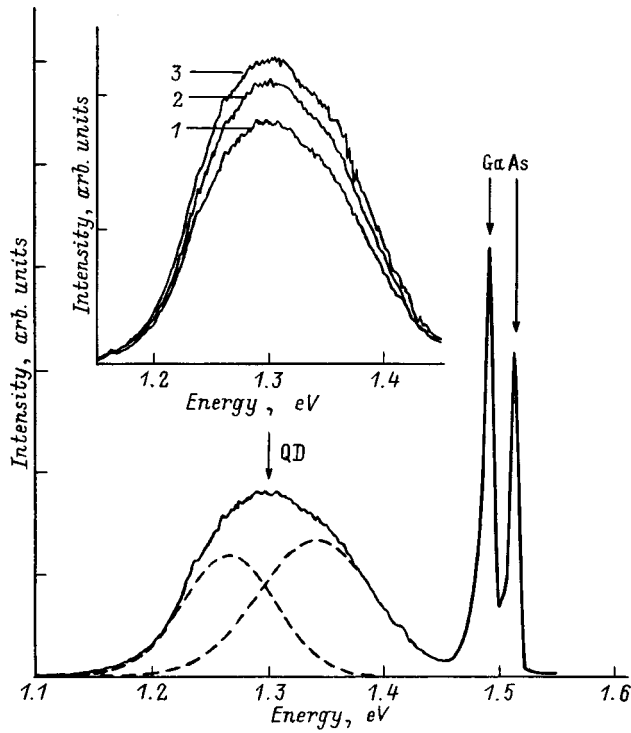
FIG. 2. Photoluminescence spectrum of quantum dots (QD) in zero magnetic field. The dotted curves are two Gaussian profiles used to approximate the photoluminescence spectrum. The narrow lines at an energy ~1.5 eV correspond to exciton recombination in bulk GaAs. The inset shows the photoluminescence spectra of quantum dots in magnetic fields $B=0$ (*1*), 0.9 (*2*), and 2 T (*3*).
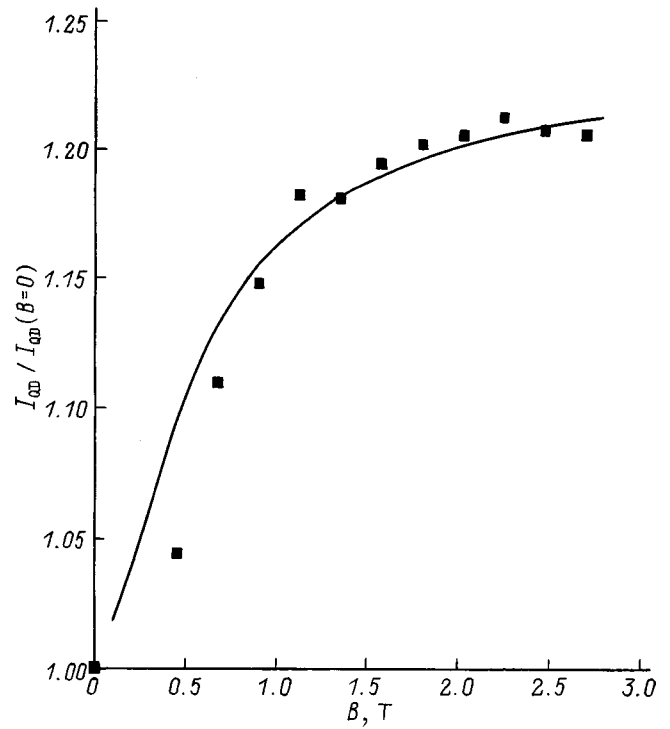


FIG. 3. Relative change in the photoluminescence intensity of quantum dots in a magnetic field, $I_{QD}/I_{QD}(B=0)$. The points are experimental data and the curve, a calculation according to Eq. (12).

which are formed through coalescence of QD-1 and QD-2. The QD-3 do not contribute to the photoluminescence because of strong nonradiative recombination of photocarriers due to the development of dislocations in them. We assume that capture of photogenerated carriers in QD-3 is suppressed in a magnetic field, so that the photoluminescence intensity increases. The effect of magnetic field on the photoluminescence intensity is explained as follows. Let $n_{1,2}=n_1+n_2$ denote the total concentration of QD-1 and QD-2 and $n_3$ denote the concentration of QD-3. We assume that the probability of capture in a quantum well is proportional to its area. The average sizes of QD-1 and QD-2 differ by only 15% and are on the order of 150 Å. In this case, we can assume that the probabilities of capture into QD-1 and QD-2 are roughly equal, with $\omega_2 \approx \omega_1 = \omega_{1,2}$. The sizes of the QD-3 are limited by the terrace size. For a disorientation angle of 4°, the terrace size is 300 Å(Ref. 3), and the probability of capture into QD-3 ($\omega_3$) is different. Thus, $\omega = \omega_3/\omega_{1,2} \approx 4$.

Under steady-state conditions, the photoluminescence is determined by the extent of capture (the fractions of photogenerated carriers captured into QD-1 and QD-2),

$$\chi_0 = \frac{n_{1,2}\omega_{1,2}}{n_{1,2}\omega_{1,2}+n_3\omega_3}. \tag{1}$$

In an external magnetic field $B$, the photocarriers are localized in a region whose size is on the order of the magnetic length $L$ ($L^2 = \hbar c/eB$). In a strong magnetic field, $1/\pi L^2 \gg n_1+n_2+n_3$ ($B \gg B_0 = \pi \hbar c(n_1+n_2+n_3)/e$) and the pho-

tocarriers are captured into the nearest QD. In this case the capture of photocarriers is determined only by the concentration of QDs and the extent of capture is given by

$$\chi_\infty = \frac{n_{1,2}}{n_{1,2}+n_3}. \tag{2}$$

Comparing Eqs. (1) and (2), we find

$$\frac{\chi_\infty}{\chi_0} = \frac{1+\omega n_3/n_{1,2}}{1+n_3/n_{1,2}}. \tag{3}$$

For $\omega > 1$, the QD photoluminescence intensity increases in a magnetic field and for $\omega < 1$, it decreases.

Figure 3 shows the QD photoluminescence intensity as a function of magnetic field, normalized to the intensity in zero field. Two characteristic regions can be distinguished. For $B < 2$ T, the photoluminescence intensity increases with rising magnetic field. For $B > 2$ T, the photoluminescence intensity saturates to 1.2 times the QD photoluminescence intensity for $B=0$. This dependence can be used to estimate the characteristic magnetic field $B_0 \approx 1$ T and to determine the total QD concentration: $n_1+n_2+n_3 = eB_0/\pi \hbar c \approx 10^{11}$ cm$^{-2}$. This concentration is consistent with the value determined previously by atomic-force microscopy in similar structures.

## 3. THEORY. DISCUSSION OF RESULTS

By analogy with Ref. 5, the fraction of photocarriers captured into a QD of type $i$ ($i=1,2,3$) can be written as

$$J_i \sim n_i \Gamma_i^{(0)}, \tag{4}$$

where $\Gamma_i^{(0)}$ is the rate of generation of particles in a QD of the corresponding type. Then the extent of capture is

$$\chi = \frac{n_{1,2}\Gamma_{1,2}^{(0)}}{n_{1,2}\Gamma_{1,2}^{(0)}+n_3\Gamma_3^{(0)}}, \tag{5}$$

and the relative change in the intensity in a magnetic field can be written as

$$\frac{I_{QD}}{I_{QD}(B=0)} = (1+\omega n_3/n_{1,2})\frac{n_{1,2}\Gamma_{1,2}^{(0)}}{n_{1,2}\Gamma_{1,2}^{(0)}+n_3\Gamma_3^{(0)}}. \tag{6}$$

In the case of low excitation intensities, when filling of the QDs can be ignored, the generation rate is given by[5,6]

$$\Gamma_i^{(0)} = \int_0^\infty dr\, 2\pi r\, \Omega_i(r)T[\Omega_i(r)]. \tag{7}$$

Here $\Omega_i(r)$ is the rate of capture of particles into a QD of type $i$ located at distance $r$ and $T[\alpha]$ is the average lifetime of a particle whose probability of capture into clearly available QDs is equal to $\alpha$,[6]

$$T[\alpha] = \int_0^\infty dt\, \exp\left\{-\alpha t - \sum_i \int dr\, 2\pi r n_i [1-e^{-\Omega_i(r)t}]\right\}. \tag{8}$$

In an external magnetic field, the photocarriers are localized in the planes of the system with the wave function

$$\varphi_0^2(r) = \frac{1}{\pi L^2}\exp\left(-\frac{r^2}{L^2}\right). \tag{9}$$

In this case, the particle capture rate is given by

$$\Omega_i(r) = \varphi_0^2(r)\omega_i. \tag{10}$$

The relative change in the photoluminescence in a magnetic field can be found by successive substitution of Eqs. (7)–(10) in Eq. (6). The final expression is cumbersome and unsuitable for the analysis. To simplify it, we propose approximating the wave function in a magnetic field by the Heaviside unit step function

$$\varphi_0^2(r) \approx \frac{1}{\pi L^2}h(L-r). \tag{11}$$

Then, after successive substitutions, we obtain the following expression for the generation rates:

$$n_{1,2}\Gamma_{1,2}^{(0)} = \frac{B_{01}+B_{02}}{B}e^{-B_0/B}$$

$$\times \int_0^1 dq\, \exp\left(\frac{B_{01}+B_{02}}{B}q + \frac{B_{03}}{B}q^\omega\right),$$

$$n_3\Gamma_3^{(0)} = \frac{B_{03}}{B}e^{-B_0/B}$$

$$\times \int_0^1 dq\, \exp\left(\frac{B_{01}+B_{02}}{B}q^{1/\omega} + \frac{B_{03}}{B}q\right), \tag{12}$$

where $B_{0i} = \pi\hbar cn_i/e$. Substituting Eq. (12) in Eq. (6), we obtain the relative change in the QD photoluminescence intensity as a function of the magnetic field.

The QD photoluminescence intensity, calculated using Eqs. (6) and (12), is shown in Fig. 3. The best agreement with experiments is reached for $B_{01}+B_{02}=2$ T and $B_{03}=0.18$ T, which corresponds to concentrations $n_1+n_2=10^{11}$ cm$^{-2}$ and $n_3=9\times10^9$ cm$^{-2}$. Since $n_1\approx n_2$, we find $n_1=5\times10^{10}$ cm$^{-2}$ and $n_2=5\times10^{10}$ cm$^{-2}$.

## 4. CONCLUSIONS

We have shown that three types of QDs can develop in structures with InAs QDs grown on disoriented GaAs substrates (Fig. 1). Quantum dots formed in the corners of the terraces (QD-1) and in the centers of the terraces (QD-2) predominate. Quantum dots located on a single terrace can also coalesce (QD-3). The relative concentrations of QD-1 and QD-2 have been determined by analyzing the shape of the photoluminescence lines. The photoluminescence intensity of the quantum dots is found to increase in an external magnetic field. This effect is explained by suppression of photocarrier trapping in QD-3 quantum dots, which are nonradiative recombination centers. The QD photoluminescence calculated as a function of magnetic field using a simple model agrees well with experiment. The QD concentration is determined by analyzing the behavior of the photoluminescence in a magnetic field.

*)E-mail: George.Astakhov@pop.ioffe.rssi.ru

[1] L. V. Asryan and R. A. Suris, Semicond. Sci. Technol. **11**, 554 (1996).
[2] D. Leonard, K. Pond, and P. M. Petroff, Phys. Rev. B **50**, 11 687 (1994).
[3] V. P. Evtikhiev, V. E. Tokranov, A. K. Kryzhanovskiĭ, A. M. Boĭko, R. A. Suris, A. N. Titkov, A. Nakamura, and M. Ichida, Fiz. Tekh. Poluprovodn. **32**(7), 860 (1998) [Semiconductors **32**, 765 (1998)].
[4] D. G. Vasil'ev, V. P. Evtikhiev, V. E. Tokranov, I. V. Kudryashov, and V. P. Kochereshko, Fiz. Tverd. Tela (St. Petersburg) **40**(5), 855 (1998) [Phys. Solid State **40**, 762 (1998)].
[5] G. V. Astakhov, A. A. Kiselev, V. P. Kocheresko, M. M. Moiseeva, and A. V. Platonov, Semicond. Sci. Technol. **13**(14), (1999).
[6] A. A. Kiselev, Fiz. Tekh. Poluprovodn. **32**(5), 564 (1998) [Semiconductors **32**, 504 (1998)].

Translated by D. H. McNeill

# Short-wavelength current tuning of InAsSb/InAsSbP heterostructure lasers caused by an injection nonuniformity

A. P. Danilova, T. N. Danilova, A. N. Imenkov, N. M. Kolchanova, M. V. Stepanov, V. V. Sherstnev, and Yu. P. Yakovlev*)

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

A reduction in the emission wavelength in the preferred mode of InAsSbP/InAsSb/InAsSbP heterostructure lasers by 50 Å is observed when the current is raised from 1.8 to 5 times the threshold with dc and pulsed power. A comparison of the spectral and spatial distributions of the output as functions of current shows that this short-wavelength tuning is caused by a change in the distribution of the nonequilibrium charge carrier concentration over the strip width as the current is varied. This effect is modeled mathematically, taking into account the increase in the injection density and the drop in the output intensity from the middle to the sides of the waveguide. The results of the model calculation are in good agreement with experiment. © *1999 American Institute of Physics.* [S1063-7826(99)01609-9]

## 1. INTRODUCTION

Current tuneable semiconductor lasers emitting at wavelengths near $\lambda = 3.3\,\mu$m are of interest for high-resolution diode laser spectroscopy because characteristic absorption lines of natural and industrial gases such as methane, propane, ethylene, benzene, etc. lie in this region. These lasers can be used in fast instrumentation for active environmental monitoring to determine the composition and concentration of pollutants, as well as in medical diagnostics and industry for precision control of technological processes.

In earlier papers[1–3] only thermal tuning of the emission wavelength due to heating of the active medium during a current pulse has been discussed.

This paper is a continuation of our studies of smooth current tuning of the output wavelength of double heterostructure lasers[4–7] with an InAsSb active region surrounded by layers of InAsSbP and represents a new concept in laser wavelength tuning by nonlinear optical effects.

## 2. EXPERIMENTAL TECHNIQUE

Laser diodes based on double $N$-InAs$_{0.48}$Sb$_{0.17}$P$_{0.35}$/$n$-InAs$_{0.95}$Sb$_{0.05}$/$P$-InAs$_{0.48}$Sb$_{0.17}$P$_{0.35}$ heterostructures (Fig. 1) were grown by liquid-phase epitaxy on an InAs substrate with a thickness $h = 100\,\mu$m and a hole concentration of $(5-8) \times 10^{18}\,\text{cm}^{-3}$. The thickness of the active region of the laser was $\sim 1\,\mu$m and that of the wide-gap emitters, $\sim 3\,\mu$m. The active region was intentionally not doped, and the electron concentration in it was $\sim 10^{16}\,\text{cm}^{-3}$. The $N$-InAsSbP layer was doped with Sn to an electron concentration of $n \approx 1 \times 10^{18}\,\text{cm}^{-3}$ and the $P$-InAsSbP layer was doped with Zn to a hole concentration of $p \simeq (2-5) \times 10^{18}\,\text{cm}^{-3}$.

A mesa-strip of width $b \simeq 16\,\mu$m was formed on the grown structures by photolithography. Fabry–Perot cavities with lengths of $L = 250 - 375\,\mu$m were obtained by cleaving. The width of the laser structure in the substrate region was $c \simeq 500\,\mu$m.

These studies were done at liquid-nitrogen temperature in a continuous operating mode with the laser powered by short current pulses of duration 0.04 ms with a reciprocal duty factor of 100, sawtooth pulses with a repetition rate ranging from $10^2$ to $10^4$ Hz, and square pulses with a reciprocal duty factor of 2 (meander) and a repetition rate of 36 Hz.

## 3. EXPERIMENTAL RESULTS

We studied the current ($I$) dependence of the coherent emission spectra and directional diagrams of lasers with different current supplies. The measurements were done over a wide range of currents between 1 and 4 times threshold. The threshold current for the best samples was $I_{\text{th}} = 30 - 35$ mA. The lasers that were tested usually had a single spectral mode for currents of $(1.1-3)I_{\text{th}}$, while a different mode predominated at higher currents (Fig. 2). The same mode composition was observed with the various power supply variants.

Figure 3 shows the output wavelength as a function of the pump current for the different power supply variants: dc, meander, sawtooth, and short current pulses. The $\lambda(I)$ curves behave similarly for the different supply variants. At low currents, $I < 1.81 I_{\text{th}}$, there is a slight shift of the wavelength to longer wavelengths, by less than 10 Å; with further increases in the current, there is a smooth shift to shorter wavelengths by a substantially larger amount. The maximum shift in the mode toward shorter wavelengths was $\sim 50$ Å. For currents above $3I_{\text{th}}$, the slope of the $\lambda(I)$ curve decreases and the spectrum ceases to be single-mode; for currents $I > 4I_{\text{th}}$, the emission wavelength ceases to change with current at all. This consistence of the spectral characteristics for the different supply variants indicates that the active region is not heated during the current pulse. The shorter wavelength modes (than the predominant mode), which show up when $I > 3I_{\text{th}}$, shift with the current in the same way as the predominant mode. We have already ob-
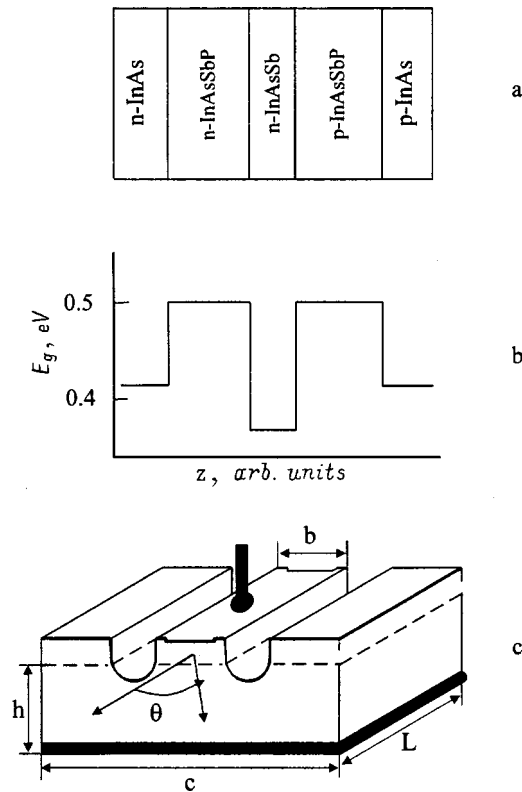
FIG. 1. The diode laser structure: a — location of the epitaxial layers, b — layer-by-layer profile of the gap width $E_g$ along the $z$ axis, c — design of the mesa-strip diode laser.
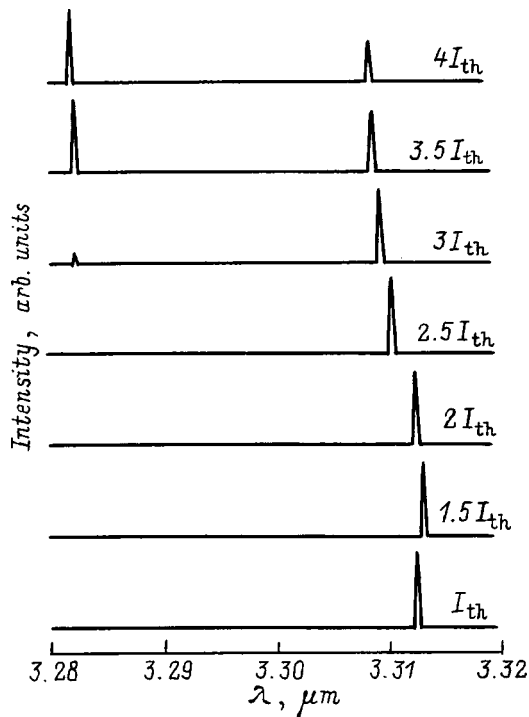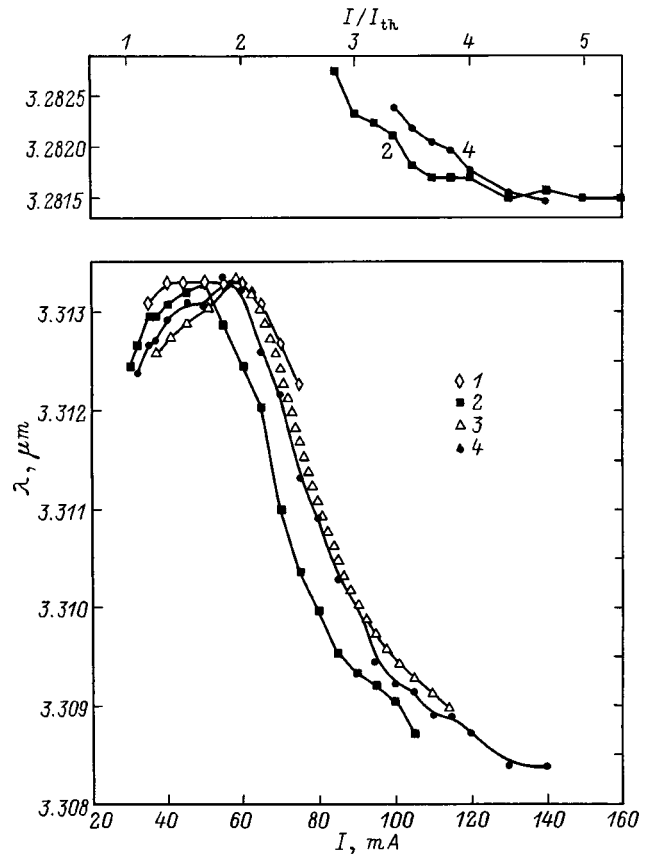


FIG. 3. Emission wavelength of the V12152-71 laser as a function of current when the sample is powered by dc (*1*), short pulses (*2*), sawtooth pulses (*3*), and meander (*4*).



FIG. 2. Emission spectra of the V12152-71 laser for different currents powered by short pulses of duration 0.04 ms.

served such a large short wavelength shift of the mode with rising current in similar lasers with a strip width of 10 $\mu$m.[6,7]

Figure 4 shows the spatial (with respect to the angle $\theta$) distribution of the laser output intensity. The smooth curves are experimental data and the dashed curves are theoretical calculations. At low currents (near $I_{th}$), the directional diagram corresponds to a single longitudinal mode with about 10% additional emission in transverse modes. As the current is raised to $2I_{th}$, the half-width of the directional diagram increases to 20° and the pedestal extends to $35-40°$ in the wings. This is explained by amplification of the transverse modes, whose contribution increases further when the current is raised to $I=3I_{th}$. A first-order transverse mode was also observed[7] in the output of the laser with a strip width of 10 $\mu$m, mentioned above, when $I>4I_{th}$.

## 4. DISCUSSION OF EXPERIMENTAL RESULTS

A comparison of the spectral and spatial distribution of the radiation as functions of current shows that the strongest retuning of the wavelength occurs at currents where output is generated in both longitudinal and transverse modes, with an intensity that falls off with mode number.

Let us consider the shift in the wavelength of the predominant mode together with the variation in the directional diagram with current. For currents $I<1.81I_{th}$ the radiation is concentrated mainly in a single longitudinal spatial mode
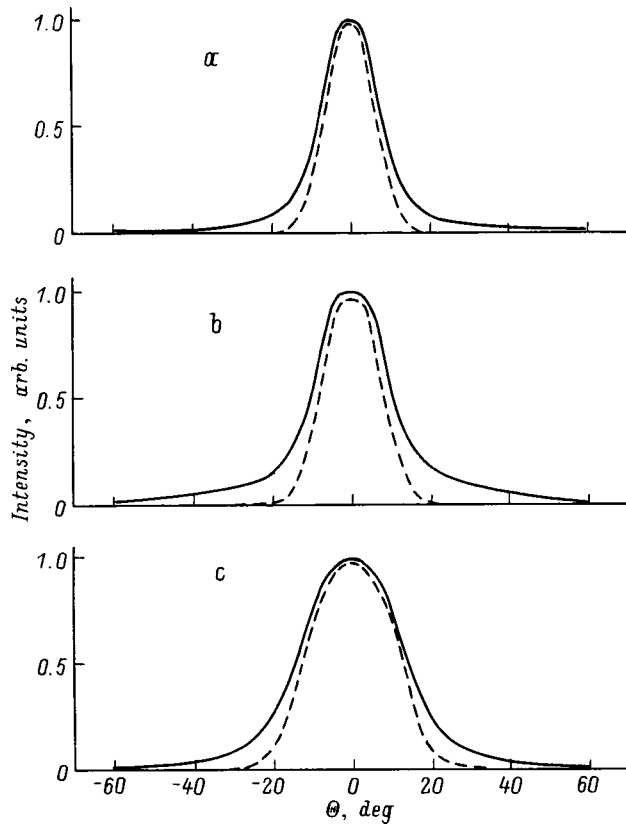
FIG. 4. Spatial distribution of the far-field emission from laser V12152-2. The smooth curves are experimental plots and the dashed curves are calculated for the case of a cosine distribution of the amplitude of the electromagnetic wave at the cavity mirror (a) and for wave packets with $\alpha y_0 = 1$ and $\alpha = 0.2 \times 10^4$ cm$^{-1}$ (b) and $\alpha = 0.3 \times 10^4$ cm$^{-1}$ (c). $I/I_{th}$: a — 1.1, b — 2, c — 3.

(Fig. 4a). Here nonequilibrium charge carriers are "burnt out" in the center of the active region, their concentration falls off somewhat, and, according to our calculations,[8] this leads to an increase in the propagation constant $\tilde{\varepsilon}$ of the electromagnetic wave and, therefore, to a shift in the output wavelength $\lambda$ to longer wavelengths.

As the current is increased further ($I > 2I_{th}$), the side lobes in the directional diagram of the output increase (Figs. 4b and 4c) and it can be assumed that nonequilibrium current carriers accumulate at the edges of the strip because of low surface recombination and the low intensity of the laser light; this creates conditions for the generation of transverse modes. Increasing the carrier concentration in the strip reduces the propagation constant $\tilde{\varepsilon}$ of the electromagnetic wave and, therefore, reduces the wavelength $\lambda$.

There is yet another reason for buildup of charge at the edges of the strip besides low surface recombination. It is hidden in the inhomogeneity of the pumping by the current in the region of the strip. Because of the large difference in area between the substrate (its width $c \simeq 500\,\mu$m) and the strip (its width $b \simeq 16\,\mu$m) (Fig. 1), the current density at the edges of the strip is very much higher than in its center, and this causes buildup of nonequilibrium charge carriers injected from the substrate, an associated reduction in the re-

fractive index, the appearance of lateral optical clipping, and the generation of transverse spatial modes.

We now examine a mathematical model of the effect of the nonlinearity of an optical medium on the propagation constant $\tilde{\varepsilon}$ of an electromagnetic wave that takes into account the pumping inhomogeneity over the strip width due to a voltage drop in the depth of the substrate.

We have shown[8] that the Helmholtz wave equation for the amplitude $u$ of an electromagnetic wave, averaged over the cavity length and in the absence of any spreading of charge carriers in the active region, has the form

$$\left(\frac{\lambda}{2\pi}\right)^2 \frac{d^2 u}{dy^2} + (\varepsilon_i - \tilde{\varepsilon})u + \delta\varepsilon_0 \frac{i+1}{1+\beta\tau N_\Phi} u = 0, \tag{1}$$

where $y$ is the coordinate along the cavity length, $\lambda$ is the vacuum wavelength, $\varepsilon_i$ is the dielectric permittivity at the inversion threshold, $\tilde{\varepsilon}$ is the propagation constant of the electromagnetic wave, $\delta\varepsilon_0$ is the change in the dielectric permittivity when the excitation level is increased from the inversion threshold to the lasing threshold, $i = (G\tau - N_{th})/(N_{th} - N_0)$ is the relative excess in the pumping density above threshold, $\beta$ is the differential gain coefficient, $\tau$ is the spontaneous lifetime of the charge carriers, $N_\Phi$ is the photon concentration, $G$ is the volume pump density, and $N_{th}$ and $N_0$ are the charge carrier concentrations at the lasing and inversion thresholds, respectively.

The distribution of the current over the width of the strip has been calculated[9] by conformal mapping for the case of an infinite laser width $c$. We used the transformation $w = \sin z$ and obtain a solution for a laser of finite width. The calculations showed that the current density increases nearly quadratically toward the edges of a mesastructure. It should be noted that this solution is true in the case where the substrate resistance exceeds the differential resistance of the $p-n$-junction. By conformal mapping we obtained the following expression for the substrate resistance:

$$R_s = \frac{\ln(4h/b)}{\sigma L(\pi/2 - 2h/c)}, \tag{2}$$

where $\sigma$ and $h$ are the conductivity and thickness of the substrate, respectively.

In our lasers $R_s$ is commensurate with the differential resistance of the $p-n$-junction of the laser in the stimulation regime. During the transition to the lasing regime, the differential resistance of the laser decreases sharply and becomes much less than the substrate resistance. Thus, we assume that the pumping at the edge of the strip rises compared to its middle only in the lasing regime. Then the relative excess of the pumping density above threshold, $i$, can be expressed as a function of the coordinate along the width of the strip by

$$i = i_0 \left[ 1 + k\left(\frac{2y}{b}\right)^2 \right], \tag{3}$$

where $i_0$ is the pump density at the center of the strip, and $k = (c/2h) - 1$ is a coefficient which depends on the ratio of the width of the substrate to its thickness and is $\sim 1.5$ in our case. This value of $k$ yields a ratio of the current densities at the edges and center of the strip of $\sim 2.5$.

Let us assume that many spatial modes are generated, so that the emission intensity is roughly the same over the width of the strip and is specified by the pumping at its center, $i_0$. The pumping at the edges of the strip is expended mainly in raising the concentration of nonequilibrium carriers and spontaneous emission. As a result, we have $\beta\tau N_\Phi = i_0$. Now the wave equation (1) takes the form of the linear harmonic oscillator equation

$$-\left(\frac{\lambda}{2\pi}\right)^2\frac{d^2u}{dy^2} + \frac{1}{2}2(-\delta\varepsilon_0)\frac{i_0}{1+i_0}k\left(\frac{2}{b}\right)^2 y^2 u = (\varepsilon_{th}-\widetilde{\varepsilon})u,$$
(4)

where $\varepsilon_{th}$ is the dielectric permittivity at the lasing threshold. Its solution can be written in the form of a sum of the eigenfunctions[10]

$$u_n = N_n H_n e^{-\alpha^2 y^2/2},$$
(5)

where $n$ is an integer $0,1,2,\ldots$; $n=0$ corresponds to the longitudinal mode, $n=1$ to the first transverse mode, etc.;

$$N_n = \left(\frac{\alpha}{\sqrt{\pi}2^n}\right)^{1/2},$$

$$H_n = (-1)^n e^{\xi^2}\frac{\partial^n}{\partial\xi^n}e^{-\xi^2}, \quad \xi = \alpha y,$$

$$\alpha^2 = \frac{4\pi}{\lambda b}\sqrt{(-\delta\varepsilon_0)k\frac{i_0}{1+i_0}}.$$

A preliminary estimate of $\alpha$ in our lasers for $\lambda = 3.3\ \mu m$, $b = 16\ \mu m$, $\delta\varepsilon_0 = -0.04$ and $i_0 = 1$ gives $0.2\times10^4\ cm^{-1}$. Only the longitudinal and first transverse modes can lie within the confines of the active region for these values of $\alpha$.

The propagation constant of the electromagnetic wave along the cavity, $\widetilde{\varepsilon}$, depends on the quantum number $n$, on $i_0$ and, therefore, on the current,

$$\widetilde{\varepsilon}_n = \varepsilon_{th} - \left(n+\frac{1}{2}\right)\frac{2\lambda}{\pi b}\sqrt{(-\delta\varepsilon_0)\frac{i_0 k}{1+i_0}}.$$
(6)

The propagation constant $\widetilde{\varepsilon}_n$ decreases sublinearly as the current rises and approaches saturation, in qualitative agreement with experiment.

At the same time, the general solution of the Schrödinger equation for a linear harmonic oscillator has the form of a minimizing wave packet oscillating around the coordinate origin. The wave packet[10] corresponds to the wave number

$$n_0 \approx \frac{1}{2}(\alpha y_0)^2,$$
(7)

where $y_0$ is the deviation of the center of gravity of the packet from the coordinate origin. For our values of $\alpha$, only the longitudinal and first transverse modes of the linear harmonic oscillator lie within the confines of the strip. Thus, $y_0$ corresponds to the position of the maximum of the first transverse mode, for which $\alpha y_0 = 1$. Equation (7) yields $n_0 = 1/2$. The directional diagrams for wave packets with $\alpha y_0 = 1$ are shown in Figs. 4b and 4c for $\alpha = 0.2\times10^4$ and $0.3\times10^4$

$cm^{-1}$, respectively. They give a lower emission intensity at the edges than observed experimentally. This difference may be explained by the simplifications in our calculations or by the independent generation of higher-order transverse modes.

It should be noted that this solution is valid for currents that ensure localization of the wave through an enhanced charge carrier concentration at the edges of the strip as the current rises. This corresponds to the condition $1/2(\alpha b/2)^2 > 1$. Then the minimum current at which the solution will be valid is given by

$$i_{0,min} = \left[k(-\delta\varepsilon_0)\left(\frac{b\pi}{2\lambda}\right)^2 - 1\right]^{-1}.$$
(8)

At lower currents, the solution is more complicated.

In our lasers, with $k = 1.5$, $\delta\varepsilon_0 = -0.04$, $\lambda = 3.3\ \mu m$, and $b = 16\ \mu m$ we obtain a minimum relative excess pumping above threshold, $i_{0,min} = 0.4$, in the middle of the strip. The condition $i_{0,min}\to\infty$ yields a minimum strip width for containment of a wave of

$$b_{min} = \frac{2\lambda}{\pi\sqrt{(-\delta\varepsilon_0)k}},$$
(9)

which is of order $b_{min} = 9\ \mu m$ for our structures.

The maximum reduction in the propagation constant of a wave packet as $i_0$ increases from 0 to $\infty$ is given by

$$\Delta\widetilde{\varepsilon}_{max} = \left(n_0+\frac{1}{2}\right)\frac{2\lambda}{\pi b}\sqrt{(-\delta\varepsilon_0)k}$$
(10)

and for $n_0 = 1/2$, $\Delta\widetilde{\varepsilon}_{max} = 0.032$. We find the corresponding reduction in the laser wavelength using the formula

$$\Delta\lambda = \frac{\Delta\widetilde{\varepsilon}\lambda}{2\varepsilon}.$$
(11)

The calculated value of the maximum reduction in the wavelength due to nonuniform inflow of current into the strips from the substrate is $\Delta\lambda_{max} = 41\ \text{Å}$ for our lasers, or only 22% less than the experimentally determined 50 Å. However, it should be pointed out again that not all orders of spatial modes show up in our samples, but primarily the longitudinal and first transverse mode. Because of this circumstance, the emission intensity near the edges of the strip is lower than in the center, which further increases the carrier concentration at the edges of the strip. In the first approximation, the coefficient $k$ thus increases by 1, and $\Delta\lambda_{max}$ increases by 30%. Thus, we can say that there is agreement between theory and experiment.

## CONCLUSIONS

In the lasers based on InAsSb/InAsSbP solid solutions that we have studied, the emission wavelength retunes when the current is changed. The overall shift in the spectral position of the mode to shorter wavelengths was $\sim50$ Å. Since the retuning was the same for different ways of powering the samples, we may assume that there was negligible heating of the laser during a current pulse. The directional diagram contained one longitudinal mode, which was supplemented by transverse modes as the current was increased. The strongest

retuning toward shorter output wavelengths occurred at those currents for which transverse modes were present.

A relationship between the propagation constant of an electromagnetic wave and the current has been established theoretically for the case of nonuniform injection of nonequilibrium current carriers along the width of a strip due to large differences over the area of the strip and substrate. This analysis also took into account the generation of many spatial modes. Good agreement has been obtained between the experimental data and the calculations.

*)E-mail: yak@iropt1.ioffe.rssi.ru; Fax: (812)247 0006

[1] V. G. Avetisov, A. N. Baranov, A. N. Imenkov, A. I. Nadezhdinskiĭ, A. N. Khusnutdinov, and Yu. P. Yakolev, Pis'ma Zh. Tekh. Fiz. **16**, 66 (1990) [Sov. Tech. Phys. Lett. **16**, 85 (1990)].

[2] Yu. P. Yakovlev, A. N. Baranov, A. N. Imenkov, V. V. Sherstnev, E. V. Stepanov, and A. Ya. Ponurovskiĭ, Kvant. Elektron. (Moscow) **20**, 839 (1993).

[3] A. A. Popov, Yu. P. Yakovlev, A. N. Baranov, A. N. Imenkov, and V. V. Sherstnev, Proc. SPIE **2112**, 50 (1994).

[4] T. N. Danilova, O. I. Evseenko, A. N. Imenkov, N. M. Kolchanova, M. V. Stepanov, V. V. Sherstnev, and Yu. P. Yakovlev, Pis'ma Zh. Tekh. Fiz. **22**, 7 (1996) [Tech. Phys. Lett. **22**, 25 (1996)].

[5] T. N. Danilova, O. I. Evseenko, A. N. Imenkov, N. M. Kolchanova, M. V. Stepanov, V. V. Sherstnev, and Yu. P. Yakovlev, Fiz. Tekh. Poluprovodn. **31**, 662 (1997) [Semiconductors **31**, 563 (1997)].

[6] T. N. Danilova, A. P. Danilova, O. G. Ershov, A. N. Imenkov, M. V. Stepanov, V. V. Sherstnev, and Yu. P. Yakovlev, Fiz. Tekh. Poluprovodn. **31**, 1392 (1997) [Semiconductors **31**, 1489 (1997)].

[7] T. N. Danilova, A. P. Danilova, O. G. Ershov, A. N. Imenkov, V. V. Sherstnev, and Yu. P. Yakovlev, Fiz. Tekh. Poluprovodn. **32**, 373 (1998) [Semiconductors **32**, 339 (1998)].

[8] A. P. Danilova, T. N. Danilova, A. N. Imenkov, N. M. Kolchanova, M. V. Stepanov, V. V. Sherstnev, and Yu. P. Yakovlev, Fiz. Tekh. Poluprovodn. **33**(2), 243 (1999) [Semiconductors **33**, 232 (1999)].

[9] A. P. Bogatov and M. P. Rakhval'skiĭ, Kvant. Elektron. (Moscow) **15**, 1720 (1988).

[10] L. Schiff, *Quantum Mechanics*, Inostr. Lit., Moscow (1957), p. 57.

Translated by D. H. McNeill

# Giant linear polarization of photoluminescence in type-II ZnSe/BeTe superlattices

A. V. Platonov,*) V. P. Kochereshko, and G. V. Mikhaĭlov

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

D. R. Yakovlev

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia;*
*Physikalisches Institut der Universität Würzburg, 97074 Würzburg, Germany*

W. Ossau, A. Waag, and G. Landwehr

*Physikalisches Institut der Universität Würzburg, 97074 Würzburg, Germany*

The photoluminescence spectra of type-II ZnSe/BeTe superlattices are investigated. Giant linear polarization of the luminescence is observed with unpolarized excitation in the spatially indirect exciton region. The effect is interpreted in a model of the general optical anisotropy of heterostructures with no common atom at the interfaces. © *1999 American Institute of Physics.* [S1063-7826(99)01709-3]

Several papers on the study of in-plane optical anisotropy of heterostructures with no common atom at the interfaces such as InAs/GaSb and (InGa)/InP have recently been published.[1–3] The optical anisotropy is attributed to a lower, compared with the interior volume, symmetry of an isolated interface between two semiconductors with zinc blende structure. In heterostructures with a common atom at the interfaces, for example, AlAs/GaAs, the interfaces in a quantum well (QW) transform into one another under a rotation by 90° and a mirror reflection in the well plane. This increases the general symmetry of the structure to $D_{2d}$. For structures with no common atom at the interface, the interface consists of two atomic layers — anionic and cationic. There are therefore four different combinations of direct and inverse interfaces (for example, for InAs/GaSb the combinations are GaAs–GaAs, InSb–InSb, GaAs–InSb, and InSb–GaAs). For identical interfaces the general symmetry of the structure remains $D_{2d}$, just as for QWs with a common atom, and no optical in-plane anisotropy arises. At the same time, for QWs with different interfaces a transformation which transforms interfaces into one another does not exist and the general symmetry of the system is lowered to $C_{2v}$, for which polarization anisotropy is now possible. Microscopically, the appearance of such anisotropy can be attributed to mixing of the light and heavy holes under normal incidence at the interface, as predicted in Ref. 4. Experimentally, such anisotropy was first observed by Voisin[1] for structures with quantum wells based on III–V semiconductors.

In this paper we present the results of an experimental study of this effect for type-II ZnSe/BeTe quantum wells. This system is chosen because the spatially indirect excitonic transition can be strongly tied to an interface and thereby be more sensitive to the interfacial properties.

The samples were grown by molecular-beam epitaxy on a GaAs (100) substrate. The types of interfaces were monitored during growth. Three types of ZnSe/BeTe (100/50) Å

× 20 superlattices with ZnTe–ZnTe, BeSe–BeSe, and BeSe–ZnTe interfaces were investigated.

The unpolarized photoluminescence (PL) spectra of such superlattices have been investigated in Refs. 5 and 6. Two excitonic resonances with energies of the order of 2.8 and 2.0 eV appear in the PL spectra. These resonances are due to a spatially direct exciton in ZnSe and an indirect exciton with a hole localized in the BeTe layer and an electron in the ZnSe layer.

We measured the degree of polarization of the PL spectra in the temperature range 6–70 K. Both linear and circular polarizations of the PL were investigated. A 441.6-nm He–Cd laser was used for excitation; this is only negligibly higher than the energy of a direct exciton in ZnSe. Linearly polarized, circularly polarized, and unpolarized excitations were used.

It was observed that irrespective of the excitation polarization, the photoluminescence signal of the indirect exciton is linearly polarized along the [1̄10] axis. The degree of polarization depends strongly on the types of interfaces. For a structure with nominally equivalent ZnTe–ZnTe interfaces it does not exceed 50% and for BeSe–BeSe it does not exceed 15%, while the polarization reaches 70% and higher for the nonequivalent pair BeSe–ZnTe (Fig. 1).

In the theoretical studies[3,4] the effective-mass method with boundary conditions of a special form was used to describe interfacial mixing of light and heavy holes. It was shown that for typical III–V type-I heterostructures the degree of polarization does not exceed 20%. In the experimental studies, the same effect reaches 40% for the same structures. Therefore the effect is much stronger in structures of our type than in previously investigated structures.

Such a large enhancement of the effect could be attributed to a strong localization of a spatially indirect exciton at the interface. The possibility of an exciton being localized near an interface in these structures was already pointed out
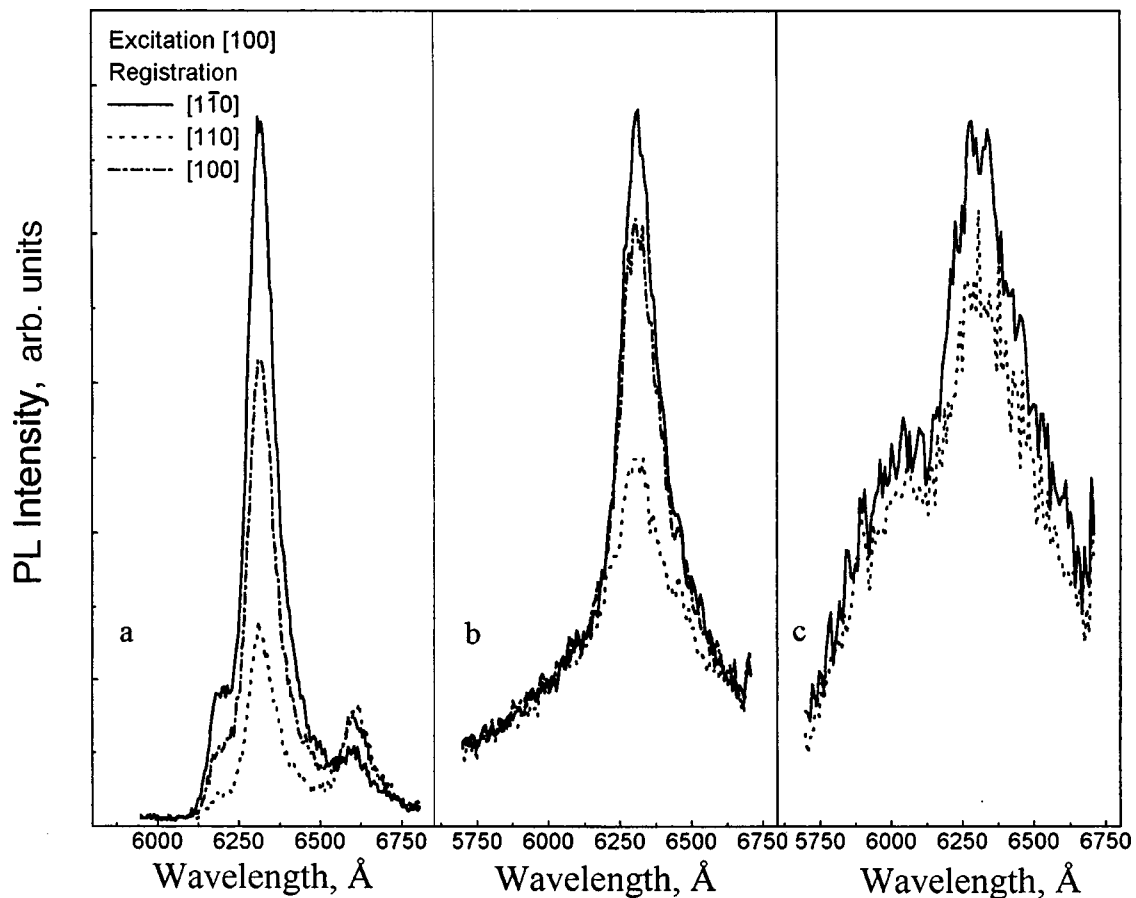
FIG. 1. Polarized photoluminescence spectra at 6 K temperature from a spatially indirect exciton in the superlattice ZnSe/BeTe (100 Å /50 Å) × 20 periods with different types of interfaces: a — (BeSe)–(ZnTe); b — (ZnTe)–(ZnTe); c — (BeSe)–(BeSe). The excitation light is linearly polarized in the [100] direction.

in Ref. 7 on the basis of an analysis of the photoluminescence spectra with a high level of excitation. In Ref. 8 it was indicated on the basis of secondary-electron emission that band bending near a heterojunction is a possible localization mechanism. To check this conjecture we investigated short-period superlattices (SLs) ZnSe/BeTe 40/20 Å . Localization should be much weaker in such a superlattice because the quantum-well energy of the hole is higher and, as a result, the hole wave function should be smeared over the entire well. We did not observe measurable linear polarization of photoluminescence for such short-period SLs for any combination of interfaces.

It follows from symmetry considerations that there should be no linear polarization in an ideal structure with equivalent interfaces. However, we observed quite strong linear polarization in our test samples. The reason could be that in real structures the direct and inverse interfaces are different. Such nonequivalence is due to the large difference in the chemical activity of Be with respect to Zn and Se with respect to Te. The high chemical activity of Be degrades the quality of both interfaces grown on a BeTe layer as compared with growth on a ZnSe layer. Low interfacial quality leads to an increase in nonradiative recombination. Consequently, the direct and inverse interfaces make different contributions to the total signal. In turn, this makes possible
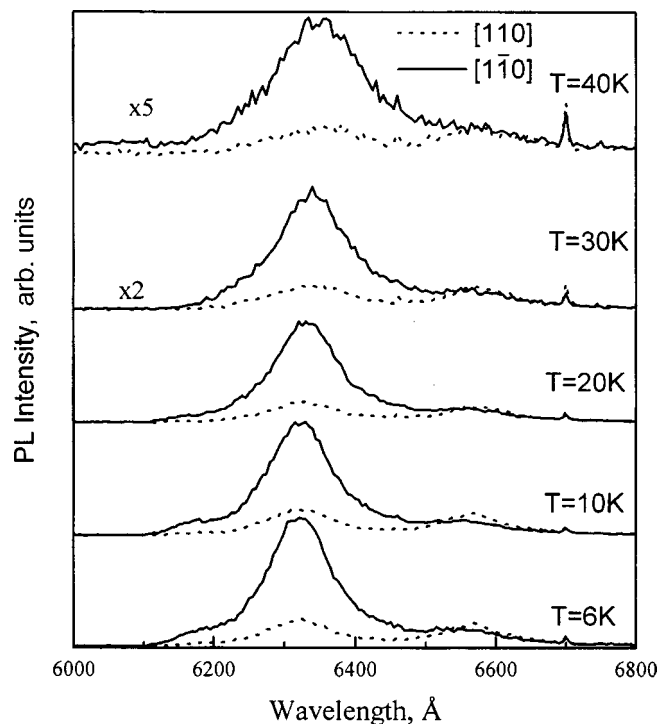


FIG. 2. Degree of polarization of PL versus the temperature, along the principal axes of a structure for the superlattice ZnSe/BeTe (100 Å /50 Å) × 20 with asymmetric interfaces (BeSe)–(ZnTe).

linear polarization of PL for other structures with nominally equivalent interfaces. A built-in electric field, present because of band bending, as previously indicated, could be another reason for symmetry lowering.[3]

Figure 2 shows the temperature dependence of the PL spectra. It is evident from the figure that in the temperature range considered here the degree of linear polarization remains the same and only the signal amplitude decreases. This behavior indicates that the polarization of PL is caused not by thermalization between two weakly split levels, but rather by the difference in the transition probabilities between the [110] and [1$\bar{1}$0] orientations. This makes it possible to compare our results directly with the results obtained by other authors using absorption spectroscopy. Moreover, this makes it possible to rule out the influence of states localized on interfacial defects (for example, well-width fluctuations), which can be strongly polarized because of the geometry of the defects. If the contribution of these states to the polarization signal were the determining contribution, then we would have observed a decrease in the general polarization of the signal with increasing temperature and carrier delocalization.

In summary, we have reported the results of an experimental study of PL spectra for type-II superlattices ZnSe/BeTe. We found that the PL signal from a spatially indirect transition is strongly polarized in the [1$\bar{1}$0] direction. We showed that the effect is attributable to the general in-plane anisotropy of heterostructures with no common atom at the interface. The large magnitude of the effect can be explained by strong hole localization at interfaces in type-II structures.

*)E-Mail: Alexei.Platonov@pop.ioffe.rssi.ru

[1] O. Krebs and P. Voisin, Phys. Rev. Lett. **77**, 1829 (1996).

[2] O. Krebs, W. Seidel, and P. Voisin, Inst. Phys. Conf. Ser. **155**, Chapter 12 (*Proceedings of the 23rd International Symposium on Compound Semiconductors,* St. Petersburg 1996, edited by M. S. Shur and R. A. Suris), p. 859.

[3] E. L. Ivchenko, A. A. Toropov, and P. Vuazen, Fiz. Tverd. Tela (St. Petersburg) **40**, 1925 (1998) [Phys. Solid State **40**, 1748 (1998)].

[4] E. L. Ivchenko, A. Yu. Kaminskii, and U. Rossler, Phys. Rev. B **54**, 5852 (1996).

[5] A. V. Platonov, D. R. Yakovlev, U. Zehnder, V. P. Kochereshko, W. Ossau, F. Fischer, T. Litz, A. Waag, and G. Landwehr, Acta Phys. Pol. A **92**, 1063 (1997).

[6] A. V. Platonov, D. R. Yakovlev, U. Zehnder, V. P. Kochereshko, W. Ossau, F. Fischer, T. Litz, A. Waag, and G. Landwehr, J. Cryst. Growth **184/185**, 801 (1998).

[7] S. V. Zaĭtsev, V. D. Kudakovskiĭ, A. A. Maksimov, D. A. Pronin, I. I. Tartakovskiĭ, N. A. Gippius, D. R. Yakovlev, V. Ossau, and K. Landwehr, JETP Lett. **66**, 376 (1997).

[8] M. Nagelstrasser, H. Droge, F. Fischer, T. Litz, A. Waag, G. Landwehr, and H.-P. Steinruck, J. Appl. Phys. **83**, 4253 (1998).

Translated by M. E. Alferieff

# Wide-gap semiconductors for high-power electronics

A. A. Lebedev*) and V. E. Chelnokov

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

The latest results obtained in the development of high-power devices based on wide-gap semiconductors are examined. It is shown that at present silicon carbide remains the most promising material for high-temperature, radiation-resistant, high-power electronics. Certain factors involving a wide commercial adoption of SiC-based devices are examined.
© *1999 American Institute of Physics.* [S1063-7826(99)01809-8]

In modern semiconductor physics there are now two very rapidly developing directions of work 1) changing the properties of a material by modifying the geometric dimensions of the structures, i.e., the physics of nanostructures, and 2) the development and study of new semiconductor materials.

Important work in the second direction involves the study of wide-gap semiconductors. The potential possibilities of wide-gap materials for creating semiconductor devices were analyzed a long time ago.[1–3] Band gaps larger than in Si and GaAs give these materials the following advantages:

- a larger operating temperature range;
- ability to construct visible-range, light-emitting devices based on these materials;
- high critical breakdown fields ($E_{cr}$);
- high radiation resistance.

As we know, special quality criteria — figures of merit (fm) calculated using the basic physicochemical properties of a semiconductor — have been proposed for comprehensive accounting of the possibilities of a semiconductor material.

Johnson[1] has proposed to use the product of $E_{cr}$ and the saturation rate $V_{sat}$ as a figure of merit JM, which determines the operating limit of a conventional transistor: $JM = (E_{cr}V_{sat}/\pi)^2$. Later, the following criterion was defined: $KM = \lambda(V_{sat}/\varepsilon)^{1/2}$ ($\varepsilon$ — permittivity of the semiconductor, $\lambda$ — thermal conductivity), where the switching rate of a transistor operating as a logic element of a computer was taken into account.[4] Baliga[5] has proposed yet another figure of merit BM for evaluating a semiconductor material. It is related to the operating losses of high-power field-effect transistors: $BM = \mu\varepsilon E_{cr}^3$ ($\mu$ — mobility of current carriers). However, this criterion was associated primarily with ohmic losses and was used to assess the possibilities of a semiconductor from the standpoint of low-frequency devices. For assessment of high-frequency devices, losses associated with switchings must also be examined. The criterion $BH = \mu E_{cr}^2$, based on the assumption that losses due to switchings are caused by the recharging input capacitance of a device, was proposed in Ref. 6. Other important parameters for selecting materials for high-powered devices are the thermal properties of a semiconductor. A figure of merit which takes into account these properties has been proposed in Ref. 7: $QF_1 = \lambda\mu\varepsilon E_{cr}^3$.

Published data on the average values of the parameters of silicon carbide, which are used to calculate the figures of merit listed above, are presented in Table I.

The computed relative values of five figure-of-merit (fm) parameters for 6*H*- and 4*H*-SiC, Si, GaAs, GaP, GaN, AlN, and diamond are presented in Table II. In the calculation the values of the parameters for Si were taken as the unit of measurement. Table II also indicates whether a given semiconductor is a direct- or indirect-gap semiconductor and whether a large-diameter substrate consisting of the same material is present ($+$) or absent ($-$). The first factor largely determines the application of a given material for optoelectronic devices. The second factor determines the application possibilities of the scientific results in commercial production. As one can see from the tables, silicon carbide is better than Si, GaAs, and GaP with respect to all criteria considered by us. At the same time, GaN and AlN are better than SiC in terms of several parameters. After *p–n* structures and then LEDs and GaN-based lasers were developed,[8] GaN was found to be unequaled for optoelectronic devices based on wide-gap materials. GaN is also a strong competitor for silicon carbide in creating field-effect transistors.[9] However, GaN and AlN so far do not have substrates made of the same material; they are grown by heteroepitaxy using substrates consisting of other materials, including silicon carbide.[6,10] This leads to a very high dislocation density in the films ($>10^7$ cm$^{-2}$), which for now limits the application of these

TABLE I. The basic parameters of some semiconductors.

| | Parameter | | | | | |
|---|---|---|---|---|---|---|
| Material | $E_g$, eV | $E_{cr}$ 10$^6$ V/cm | $V_{sat}$, 10$^7$ cm/s | $\lambda$, W/(cm·K) | $\mu_p$, cm$^2$/(V·s) | $\mu_n$, cm$^2$/(V·s) |
| Si | 1.1 | 0.3 | 1 | 1.5 | 600 | 1500 |
| GaAs | 1.43 | 0.6 | 1 | 0.45 | 400 | 8500 |
| GaP | 2.2 | 0.5 | 1.5 | 0.7 | 150 | 250 |
| 4*H*-SiC | 3.2 | 3 | 2 | 4.9 | 50 | 1000 |
| GaN | 3.45 | >1 | 2.2 | 1.3 | 850 | 1250 |
| Diamond | 5.45 | 10 | 2.7 | 22 | 1600 | 2200 |
| AlN | 6.2 | ? | ? | 2 | 14 | ? |

*Note:* $\mu_p$ and $\mu_n$ — hole and electron mobilities, respectively; $E_g$ — band gap.

TABLE II. The values of normalized figures of merit and a number of other parameters for certain semiconductors.[1,2].

| Material | $JM = (E_{cr}V_{sat}/\pi)^2$ | $KM = \lambda(V_{sat}/\varepsilon)^{1/2}$ | $BM = \varepsilon\mu E_{cr}^3$ | $BH = \mu E_{cr}^2$ | $QF_1 = \lambda\varepsilon\mu E_{cr}^3$ | $T_{oper}$, K | $T_D$, K | Band-gap structure | Existence of a like substrate with diameter $\geq 1$ inch |
|---|---|---|---|---|---|---|---|---|---|
| Si | 1 | 1 | 1 | 1 | 1 | 410 | 645 | Indirect | + |
| GaAs | 11 | 0.45 | 28 | 16 | 9.4 | 570 | 344 | Indirect | + |
| GaP | 37 | 0.73 | 16 | 3.8 | 9.4 | 800 | 445 | Indirect | + |
| 6$H$-SiC | 260 | 5.1 | 90 | 13 | 300 | 1200 | 1200 | Indirect | + |
| 4$H$-SiC | 410 | 5.1 | 290 | 34 | 950 | 1230 | 1200 | Indirect | + |
| GaN | 790 | 1.8 | 910 | 100 | 910 | 1250 | 600 | Direct | − |
| Diamond | 5330 | 31 | 14860 | 1080 | 198100 | 2100 | 1860 | Indirect | − |
| AlN | 5120 | 2.6 | 390 | 14 | 660 | 2100 | 747 | Indirect | − |

*Note:* $T_{oper}$ — operating temperature, $T_D$ — Debye temperature.

materials for high-power semiconductor devices. Moreover, as shown in Refs. 1 and 11, for wide-gap semiconductors the Debye temperature $T_D$, which determines the limit of thermal stability of a semiconductor, should be considered on par with the maximum operating temperature $T_{oper}$, which is determined by the band gap of the material. If $T_D < T_{oper}$, the maximum working temperature decreases. As one can see from Fig. 1 and Table II, with respect to $T_D$ and $T_{oper}$ the semiconductors GaN and AlN are not as good as SiC.

Diamond is superior with respect to the figures of merit and the maximum operating temperatures. However, for now it is impossible to obtain single-crystal diamond layers by heteroepitaxy, and diamond substrates have a small area and are rather expensive.[12] Moreover, there are definite difficulties in obtaining high-quality $p$–$n$ structures for diamond.

In the last decade the full potential application of silicon carbide has been realized. The maximum values of the parameters obtained for certain types of semiconductor devices are presented in Table III. The first industrially produced SiC-based devices were blue LEDs. However, after high-quality GaN-based $p$–$n$ structures were produced GaN took the leading position in the development of optoelectronic devices.

Single-crystal SiC substrates are now produced commercially. They are used for growing SiC itself and for het-eroepitaxy of GaN. In contrast with growth on sapphire, the smaller lattice mismatch and closer match of thermal expansion coefficients make it possible to obtain structurally more perfect epitaxial GaN layers based on SiC.[18] Moreover, a conducting SiC substrate makes it possible to use a vertical device geometry, which greatly simplifies the technology.

One of the main problems in producing SiC-based semiconductor devices is the low quality of single-crystal substrates. The dislocation density in commercially produced (CREE Inc., USA) substrates is $10^3 - 10^5$ cm$^{-2}$, and they contain extended micropipes. The existence of such defects limits the operating area of the devices produced. For high-frequency transistors this deficiency is not so limiting and devices of this type based on SiC can operate in the gigahertz range, and, with respect to their parameters, they are essentially at least as good as Si- and GaAs-based transistors.[16] At the same time, even though it has no material advantages from the standpoint of the breakdown voltage, carrier mobility, and drift saturation rate, GaN makes it possible to obtain heterostructures based on the solid solutions AlGaN–GaN, with a two-dimensional electron gas. However, because of the large number of polytypes, SiC also has great potential possibilities from the standpoint of producing heteroepitaxial structures. Thus it is difficult to say right now which of the wide-gap material will be the main material in the future for producing microwave-range devices. According to assessments by specialists, by the year 2000 SiC-based field-effect transistors will be used in commercially produced electronics.[19]

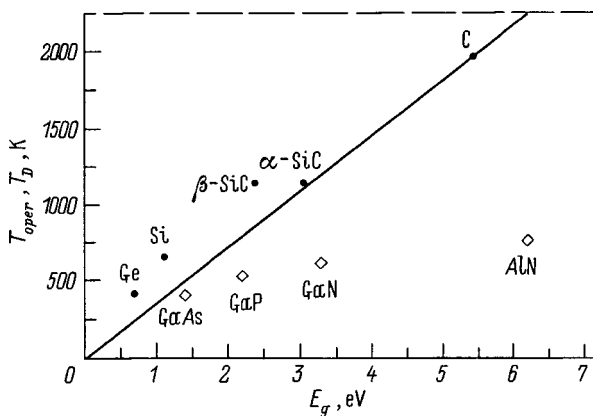The high defect density in substrates and epitaxial layers



FIG. 1. Debye temperature $T_D$ and maximum operating temperature $T_{oper}$ versus the band gap for certain semiconductors.

TABLE III. The parameters of certain devices based on SiC.

| Device type | Parameters and references |
|---|---|
| Schottky diode | $U_{br} = 3$ (Ref. 3) |
| Diode | $U_{br} = 5.5$ (Ref. 14) |
| Thyristor | $U_s = 700$, $I_{max} = 1.5$ (Ref. 15) |
| MESFET | $f_{max} = 40$, $W_{max} = 3.3$ (Ref. 16) |
| MOSFET | $U_{br} = 2.6$ (Ref. 17) |

*Note:* $U_{br}$ — maximum operating voltage, $U_s$ — switching voltage, $I_{max}$ — maximum operating current, $f_{max}$ — maximum operating frequency, $W_{max}$ — maximum power level.

of SiC is still one of the main limiting factors for the development of high-power devices. The maximum breakdown voltages, on the order of or greater than 1 kV, generally have been attained for $p$–$n$ structures with a diameter on the order of 100 $\mu$m. The small working areas made it impossible to obtain on the basis of such devices substantial dc currents and, correspondingly, it is difficult to call such devices high power-level devices.

The advancement of the technology for growing substrates gives hope that in the future it will be possible to decrease the defect density in substrates substantially and to obtain device structures with reasonable dimensions. Besides improving the technology for volume growth of SiC, various alternative variants for increasing the structural perfection of substrates are being developed. One such variant is a method of treating substrates by liquid-phase epitaxy.[20] As a result, a buffer layer grows and the defects present in the substrate are buried. Thus, after treatment in melt an epitaxial layer was obtained on the basis of a commercially produced substrate by sublimation epitaxy, and Schottky diodes based on it possessed breakdown voltages up to 100 V and an area of 8 mm$^2$ (Ref. 21).

In addition, it should be noted that epitaxial layers of SiC obtained by sublimation epitaxy are themselves more heat- and radiation-resistant than layers obtained by CVD.[22,23] Recently[24,25] interest has reappeared in this growth method, which was proposed at A. F. Ioffe Physicotechnical Institute of the Russian Academy of Sciences[26] and traditionally used here. We note that in our laboratory the method of sublimation epitaxy has been used to obtain diode structures with breakdown voltages on the order of 500 V and a working area of 2 mm$^2$.

*)E-Mail: shura.lebe@pop.ioffe.rssi.ru

[1] E. O. Jonson, RCA Rev. **26**, 163 (1965).

[2] A. E. Otblesk and V. E. Chelnokov, in *Proceedings of the 11th Winter School of the Physiotechnical Institute* (Leningrad Institute of Nuclear Physics, Leningrad, 1979), p. 161.

[3] M. N. Yoder, IEEE Trans. Electron Devices **ED-43**, 1633 (1996).

[4] R. W. Keyes, Proc. IEEE **60**, 225 (1972).

[5] B. J. Baliga, J. Appl. Phys. **53**, 1759 (1982).

[6] R. Gaska, Q. Chen, J. Yang, A. Osinsky, M. Asif Khan, and M. S. Shur, IEEE Electron Device Lett. **18**, 492 (1997).

[7] K. Szenai, R. S. Scott, and B. J. Baliga, IEEE Electron. Dev. **10**, 85 (1989).

[8] S. Nakamura, MRS Bull. **22**, 29 (1997).

[9] M. S. Shur and M. A. Khan, MRS Bull. **22**, 44 (1997).

[10] J. T. Torvik, Qin Chang-Hua, M. Lecsono, and J. I. Pankove, Appl. Phys. Lett. **72**, 945 (1998).

[11] H. Henisch and R. Roy, *Silicon Carbide* (Mir, Moscow, 1972), p. 17.

[12] S. Sussman, J. R. Brandor, G. A. Scarsbrook, C. G. Sweeney, T. J. Valentine, A. J. Whitehaed, and C. J. Wort, Diamond Relat. Mater. **3**, 303 (1994).

[13] Q. Wahab, T. Kimoto, A. Ellison, C. Hallin, M. Timonen, R. Yakimova, A. Henry, J. P. Bergman, and E. Janzen, Appl. Phys. Lett. **72**, 445 (1998).

[14] C. H. Carter, V. F. Tsvetkov, D. Henshall, O. Kordina, K. Irvine, R. Singh, S. T. Allen, and J. W. Palmor, in *Abstracts of Reports at the 2nd European Conference on SiC and Related Materials* (Sept. 2–4, 1998, Montpellier, France), p. 1.

[15] A. K. Agarwal, J. B. Casady, L. B. Rowland, S. Seshadri, R. R. Siergeiej, W. F. Valek, and C. D. Brandt, IEEE Electron Device Lett. **18**, 518 (1997).

[16] R. R. Siergeiej, S. Sriman, R. C. Clarke, A. H. Agarwal, C. D. Brandt, A. A. Burk, T. J. Smith, A. W. Morse, and P. A. Orphanos, Inst. Phys. Conf., No. **142**, 769 (1996).

[17] J. A. Cooper, Jr., M. A. Melloch, J. M. Woodall, J. Spitz, K. J. Schoen, and J. P. Henning, Meter. Sci. Forum **264–268**, 895 (1998).

[18] R. Gaska, J. W. Yang, A. Osinsky, Asif Khan, A. O. Orlov, G. L. Shidet, and M. S. Shur, Appl. Phys. Lett. **72**, 707 (1998).

[19] Ch. Brylinskii, in *Abstracts of Reports at the 2nd European Conference on SiC and Related Materials* (Sept. 2–4, 1998, Montpellier, France), p. 45.

[20] S. V. Renadakova, I. P. Nikitina, A. S. Tregubova, and V. A. Dmitriev, J. Electron. Mater. **27**, 292 (1998).

[21] V. Dmitriev, S. Rendakova, N. Kuznetsov, N. Savkina, A. Andreev, M. Rastegaeva, M. Munbaeva, and A. Morozov, in *Abstracts 2nd European Conference on SiC and Related Materials* (Sept. 2–4, 1998, Montpellier, France), p. 251.

[22] A. A. Lebedev, M. G. Rastegaeva, A. L. Syrkin, N. S. Savkina, A. S. Tregubova, V. E. Chelnokov, and M. P. Scheglov, Inst. Phys. Conf. Ser., No. 155, 605 (1997).

[23] A. A. Lebedev, A. M. Strel'chuk, V. V. Kozlovski, N. S. Savkina, D. V. Davydov, and V. V. Solov'ev, in *Abstracts of Reports at the 2nd European Conference on SiC and Related Materials* (Sept. 2–4, 1998, Montpellier, France), p. 257.

[24] S. Nishino, T. Yoshida, K. Matsumoto, Y. Chen, and S. K. Lilov, in *Abstracts of Reports at the 2nd European Conference on SiC and Related Materials* (Sept. 2–4, 1998, Montpellier, France), p. 27.

[25] K. P. Raback, R. Yakimova, M. Syvajarui, R. Neiminer, and E. Janzen, in *Abstracts of Reports at the 2nd European Conference on SiC and Related Materials* (Sept. 2–4, 1998, Montpellier, France), p. 103.

[26] Yu. A. Vodakov, E. N. Mokhov, M. G. Ramm, and A. O. Roenkov, Krist. Tech. **14**, 729 (1979).

Translated by M. E. Alferieff