PHYSICAL REVIEW B VOLUME 50, NUMBER 8 15 AUGUST 1994-II

Defect formation and diffusion in heavily doped semiconductors

W. Walukiewicz
Center for Advanced Materials, Materials Science Division, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, California 94720

(Received 18 January 1994; revised manuscript received 2 May 1994)

The effect of high free-carrier concentrations on the formation of native defects is considered. It is shown that the onset of degenerate carrier statistics results in an abrupt reduction of the formation energy of charged native defects and an enhanced diffusion of impurities and host lattice atoms. The effects are especially important in very heavily doped semiconductors at reduced growth and processing temperatures. As an example, it is shown that the proper incorporation of the effects of the degenerate statistics is necessary for understanding the mechanisms determining the maximum attainable concentration of free carriers in p- and n-type $In_xGa_{1-x}As$.

Diffusion of impurities and self-diffusion are among the most important processes affecting the stability of semiconductor structures. Despite many years of extensive studies there is no generally accepted view of the microscopic diffusion processes even in the elemental semiconductors.[1,2] The problem is much more complex in compound semiconductors, where the diffusion processes can proceed via a larger variety of crystal lattice imperfections possible in such materials.[3-5]

In this paper we discuss the effects of heavy doping on defect formation and diffusion in semiconductors. It is shown that the onset of degenerate Fermi statistics for the electron or hole gas leads to a dramatic reduction of the formation energy of charged native defects, which in turn results in an increased compensation of shallow dopants and in a strong enhancement of the dopant and the host crystal lattice atom diffusion. This effect determines the limits of structural stability of heavily doped semiconductors and explains many diffusion-related phenomena observed in different semiconductor systems.[6]

In semiconductors the diffusion of impurities and/or native defects is thermally activated. The total activation enthalpy for a diffusion process can be divided into two parts, the enthalpy of the defect formation and the enthalpy of defect migration. The formation enthalpy determines the concentration of the diffusing species under equilibrium conditions, whereas the migration enthalpy represents the barrier for the migrating species between two equivalent lattice sites. It is commonly assumed that the entropy terms are not dependent on the Fermi energy. Since in this work we are interested in the Fermi-level dependence of the defect concentration and defect diffusion, we will treat the entropies as constants affecting only the preexponential factors.

It is quite obvious that the formation enthalpy and thus also the concentration of charged native defects depends on the location of the Fermi energy. The change in the formation enthalpy is associated with a transfer of electrons between the reservoir of free carriers and the localized native defects. In the following we will consider the formation of donorlike defects in a p-type semiconductor. However, these considerations can easily be adopted to the case of acceptorlike defects formed in an n-type semiconductor. The formation enthalpy of a donorlike defect that can donate s electrons is given by

$$H_f = H_{f0} + \sum_{n=0}^{s-1} (E_F - E^{n+}) , \quad (1)$$

where H_{f0} is the formation enthalpy of the neutral defect, E_F is the Fermi energy, and E^{n+} is the energy of the $n+1/n$ charge transition state.

Using Eq. (1) one obtains the expression for the thermal equilibrium concentration of defects:

$$C = C_0 \exp\{-[H_{f0} + s(E_F - E_D)]/kT\} , \quad (2)$$

where $E_D = \sum_{n=0}^{s-1} E^{n+}/s$ is the average energy of all the charge transition states. If the native defects participate in any self- or impurity diffusion the diffusivity for such process will be proportional to the native defect concentration given by Eq. (2), with an additional factor associated with thermal activation of the native defect migration process. If there is no charge exchange between a migrating native defect and the free-carrier gas reservoir, the migration enthalpy, H_m does not depend on the Fermi level and therefore the migration factor in the diffusion coefficient can be treated as a constant. Also, the temperature and the Fermi-level-independent terms associated with the formation and migration entropies are incorporated in the preexponential factor.

The Fermi energy can be related to free hole concentration p through the standard expression

$$p = N_v F_{1/2}(-E_F/kT) , \quad (3)$$

where $F_{1/2}(\ )$ is the Fermi Dirac integral, N_v the valence-band density of states, and E_F the Fermi energy relative to the valence-band edge:

$$N_v = 4\pi(2m_v kT/h^2)^{3/2} . \quad (4)$$

A simplifying assumption for nondegenerate free hole statistics is customarily assumed in standard treatments of the defect formation and diffusion problems. In this case Eq. (3) takes the form

0163-1829/94/50(8)/5221(5)/$06.00 ©1994 The American Physical Society

$$p=\sqrt{\pi/2}N_v\exp(-E_F/kT)\ . \tag{5}$$

Substituting the Fermi energy from Eq. (5) into Eq. (2), one obtains a well-known expression for the concentration-dependent defect concentration (see, e.g., Ref. 4):

$$C=C_0\exp(-H'_{f0}/kT)(p/p_i)^s\ , \tag{6}$$

where $H'_{f0}=H_{f0}+s(E_{Fi}-E_D^s)$; E_{Fi} and p_i are, respectively, the Fermi energy and hole concentration in the intrinsic semiconductor.

Expressions similar to that given by Eq. (6) or equivalent equations for n-type doped semiconductors are routinely used to analyze the concentration-dependent diffusion in semiconductors.[4,5] It should be noted, however, that they are applicable only to the cases where the electron or hole gases can be described in terms of nondegenerate statistics. It means that they can be used only for lightly and/or moderately doped semiconductors at high temperatures. Recent progress in achieving very high doping levels in materials grown and processed at low temperatures raises the issue of the limitation of the standard approach, and underscores the need for a more general treatment of this problem.

In general the concentration of defects given by Eq. (2) can be calculated by substituting the Fermi energy from Eq. (3). An analytic solution of this problem can be obtained only for the cases of nondegenerate and highly degenerate hole gases. In the latter case, i.e., when $-E_F \gg kT$, the hole concentration can be approximated by

$$p=(2N_v/3)(-E_F/kT)^{3/2}\ . \tag{7}$$

Substituting (4) and (7) into Eq. (2), one obtains the expression

$$C=C_0\exp\{-[H_{f0}-s(E_D+(h^2/2m_v)(3p/8\pi)^{2/3})]/kT\}\ . \tag{8}$$

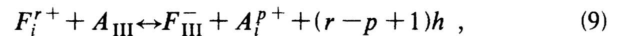
Equation (8) indicates that in the limit of very heavy doping the native defect concentration depends exponentially on the hole concentration. Therefore it is expected that the onset of degenerate statistics will result in a rapid increase of the defect concentration, and consequently it will affect all phenomena which depend on the concentrations of charged native defects. This includes impurity diffusion and self-diffusion and electrical activation of intentionally introduced dopants. Also, in contrast to the nondegenerate case, where the effect of the doping is to change the pre-exponential factor, in the case of degenerate statistics the doping leads to an apparent reduction of the defect activation energy by the amount given by the last term in the square brackets of Eq. (8). The reduction of the effective activation energy depends on the charge states of the defects and the concentration and effective mass of the free carriers, and is especially significant in the case of heavily doped n-type semiconductors with a small electron effective mass.

The transition from nondegenerate to degenerate hole statistics occurs approximately when $F_{1/2}(-E_F/kT)=1$. From Eq. (3) this corresponds to the condition $p=N_v$ for holes in the valence band, and an analogous condition $n=N_c$ for electrons in the conduction band. As we will see later, these are quite restrictive conditions for most of the semiconductor materials grown or processed at temperatures lower than 1000 K.

As an example of the semiconductor system for which the effects of electron and hole statistics are clearly discernible, we consider the case of heavily p- and n-type doped $In_{0.53}Ga_{0.47}As$ epitaxial films lattice matched to InP substrates. For a typical growth temperature of about 720 K, $N_v=3\times10^{19}\ cm^{-3}$ and $N_c=1\times10^{18}\ cm^{-3}$. Those two concentrations set the upper p- and n-type doping limits for the applicability of the standard model of the concentration-dependent defect formation. When the hole or electron concentration exceed the above limits a general carrier statistics given by Eq. (3) has to be used. To illustrate the effects of heavy doping we consider the impurity segregation in p-$In_xGa_{1-x}As$ and the doping-induced defect formation in heavily n-type doped $In_xGa_{1-x}As$ and GaAs layers. To make our considerations more transparent we will simplify the model defect reactions describing these processes. We will ignore many complexities of the processes, concentrating rather on the Fermi energy effects which are affected by the free-carrier statistics.

Recent experiments in $In_{0.53}Ga_{0.47}As$ and GaAs heavily doped with Zn (Refs. 7–9) acceptors have shown that for doping levels higher than about $10^{20}\ cm^{-3}$ a very rapid diffusion of the acceptors is observed. The onset of the diffusion-induced segregation of impurities is very abrupt. Well-defined impurity profiles were obtained for doping lower than $10^{20}\ cm^{-3}$. However, an extensive segregation was observed for doping levels higher than this threshold value. Here we will show that the rapid increase of the diffusion rate can be explained in terms of Fermi-energy-dependent formation of native defects.

It is now generally accepted that in $A_{III}B_V$ semiconductors, the diffusion of group-II acceptors, F occurs via the substitutional-interstitial process. One possible representation of the process is the kick-out mechanism[10]



$$F_i^{r+}+A_{III}\leftrightarrow F_{III}^{-}+A_i^{p+}+(r-p+1)h\ , \tag{9}$$

where the subscripts denote the lattice site location and the superscripts represent the charge state of the defects. Experiments on the group-II acceptor diffusion in III-V semiconductors indicate that $r-p+1=2$, which means that two holes at the Fermi energy are formed when reaction (9) proceeds from the left- to the right-hand side. Since A_i are mobile defects, it is reasonable to assume that at elevated temperatures the concentration of A_i attains its equilibrium value by interacting with internal and external sinks and sources. Any redistribution of F atoms will then be accommodated by proper redistribution of A_i to maintain the equilibrium for reaction (9). When these conditions are satisfied, the Fermi-level dependence of the concentration ratio of interstitial to substitutional impurity atoms is given by

$$[F_i]=[F_{III}][A_i]\exp[-2(E_F-E_{ref})]\ , \tag{10}$$

where E_{ref} is an energy reference for the substitutional-

interstitial process represented by reaction (9). As has been discussed previously the choice of the energy reference plays a crucial role in understanding of the diffusion processes at semiconductor heterointerfaces.[11] However, since here we are interested only in the relative changes induced by the Fermi level, the actual choice of the energy reference is not important in our considerations.

Migration of interstitials is a principal mode of impurity diffusion, therefore relative concentration of the group-II interstitial atoms given by Eq. (10) can be considered a measure of the Fermi-level-dependent diffusivity of group-II acceptors in p-type III-V semiconductors. This is consistent with previous analysis of the substitutional-interstitial diffusion processes[10] where it has been argued that in the standard model the diffusion of group-II acceptors in III-V semiconductors is proportional to the square of the hole concentration. Using Eq. (3) to determine the Fermi energy, one can find from Eq. (10) the relative diffusivity as a function of the hole concentration. The results of the calculations are presented in Fig. 1. Also in the same figure, the results of the standard model, assuming a nondegenerate hole gas, are shown. As expected, the results for both models agree well for the hole concentrations smaller than about 10^{19} cm^{-3}. At higher concentrations the diffusivity calculated using the general hole statistics increases very rapidly. Thus the change of the hole concentration by the factor of 2 from 1×10^{20} to 2×10^{20} cm^{-3} results in a change of the diffusivity by a factor of 50. This dramatic increase in the diffusion rate can account for the onset of an abrupt segregation of Be atoms during epitaxial growth of $In_xGa_{1-x}As$ layers.[8] It was found that spatially well-defined abrupt Be doping profiles can be obtained only for doping levels below about 10^{20} cm^{-3}. Attempts to dope at higher levels lead to a very rapid out-diffusion of the impurities in excess of the 10^{20}-cm^{-3} threshold level.

A similar saturation of the Be doping level has also been observed in GaAs.[12] This is not surprising, however, as the onset of the degenerate hole statistics is determined by the valence-band density of states N_v, which is

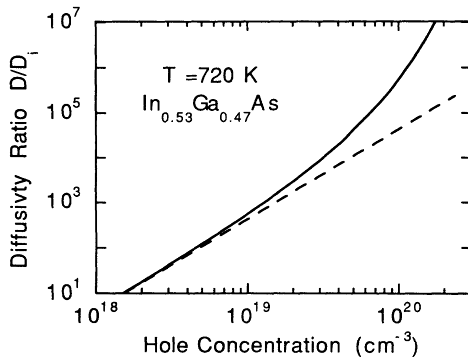


FIG. 1. Hole-concentration-dependent diffusivities of group-II acceptors normalized to the diffusivity of intrinsic $In_{0.53}Ga_{0.47}As$ at $T=720$ K. The broken line represents the diffusivity calculated using the standard model, and the solid line shows the diffusivity calculated using general carrier statistics.

practically the same for GaAs as it is for $In_xGa_{1-x}As$. It has been proposed previously that fast diffusion at high p-type doping levels results from the electrostatic repulsion of negatively charged acceptor ions in the depletion layer at the growth surface.[9] A large surface Fermi-level pinning energy of 0.9 eV had to be assumed to explain the experimentally observed segregation.[9] This explanation will have to be reevaluated in view of a recent scanning tunneling microscopy (STM) study which demonstrated that there is a very little pinning of the surface Fermi energy during molecular-beam epitaxy (MBE) growth of heavily p-type doped GaAs.[13]

Because of the small effective mass of electrons, it is anticipated that in n-$In_xGa_{1-x}As$ the effects of degenerate statistics will occur for carrier concentration much lower than in p-$In_xGa_{1-x}As$. For a growth temperature of 720 K the standard approach to the concentration-dependent diffusion is applicable only for electron concentrations smaller than 10^{18} cm^{-3}. At higher concentrations an increase of the defect formation and defect-related diffusion rates due to the onset of degenerate statistics is expected. Recent investigations of the diffusion in heterojunction bipolar transistor (HBT) structures have shown that a strong diffusion of group-II acceptors is observed in the base layer adjacent to a heavily n-type doped emitter. This effect, found in $GaAs/Al_xGa_{1-x}As$ (Ref. 14) and $In_xGa_{1-x}As/InP$ (Ref. 15) systems was explained by an increased formation rate of the Frenkel pairs on the group-III sublattice in heavily n-type doped emitter or cap layers.[5,15] The released interstitials diffuse into the p-type base causing fast diffusion of group-II acceptors via the substitutional-interstitial mechanism represented by reaction (9). In the case of the $In_xGa_{1-x}As/InP$ structure, a distinct doping threshold for this diffusion mechanism has been reported.[15] It has been found that diffusion is strongly enhanced for the electron concentration exceeding about 10^{19} cm^{-3} in the $In_xGa_{1-x}As$ layer. We will show that this diffusion enhancement is associated with an increased concentration of group-III interstitials induced by the onset of highly degenerate electron-gas conditions.

A model analysis of the diffusion of group-II acceptors in the base of a HBT structure has shown that the extent of the diffusion is determined by the excess of neutral interstitials in the n-type doped layer adjacent to the p-type base.[5] Using the results of Ref. 5, one finds that in n-type material the Fermi-level-dependent concentration of group-III interstitials can be approximated by the expression

$$[A_i^n]\simeq C\exp(E_F/2kT)\ , \qquad (11)$$

where C is a Fermi-level-independent constant.

For the standard model, i.e., for electron concentrations where nondegenerate statistics are still valid, one obtains $[A_i^n]\sim n^{1/2}$. At a temperature of 720 K this approximation is applicable only for an electron concentration smaller than 10^{18} cm^{-3}. Since all the experiments are done at much higher doping levels, it is obvious that general electron statistics have to be used in this case.

Using Eq. (3) modified to the case of an electron gas,

one can obtain the concentration of the interstitials from Eq. (11). The results of the calculations are shown in Fig. 2. In agreement with the experiment, a very abrupt increase of $[A_i^n]$ is observed for n about 10^{19} cm^{-3}. This result can be contrasted with the standard model, which predicts only a slow incremental increase of the concentration of the defects in the same electron concentration range. The results in Fig. 2 clearly indicate that in n-$In_xGa_{1-x}As$ the standard approach can be used only in a very limited concentration range. For all practically interesting electron concentrations a general description of the electron statistics is necessary.

The enhanced formation rate of defects in n-$In_xGa_{1-x}As$ can explain other phenomena observed at high doping level in this material. It has been demonstrated that an efficient intermixing of $In_{0.53}Ga_{0.47}As$/$In_{0.48}Al_{0.52}As$ superlattices occurs for electron concentrations higher than 1.3×10^{19} cm^{-3}.[16] Also, it has been shown that high-temperature annealing of a heavily doped n-$In_xGa_{1-x}As$ reduces the electron concentration to about 1.5×10^{19} cm^{-3}.[17] Intermixing of the GaAs/$As_xGa_{1-x}As$ superlattice,[4,11] as well as reduced activity of donor dopants in GaAs,[11] were previously explained by the Fermi-level-induced formation of $V_{\rm III}$. The equilibrium concentration of $V_{\rm III}^{m-}$ is given by

$$[V_{\rm III}^{m-}]=[V_{\rm III0}]\exp\left[\frac{mE_F}{kT}\right], \qquad (12)$$

where $V_{\rm III0}$ is a Fermi-level-independent constant, and m is the charge of the vacancy in an n-type material. In n-GaAs the gallium vacancies exist in the triply ionized charge state ($m=3$).[4,11]

Assuming that the formation of group-III vacancies is also responsible for an intermixing of the $In_xGa_{1-x}As/In_xAl_{1-x}As$ superlattice and saturation of the electron concentration in $In_xGa_{1-x}As$, we find that the concentration of the vacancies exponentially increases with n for $n\gtrsim10^{19}$ cm^{-3}. This exponential increase of $V_{\rm III}$ will lead to a large enhancement of the intermixing. Also, since the vacancies are triply charged acceptors, their enhanced incorporation would lead to a compensation of the intentionally introduced dopants and thus also to a saturation of free-electron concentration.[11]

The above discussion indicates that the onset of degenerate carrier statistics results in an abrupt increase of the defect concentration and hence also self- and impurity diffusion. One could thus argue that it also sets limits on the structural stability of doped semiconductor structures prepared or processed under thermal equilibrium conditions. For p- and n-type semiconductors the threshold carrier concentrations are approximately given by the conditions $p_{\rm th}=N_v$ and $n_{\rm th}=N_c$. For much higher hole and electron concentrations the semiconductor materials become intrinsically unstable to the formation of native defects. The threshold hole and electron concentrations for a number of III-V materials at $T=720$ K are shown in Figs. 3(a) and 3(b), respectively. The standard approach using nondegenerate carrier statistics is applicable only in the carrier concentration ranges $p_i<p<p_{\rm th}$ and $n_i<n<n_{\rm th}$. For concentrations exceeding the threshold values the general carrier statistics have to be used. As is seen in Fig. 3(b), in the case of n-type $In_xGa_{1-x}As$ the standard approach is valid only in a very limited concentration range between about 3×10^{17} and 10^{18} cm^{-3}.

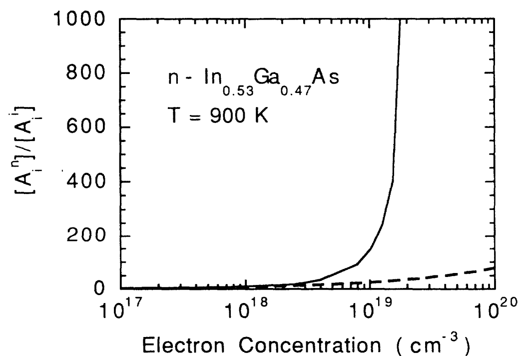


FIG. 2. Electron concentration dependence of concentration of the group-III interstitials normalized to the concentration of the interstitials in intrinsic material $[A_i^i]$. The broken line represents the concentration calculated using a standard model with nondegenerate carrier statistics wherever the solid line represents the concentration calculated using general electron statistics.

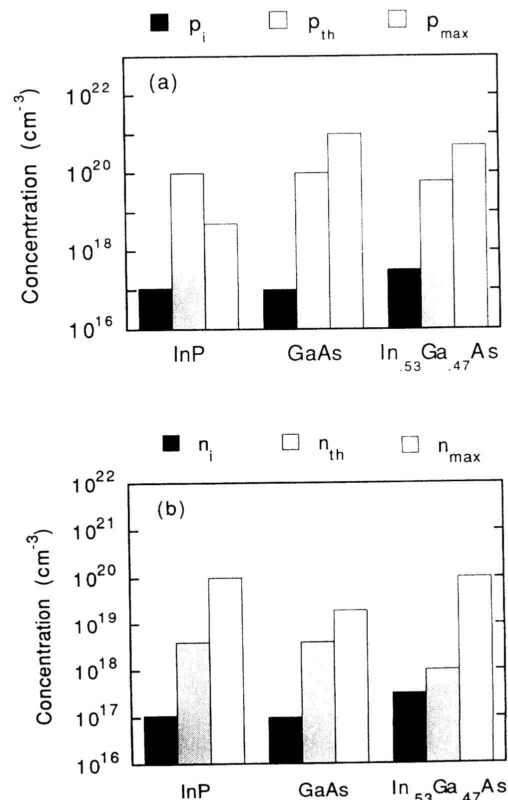


FIG. 3. Intrinsic (i) and threshold (th) carrier concentrations in different group-III-V semiconductors. The concentrations $p_{\rm max}$ (a) and $n_{\rm max}$ (b) represent the maximum hole and electron concentrations attainable in these materials.

Though we have so far discussed group III-V semiconductors, it is quite obvious that the effects of degenerate statistics are of a general nature, and they can be important is any heavily doped semiconductor. One of the processes where the effects of degenerate carrier statistics can be especially important is solid phase epitaxy (SPE). In this process a crystalline substrate induces an epitaxial regrowth of an amorphous layer. In most semiconductors SPE can occur at very low temperatures.[19] Thus for Si a regrowth of an amorphous layer occurs in the temperature range 450°–500 °C. It has been shown that intentionally introduced dopants are activated at these low temperatures, and that the recrystallization rate is dramatically enhanced at concentrations of electrically active dopants exceeding 10^{20} cm^{-3}.[20,21] Since the band density of states depends on temperature, it is expected that for typical SPE conditions the onset of degenerate statistics will already be observed at lower carrier concentrations. In Si at 750 K the Fermi energy enters the conduction (valence) band at an electron (hole) concentration of about 10^{20} cm^{-3} (5×10^{19} cm^{-3}). A rapid increase of the concentration of electrically active defect centers is expected at these threshold concentrations. This could in turn explain the enhanced regrowth rate, since it has been argued that the participation of native defects is necessary for a SPE of Si.[20]

In conclusion, we have shown that an onset of degenerate free-carrier statistics for electron or hole gases leads to a rapid increase in the formation rate of charged defects. In the limit of the highly degenerate free-carrier gas the concentration of the defects is an exponential function of the carrier concentration. The enhancement in the defect formation rate is a critical factor determining dopant activation as well as self- and impurity diffusion in semiconductors grown and/or processed at reduced temperatures.

The author would like to thank E. E. Haller, T. Kobayashi, and K. Wada for stimulating discussions. This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Science Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

[1]A. Seeger, and K. P. Chik, Phys. Status Solidi B **29**, 455 (1968).

[2]P. E. Blochl, E. Smargiassi, R. Car, D. B. Laks, W. Andreoni, and S. T. Pantelides, Phys. Rev. Lett. **70**, 2435 (1993).

[3]J.-L. Rouviere, Y. Kim, J. M. Rentschler, A. Bourret, and A. Ourmazd, Phys. Rev. Lett. **68**, 2798 (1992).

[4]T. Y. Tan, U. Gösele, and S. Yu, Crit. Rev. Solid State Mater. Sci. **17**, 47 (1991).

[5]D. G. Deppe, Appl. Phys. Lett. **56**, 370 (1990).

[6]W. Walukiewicz, in *Symposium on III-V Electronic and Photonic Device Fabrication and Performance,* edited by K. S. Jones, S. J. Pearton, and H. Kanber, MRS Symposia Proceedings No. 300 (Materials Research Society, Pittsburgh, 1993), p. 421.

[7]P. M. Enquist, J. A. Hutchby, and T. J. de Lyon, J. Appl. Phys. **63**, 4485 (1988).

[8]M. B. Panish, R. A. Hamm, D. Ritter, and H. S. Luftman, J. Cryst. Growth **112**, 343 (1991).

[9]E. F. Schubert, G. H. Gilmer, R. F. Kopf, and H. S. Luftman, Phys. Rev. **46**, 15 078 (1992).

[10]U. Gösele and F. Morehead, J. Appl. Phys. **52**, 4617 (1981).

[11]W. Walukiewicz, J. Vac. Sci. Technol. B **6**, 1257 (1988).

[12]E. F. Schubert, J. M. Kuo R. F. Kopf, H. S. Luftman, L. C. Hopkins, and N. J. Sauer, J. Appl. Phys. **67**, 1969 (1990).

[13]M. D. Pashley, K. W. Haberern, R. M. Feenstra, and P. D. Kirchner, Phys. Rev. B **48**, 4612 (1993).

[14]P. M. Enquist, J. Cryst. Growth **93**, 637 (1988).

[15]K. Kurishima, T. Kobayashi, and U. Gösele, Appl. Phys. Lett. **60**, 2496 (1992).

[16]T. Miyazawa, Y. Kawamura, and O. Mikami, Jpn. J. Appl. Phys. **27**, L1731 (1988).

[17]N. Watanabe, T. Nittono, and K. Watanabe, Appl. Phys. Lett. **61**, 1945 (1992).

[18]R. A. Hamm, M. B. Panish, R. N. Nottenburg, Y. K. Chen, and D. A. Humphrey, Appl. Phys. Lett. **54**, 2586 (1989).

[19]S. S. Lau and W. F. van der Weg, in *Thin Film Interdiffusion and Reactions,* edited by J. M. Poate, K. N. Tu, and J. W. Mayer (Wiley, New York, 1978), Chap. 12.

[20]L. Csepregi, E. F. Kennedy, T. J. Gallagher, J. W. Mayer, and T. W. Sigmon, J. Appl. Phys. **48**, 4234 (1977).

[21]C. Licoppe, and Y. I. Nissim, J. Appl. Phys. **59**, 432 (1986).

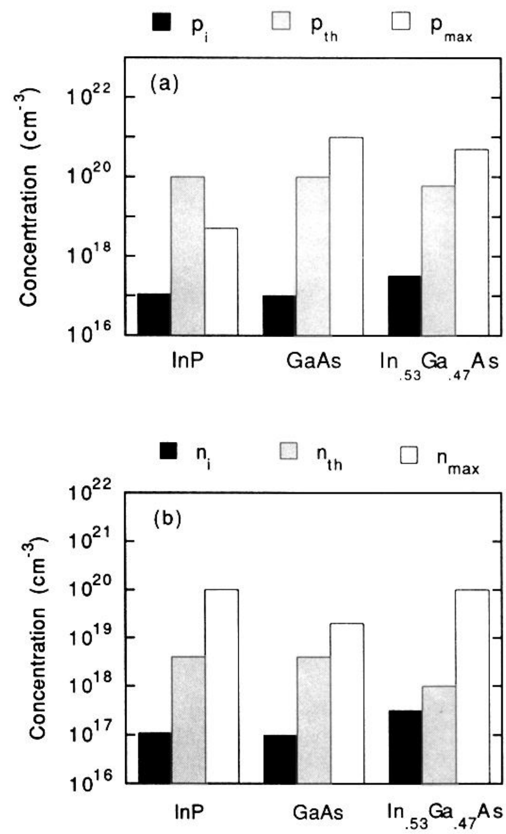


FIG. 3. Intrinsic (i) and threshold (th) carrier concentrations in different group-III-V semiconductors. The concentrations p_{max} (a) and n_{max} (b) represent the maximum hole and electron concentrations attainable in these materials.