PHYSICAL REVIEW B VOLUME 39, NUMBER 5 15 FEBRUARY 1989-I

Theory of relative native- and impurity-defect abundances in compound semiconductors and the factors that influence them

Robert W. Jansen
Complex Systems Theory Branch, Naval Research Laboratory, Washington, DC 20375-5000

Otto F. Sankey
Department of Physics, Arizona State University, Tempe, Arizona 85287-1504

(Received 6 June 1988)

An *ab initio* pseudo-atomic-orbital method [Phys. Rev. B **36**, 6520 (1987)] is used to predict relative point-defect abundances and the factors that influence them in a number of III-V and II-VI compound semiconductors. A study is presented of the trends in the native–point-defect concentrations including vacancies, interstitials, and substitutional defects as a function of stoichiometry, temperature, chemical potential, host material, and the presence of extrinsic impurities. The concentrations are predicted from equilibrium statistical mechanics by making use of the defect-formation energies.

I. INTRODUCTION

Of fundamental importance to the electronic properties of a material is the presence of defects. It has only been recently, with improvement in computational resources, that reliable quantitative studies of the energetics of defect systems have become feasible. Most of these investigations have focused on the elemental semiconductor Si, which is, by far, the most commonly used semiconductor for devices. Relatively few extensive studies have been carried out on the more complex systems involving defects in compound semiconductors. With the ever increasing technological interest in compound semiconductors, a thorough understanding of their defects is important. Of particular interest is the nature of point defects, which consist of native vacancies, antisites and interstitials, and extrinsic substitutional and interstitial impurities.

The relative abundance of point defects may in principle be predicted from equilibrium statistical mechanics once the defect reaction energies are known. The defect concentrations will depend on factors such as stoichiometry, temperature, chemical potential, and the presence of extrinsic impurities. In this paper we study the defect reactions and the resulting defect concentrations for a variety of III-V and II-VI compound semiconductors. We consider both *intrinsic* and *extrinsic* defects. Specifically, we will investigate unrelaxed tetrahedrally bonded native (intrinsic) point defects and a number of extrinsic impurities in the semiconductor compounds GaAs, GaP, ZnSe, and ZnTe. The intrinsic defects we consider are the anion and cation antisites (anion occupying a cation site and *vice versa*), A_C and C_A, the anion- and cation-site vacancies, V_A and V_C, and the anion and cation tetrahedral-site interstitial defects $A(T_A)$, $A(T_c)$, $C(T_A)$, and $C(T_C)$ at the two nonequivalent tetrahedral sites T_A (surrounded by anions) and T_C (surrounded by cations). The extrinsic impurity defects (denoted X) considered are the anion- and cation-site substitutional defects X_A and X_C, and the two interstitial-site defects $X(T_A)$ and $X(T_C)$. No other (lower symmetry) interstitial sites, such as the hexagonal site, were considered for either native or extrinsic defects.

Since defect concentrations depend exponentially on the formation energies and the formation energies vary on a scale of roughly 10 eV, our goal is to predict *trends* and to identify the *dominant* or near-dominant defects only. We are not attempting to predict absolute values of defect concentrations, but rather relative concentrations under the assumption of equilibrium. Of course, if equilibrium is not attained in the crystal, our calculations still should give an identification of which defects are most likely and which are highly unlikely.

We imagine a lattice of N_L lattice sites (there are two atoms per lattice site in the perfect zinc-blende structure) containing N_A anions, N_C cations, and N_X "X" impurities, and define a concentration $[Z]=N_Z/N_L$, where N_Z is the number of defects of type Z in the crystal. The calculation of the defect formation energies are performed using self-consistent pseudoatomic-orbital (PAO) total energies. A letter version of a few of the results presented here have previously been published.[1] In the present paper the details of the calculations are given and several new results are presented.

II. METHOD

To obtain the isolated–point-defect concentrations we first construct a complete set of independent reactions for the defects. For the eight intrinsic defects, seven (and only seven) independent defect reactions within the crystal can be constructed. A convenient set is

$$A_A + C_C \leftrightarrows A_C + C_A \ , \tag{1}$$

$$0 \leftrightarrows V_A + V_C \ , \tag{2}$$

$$A_A + V_C \leftrightarrows A_C + V_A \ , \tag{3}$$

$$A_A \leftrightarrows V_A + A(T_A)\ , \tag{4}$$

$$A_A \leftrightarrows V_A + A(T_C)\ , \tag{5}$$

$$C_C \leftrightarrows V_C + A(T_A)\ , \tag{6}$$

$$C_C \leftrightarrows V_C + A(T_C)\ . \tag{7}$$

We expect each of these reactions to be endothermic in going from left to right. The first reaction [Eq. (1)] represents the formation (or removal) of both a cation- and an anion-antisite defect, while in the second reaction [Eq. (2)] the crystal is expanded (or reduced) by one unit cell forming (or destroying) two vacancies, and so on. Three more reactions can be constructed for the single-X impurity at its four sites,

$$X_A \leftrightarrows V_A + X(T_A)\ , \tag{8}$$

$$X_A \leftrightarrows V_A + X(T_C)\ , \tag{9}$$

$$X_C \leftrightarrows V_C + X(T_A)\ . \tag{10}$$

We see that these reactions serve to couple the extrinsic defects with the intrinsic defects.

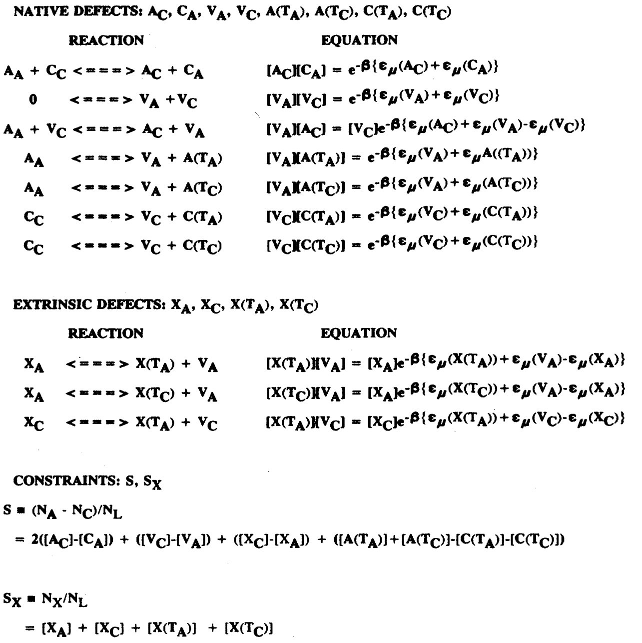
So far we have constructed $7+3=10$ equations, but there are 12 unknown defect concentrations to be determined. The final two relations come from the stoichiometry S, defined as $S=(N_A-N_c)/N_L=[A]-[C]$ and the total impurity concentration $S_X=N_X/N_L=[X]$. These two quantities are parameters, similar in the theory to temperature, and are determined in practice by the crystal grower. Be relating $[A]$, $[C]$, and $[X]$ to individual defect concentrations, we can write S as

$$S=2([A_C]-[C_A])+([V_C]-[V_A])+([X_C]-[X_A]) +\{[A(T_A)]+[A(T_C)]-[C(T_A)]-[C(T_C)]\}\ . \tag{11}$$

The quantity S is defined so that $S=0$ corresponds to perfect stoichiometry, while $S>0$ represents excess anions, and $S<0$ represents excess cations. The total-impurity concentration S_X is simply the sum of individual impurity concentrations,

$$S_X=[X_A]+[X_C]+[X(T_A)]+[X(T_C)]\ . \tag{12}$$

The ten reactions [Eqs. (1)–(10)] and corresponding mass-action equations plus the two stoichiometric constraints [Eqs. (11) and (12)] determine a unique solution for the 12 defect concentrations through use of the mass-action equations.[2-6] The ten mass-action equations are assembled opposite their corresponding reactions along with the two stoichiometric constraint equations for S and S_X in Fig. 1. For these defect systems with the two constraints, it may not be immediately apparent that the mass-action equations are appropriate. To demonstrate that the mass-action equations are in fact the correct expressions for these systems and to determine what, if any, assumptions are inherent in them, we include a formal derivation of them in the Appendix.



NATIVE DEFECTS: A_C, C_A, V_A, V_C, $A(T_A)$, $A(T_C)$, $C(T_A)$, $C(T_C)$

REACTION	EQUATION
$A_A + C_C <===> A_C + C_A$	$[A_C][C_A] = e^{-\beta\{\varepsilon_\mu(A_C)+\varepsilon_\mu(C_A)\}}$
$0 <===> V_A + V_C$	$[V_A][V_C] = e^{-\beta\{\varepsilon_\mu(V_A)+\varepsilon_\mu(V_C)\}}$
$A_A + V_C <===> A_C + V_A$	$[V_A][A_C] = [V_C]e^{-\beta\{\varepsilon_\mu(A_C)+\varepsilon_\mu(V_A)-\varepsilon_\mu(V_C)\}}$
$A_A <===> V_A + A(T_A)$	$[V_A][A(T_A)] = e^{-\beta\{\varepsilon_\mu(V_A)+\varepsilon_\mu A((T_A))\}}$
$A_A <===> V_A + A(T_C)$	$[V_A][A(T_C)] = e^{-\beta\{\varepsilon_\mu(V_A)+\varepsilon_\mu(A(T_C))\}}$
$C_C <===> V_C + C(T_A)$	$[V_C][C(T_A)] = e^{-\beta\{\varepsilon_\mu(V_C)+\varepsilon_\mu(C(T_A))\}}$
$C_C <===> V_C + C(T_C)$	$[V_C][C(T_C)] = e^{-\beta\{\varepsilon_\mu(V_C)+\varepsilon_\mu(C(T_C))\}}$

EXTRINSIC DEFECTS: X_A, X_C, $X(T_A)$, $X(T_C)$

REACTION	EQUATION
$X_A <===> X(T_A) + V_A$	$[X(T_A)][V_A] = [X_A]e^{-\beta\{\varepsilon_\mu(X(T_A))+\varepsilon_\mu(V_A)-\varepsilon_\mu(X_A)\}}$
$X_A <===> X(T_C) + V_A$	$[X(T_C)][V_A] = [X_A]e^{-\beta\{\varepsilon_\mu(X(T_C))+\varepsilon_\mu(V_A)-\varepsilon_\mu(X_A)\}}$
$X_C <===> X(T_A) + V_C$	$[X(T_A)][V_C] = [X_C]e^{-\beta\{\varepsilon_\mu(X(T_A))+\varepsilon_\mu(V_C)-\varepsilon_\mu(X_C)\}}$

CONSTRAINTS: S, S_X

$$S \equiv (N_A - N_C)/N_L$$
$$= 2([A_C]-[C_A]) + ([V_C]-[V_A]) + ([X_C]-[X_A]) + ([A(T_A)]+[A(T_C)]-[C(T_A)]-[C(T_C)])$$

$$S_X \equiv N_X/N_L$$
$$= [X_A] + [X_C] + [X(T_A)] + [X(T_C)]$$

FIG. 1. Reaction equations and stoichiometric constraints for native and extrinsic point defects in a compound semiconductor. The eight native defects considered are the anion- and cation-antisite defects A_C and C_A, the anion- and cation-site vacancies V_A and V_C, and the anion- and cation-interstitial defects at the two nonequivalent tetrahedral interstitial sites T_A (surrounded by anions) and T_C (surrounded by cations) $A(T_A)$, $A(T_C)$, $C(T_A)$, and $C(T_C)$. The extrinsic (for impurity X) defects considered are the anion- and cation-site substitutional defects X_A and X_C and the two tetrahedral site interstitial impurities $X(T_A)$ and $X(T_C)$.

A. Total-energy calculations

The reaction energies are computed within the local-density approximation (LDA) using the self-consistent pseudopotential method and a basis of pseudoatomic orbitals (PAO's). The pseudopotentials are of the Hamman, Schlüter, and Chiang type[7] and were fit to the Bachelet *et al.* form.[8] We use the Ceperly-Alder[9] exchange-correlation potential as parametrized by Perdue and Zunger.[10] A repeating supercell of defects is used so that Bloch's theorem holds. The Bloch basis states of wave vector $\mathbf{k}$ and orbital μ (μ index includes both basis index for the basis atom at τ and the s orbital index) are expanded as

$$\Phi_\mu(\mathbf{k},\mathbf{r})=(1/\sqrt{N}\,)\sum_{l} e^{i\mathbf{k}\cdot(l+\tau)}\Phi_\mu^{\mathrm{PAO}}(\mathbf{r}-l-\tau)\ , \tag{13}$$

where l are the Bravais-lattice sites. The expression for the Bloch states is identical to that used in the empirical tight-binding method, but here the orbitals $\Phi_\mu^{\mathrm{PAO}}(\mathbf{r})$ are explicitly constructed from the atomic pseudopotentials for the isolated atoms using same LDA as the solid. The atomic s and p orbitals used are those of the ground-state configuration of the atom. To avoid the need for evaluating multicenter integrals we expand the PAO's into their Fourier components and work in the formalism of plane waves.[11,12]

The PAO approach takes advantage of the desirable features of both the momentum-space formalism and the local orbital approaches. The computation of the Hamiltonian matrix elements is performed efficiently and straightforwardly in the momentum-space formalism, while the tight-binding framework is not only more phyiscally appealing and interpretable, but also drastically reduces the dimension of the Hamiltonian matrix. The PAO method is described in detail and tested extensively in a previous publication.[12] The psdueopotentials and PAO's used here are the same as those used in Ref. 12.

To perform the defect calculations, we use a basis of s and p PAO's on each of the basis atoms in a body-centered cubic (bcc) supercell consisting of 16 unit cells of the perfect crystal. A single special $\mathbf{k}$ point was used to perform the sum over an irreducible sector of the first Brillouin zone for the supercell and a cutoff of 80, 90, 80, and 90 eV was used for the plane-wave expansion of the PAO's for GaAs, GaP, ZnSe, and ZnTe, respectively. The lattice constants used for these materials were the experimental values of 5.65, 5.45, 5.66, and 6.09 Å, respectively. Lattice relaxation was not included and is expected to make some quantitative changes in these results but is not likely to upset the qualitative behavior since such effects are typically small compared to differences in formation energies. Because we use an sp^3 PAO basis and a small cutoff, the computed bandgaps are in fortuitously close agreement with experiment even though the calculations are within the LDA. Our computed bandgaps are 1.2, 2.2, 2.7, and 2.7 eV compared with the experimental values of 1.4, 2.4, 2.8, and 2.4 eV, for GaAs, GaP, ZnSe, and ZnTe, respectively. The full consequences of the use of PAO's and the LDA are difficult to assess in general. For particular situations they may become important, but here we will focus on trends which are relatively insensitive to these errors.

To test the validity of the method in reference to defects, we have used this same PAO basis and cutoff energy to make predictions of the deep levels and formation energies of the vacancy and T_2 symmetric self-interstitial in Si. These are prototype defect systems for which a number of calculations already exist.[13–17] Our calculations find that the Si vacancy forms a deep level about midgap ($E_V+0.66$ eV), and the interstitial forms a conduction-band resonance in agreement with previous calculations. Our predicted formation energies for the neutral vacancy and self-interstitial in Si were found to be 5.2 and 7.2 eV, respectively, in good agreement with those of Car *et al.*[17] who also found the interstitial to be a higher energy defect. In reference to compound semiconductors, which we are more concerned about here, we have also compared our results for the native-defect concentrations in GaAs to those of the Green's function calculations of Baraff and Schluter.[18] The predicted dominant-defect orderings and their trends are found to be identical for these two independent calculations.

B. Extracting formation energies from total energy calculations

An important consideration in predicting formation energies is the charge state of the defect.[1,16–19] A defect can act as a shallow donor or acceptor or form a deep level in the bandgap. The occupation of these levels depends on the position of the chemical potential in the gap. The energy associated with the exchange of the electrons or holes with the chemical potential is of the order of the gap and is important in most materials, especially those materials such as ZnSe which have large bandgaps. For such large-bandgap materials one cannot speak of defect-formation energies without stating the chemical potential.

To find the formation energies for a specified chemical potential μ, we first must calculate the relevant self-consistent total energies. Let us start with the native defects. We define $E(N,N)$ as the perfect crystal bcc supercell energy with N $(=16)$ anions and N cations (note that there is no dependence on chemical potential for this quantity), $E_\mu(N-1,N;V_A)$ as the total energy of the supercell at chemical potential μ with one anion removed from its site to form an anion vacancy, and so forth. For a given chemical potential, there are nine independent native-defect supercell energies that must be obtained:

$$\begin{gathered}E(N,N),\quad E_\mu(N-1,N;V_A),\quad E_\mu(N,N-1;V_C)\,,\\ E_\mu(N+1,N-1;A_C),\quad E_\mu(N-1,N+1;C_A)\,,\\ E_\mu(N+1,N;A(T_A)),\quad E_\mu(N+1,N;A(T_C))\,,\\ E_\mu(N,N+1;C(T_A)),\quad E_\mu(N,N+1;C(T_C))\,.\end{gathered}$$

The supercell calculations are performed for a given charge state n, by removing n electrons from the system and placing them at the noninteracting vacuum level (to preserve charge neutrality).[16] That is, the n electrons are not allowed to enter into the exchange correlation or Hartree potential except through the $\mathbf{g}=0$ ($\mathbf{g}$ is a reciprocal lattice vector) component of the charge density. For, say the antisite defect A_C, the true chemical-potential-dependent supercell energy $E_\mu(N+1,N-1;A_C)$ is related to the computed supercell energies for charge state n, $\xi(N+1,N-1;A_C^{+n})$, by adding the n electrons back into the system at the Fermi level,

$$\begin{aligned}E_\mu(N+1,N-1;A_C)=\xi&(N+1,N-1;A_C^{+n})\\ &+n(E_V+\mu)\,,\end{aligned}\tag{14}$$

where E_V is the valence-band edge and μ is the chemical potential (measured relative to E_V). The system will switch from the nth to the $(n+1)$th charge state in the limit where the chemical potential $\mu\rightarrow\mu_0$ satisfies

$$\begin{aligned}&\xi(N+1,N-1;A_C^{+n})+n(E_V+\mu_0)\\ &\quad=\xi(N+1,N-1;A_C^{+(n+1)})+(n+1)(\mu_0+E_V)\,.\end{aligned}$$

From this expression we see that the $n/(n+1)$ transition-state energy $\varepsilon_{\mathrm{TS}}^{n/(n+1)}$ may be defined as

$$\varepsilon_{TS}^{n/(n+1)}=\mu_0 =\xi(N+1,N-1;A_C^n) -\xi(N+1,N-1;A_C^{(n+1)})-E_V \ . \qquad (15)$$

The transition-state energies determine upper and lower bounds on the chemical potential for a given charge state. Thus the charge state is determined once the chemical potential is stated.

The defect-reaction energies are derived for a given chemical potential (charge state) by taking appropriate total-energy differences making sure to maintain particle conservation. To illustrate the procedure, consider the first reaction [Eq. (1)]. In the forward sense, this reaction may be called the "antisite-creation" reaction since two antisites are formed. The energy associated with this reaction is the sum of the anion and cation formation energies, $\varepsilon_\mu(A_C)+\varepsilon_\mu(C_A)$. We refer to such a sum as a *reaction* energy, while the individual terms in the sum are referred to as *formation* energies. In terms of supercell energies, this reaction energy becomes

$$\varepsilon_\mu(A_C)+\varepsilon_\mu(C_A)=E_\mu(N+1,N-1;A_C) +E_\mu(N-1,N+1;C_A) -2E(N,N) \ . \qquad (16a)$$

Notice that, in writing this reaction energy, care has been taken to ensure that the number of anions and cations remains constant. Observe also that it is not possible to write down an equation for $\varepsilon_\mu(A_C)$ [or $\varepsilon_\mu(C_A)$] alone in terms of antisite supercell energies which conserves the number of anions and cations. Only the *sum* of the cation- and anion-formation energies is defined by this reaction (see also the Appendix).

Consider now the vacancy-creation reaction, $0\rightarrow V_A+V_C$, which produces a pair of infinitely separated vacancies. The energy required for this reaction $\varepsilon_\mu(V_A)+\varepsilon_\mu(V_C)$ is computed by

$$\varepsilon_\mu(V_A)+\varepsilon_\mu(V_C)=E_\mu(N-1,N;V_A) +E_\mu(N,N-1;V_C) -[(2N-1)/N]E(N,N) \ . \qquad (16b)$$

Analogous to the first reaction, we see that only the sum of the anion- and cation-site vacancy energies (and not the individual energies) is definable in terms of the vacancy supercell energies if the number of anions and cations are to be conserved. The anion and cation atoms, removed from their sites to create the vacancies, are entered back into the crystal at the normal perfect-crystal lattice sites thus expanding the (infinite) crystal by one unit cell.

Continuing in this manner we obtain the equations for the remaining five reactions. They are, respectively,

$$\varepsilon_\mu(A_C)+\varepsilon_\mu(V_A)-\varepsilon_\mu(V_C) =E_\mu(N+1,N-1;A_C)+E_\mu(N-1,N;V_A) -E_\mu(N,N-1;V_C)-E(N,N) \ , \qquad (16c)$$

$$\varepsilon_\mu(V_A)+\varepsilon_\mu(A(T_A))=E_\mu(N-1,N;V_A) +E_\mu(N+1,N;A(T_A)) -2E(N,N) \ , \qquad (16d)$$

$$\varepsilon_\mu(V_A)+\varepsilon_\mu(A(T_C))=E_\mu(N-1,N;V_A) +E_\mu(N+1,N;A(T_C)) -2E(N,N) \ , \qquad (16e)$$

$$\varepsilon_\mu(V_C)+\varepsilon_\mu(C(T_A))=E_\mu(N,N-1;V_C) +E_\mu(N,N+1;C(T_A)) -2E(N,N) \ , \qquad (16f)$$

$$\varepsilon_\mu(V_A)+\varepsilon_\mu(C(T_C))=E_\mu(N,N-1;V_C) +E_\mu(N,N+1;C(T_C)) -2E(N,N) \ . \qquad (16g)$$

These seven equations [Eqs. (16)] serve to define the seven *reaction* energies uniquely. However, the eight *formation* energies are defined only to within an arbitrary constant. One should observe that the formation energies of individual defects need never have been defined since it is the reaction energies which are used in the mass-action equations and determine the defect abundances. Nevertheless, it is appealing to think in terms of formation energies which describe single defects. Individual defect-formation energies are also convenient since, while the reaction energies are normally complicated functions of the chemical potential, they can be defined in terms of formation energies which vary with chemical potential in the same manner as their corresponding supercell energies. Defined in this way, the formation energies of a defect can be found at all chemical potentials once the formation energy is known for one chemical potential. We can fix the arbitrary constant in the formation energies by, for example, setting the cation- and anion-site vacancy-formation energies equal to each other in their zero-charge state,

$$\varepsilon(V_A^0)=\varepsilon(V_C^0) \ . \qquad (17)$$

The formation energies are now uniquely determined, and because of the arbitrary constant may be negative.

The chemical-potential-dependent formation energies ε_μ are then defined to vary from the zero-charge-state value in the same way as the total supercell energies vary with chemical potential μ. Thus the slope of ε_μ at chemical potential μ is equal to the charge state n at μ and the transition-state energies ε_{TS} determine the chemical potentials at which the slope of ε_μ changes by ±1. With this prescription, all eight native-defect-formation energies may be individually determined for all important charge states. Since many of the defects involve several charge states, which are computed independently, the prediction of the concentrations for a single-compound semiconductor typically involves ~40 self-consistent supercell calculations, which illustrates the need for an expedient means of calculation such as the PAO method.

C. Extracting concentrations from formation energies

Using formation energies derived as explained in Sec. II B, we numerically solve the nonlinear mass-action equations subject to the stoichiometric constraints for S and S_X. Our goal is only to determine which of the defects are dominant and investigate the major trends in the concentrations of these defects. We are not concerned with absolute values for the concentrations, which are highly sensitive to small errors in formation energies.

In practice, it is extremely difficult to obtain useful general solutions to the mass-action equations due to the rapid variation of the exponential factor for small variations in the formation energies.[20] The concentrations generally differ from one another by several orders of magnitude. At room temperature, for instance, a 1-eV difference in formation energies roughly corresponds to a 17 order-of-magnitude difference in the concentrations. Since formation energies typically range over several ev's, utmost caution must be exercised in deriving useful solutions for a particular situation. A normally safe procedure is to decide which of the concentrations is dominant and solve for it first; subsequently deriving the remaining concentrations. The procedure we use is to first guess at the dominant *native* defect and its concentration and determine a solution for the coupled mass-action and constraint equations in a self-consistent manner. Writing the dominant native defect concentration $C_{\max}$ as

$$C_{\max}=[i]=e^{-\beta\varepsilon} ,$$

where $\beta=1/k_BT$ and i represents the dominant native defect, we start by guessing at i and ε (or equivalently $[i]$). Once $C_{\max}$ is stated, all other native defects can be directly solved for through the mass-action equations. If i was chosen correctly we should find that the computed concentration $[i]$ is indeed largest. If, for this guessed ε, the assumed i was incorrect, the native defect concentrations are recalculated with the correct (computed) i. The value of ε is then fixed by the stoichiometric constraint for S. The correct ε is the one which produces a physically reasonable solution to Eq. (11) and is determined through a root-finding procedure. The addition of extrinsic impurities can be handled analytically (with care) within the iterative process involved in determining the native defect concentrations.

III. RESULTS

We now discuss our predictions for defect concentrations on a number of systems. We first present results for the native defects in GaAs, GaP, ZnSe, and ZnTe. Second, we discuss our results for the extrinsic impurity systems Zn in GaAs (GaAs:Zn), Si in GaAs (GaAs:Si), and the group I impurities Li, Na, and K in ZnSe. The defect concentrations C are highly temperature dependent, although their *ordering* is not. We remove much of the temperature dependence from the major defects by investigating only the relative concentrations $\bar{C}=C/C_{\max}(\varepsilon_0)$ and $\bar{S}=S/C_{\max}(\varepsilon_0)$, where ε_0 is the value of ε at perfect stoichiometry ($S=S_X=0$). The dominant-defect concentration at stoichiometry, $C_{\max}(\varepsilon_0)=e^{-\beta\varepsilon_0}$, is exponentially dependent on ε_0 and may be off by orders of magnitude for small errors in ε_0. The errors in ε_0 depend strongly on chemical potential and charge state, among other factors, but are generally on the order of a few tenths of an eV. Since a 0.1-eV error in ε leads to about two orders of magnitude error in the concentrations at room temperature, there is no hope of accurately predicting absolute values for defect concentrations. Nevertheless, a great deal can be learned just by knowing the *ordering* and approximate *ratios* of the defect concentrations under various conditions. The formation-energy differences are usually much greater than the errors in ε_0. We, therefore, expect the ordering of the defect concentrations to be far less sensitive to errors in the method than the absolute values are. In deriving the defect concentrations we have used a temperature corresponding to $\beta=7$ eV^{-1}. This high temperature has no important affect on the ordering or trends of the defects.

A. Native defects

1. GaAs

The formation energies for the eight native point defects in GaAs are shown in Fig. 2. As explained previously, the deep trap associated with a given defect can transfer electrons to and from the Fermi level giving rise to a variety of charge states for the defect. Thus, the defect-formation energy is dependent on the position of

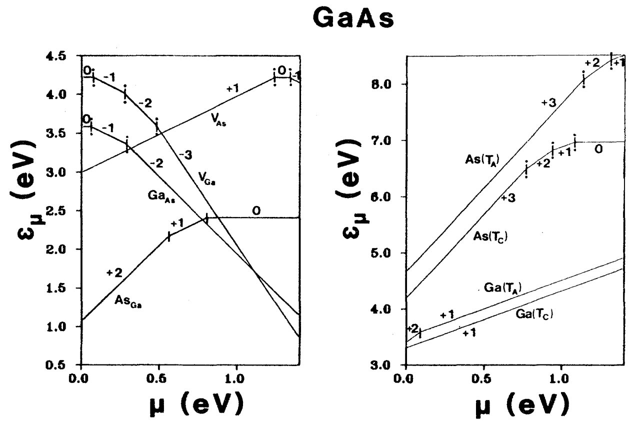


FIG. 2. Formation energies ε_μ vs chemical potential μ for native point defects in GaAs. The chemical potential is taken relative to the valence-band edge and the experimental band gap of GaAs is used. The vertical single lines represent A_1 transition-state levels and the vertical lines shouldered by dots are T_2 transition-state levels. The charge state of the defects at any chemical potential is also given.

the Fermi level. We, therefore, plot the formation energy ε_μ as a function of the chemical potential through the gap region. From the figure one can see that charge-state effects can be quite dramatic. The formation energy for the vacancy defect V_{Ga}, for instance, ranges over more than 3 eV in going from p-type GaAs, where it appears to be a relatively unimportant neutral defect, to n-type material, where it is a triple acceptor and has the lowest formation energy overall. In an analogous manner, the antisite defect $\mathrm{As}_{\mathrm{Ga}}$ goes from being a double donor with the lowest formation energy in p-type material to a neutral, relatively high-energy defect in n-type material. Except for a few special cases, we find that, overall, the defects are well separated from one another in energy for a given chemical potential. Thus, in most instances, the errors in our method, which are typically a few tenths of an eV, should not affect the ordering in the energies. In the limit of p-type GaAs, we see that the antisite defect $\mathrm{As}_{\mathrm{Ga}}$ is unquestionably the lowest energy defect. In n-type material, however, the system contains two defects, V_{Ga} and $\mathrm{Ga}_{\mathrm{As}}$, which are competitively close to one another in their formation energies. Modest errors in the deep-trap levels or other effects could, therefore, conceivably switch the ordering shown at this limit. We will find that stoichiometry plays a major role in determining which of these two defects are important, since the V_{Ga} defect will tend to favor As-rich material, while the $\mathrm{Ga}_{\mathrm{As}}$ defect will tend to form in Ga-rich material.

Figure 3 shows the native defect concentrations (with $S_X=0$) as a function of stoichiometry for both n- and p-type GaAs. Only the major defects appear in the figure; the remaining defects being below the range of the plot and relatively unimportant. The quantities are in units of $C_{\max}(\varepsilon_0)$, where ε_0 is found to be 0.9 and 2.3 eV for n- and p-type GaAs, respectively, with $\beta=7\ \mathrm{eV}^{-1}$. The dominant defect in p-type material is found to be the As on Ga antisite defect $\mathrm{As}_{\mathrm{Ga}}$ for stoichiometric material ($S=0$) and throughout much of the important range of the stoichiometric parameter S. For sufficient excess Ga, however, the dominant defect eventually changes to the $\mathrm{Ga}_{\mathrm{As}}$ antisite defect, although this may be at an unphysically high value of S. The dominant defect in n-type material is found to be the vacancy on Ga, V_{Ga}, except in the high-excess-Ga region where the antisite becomes dominant. It is interesting to observe that in both n-type and p-type material, we find that acceptor- and donor-type defects, respectively, form in an attempt to self-compensate the material. It has recently been proposed that this type of behavior may be responsible for Fermi level pinning in Schottky barriers.[21] The figure also shows that for extremely high concentrations of As, the antisite defect, $\mathrm{As}_{\mathrm{Ga}}$, becomes dominant. These results for GaAs are consistent with those of Baraff and Schlüter[18] who looked at the large nonstoichiometric limits.

A more physical interpretation of the energy ε_0 can be obtained by including only the most important defects in the mass-action equations and stoichiometric constraint S. At stoichiometry, the constraint on S [Eq. (11)] for GaAs can be written approximately as

$$2[\mathrm{As}_{\mathrm{Ga}}]+[V_{\mathrm{Ga}}]\approx 2[\mathrm{Ga}_{\mathrm{As}}]+[V_{\mathrm{Ga}}] \tag{18}$$

since interstitial defects are relatively unimportant here. From Figs. 2 and 3 we see that for p-type GaAs we may neglect $[V_{\mathrm{Ga}}]$ and assume $[V_{\mathrm{As}}]\approx[\mathrm{Ga}_{\mathrm{As}}]$. With these assumptions, the stoichiometric constraint equation [Eq. (18)] for p-type material reduces to $[\mathrm{Ga}_{\mathrm{As}}]\approx(\frac{2}{3})[\mathrm{As}_{\mathrm{Ga}}]$, so that the first mass-action equation at stoichiometry can be written as $[\mathrm{As}_{\mathrm{Ga}}]=e^{-\beta\varepsilon_0}$ with ε_0 given by

$$\varepsilon_0(p\text{-type})\approx(\tfrac{1}{2})[\varepsilon_\mu(\mathrm{As}_{\mathrm{Ga}})+\epsilon_\mu(\mathrm{Ga}_{\mathrm{As}})-\beta^{-1}\ln(\tfrac{3}{2})]\ .$$

Thus, we find at stoichiometry, the concentration of the dominant-defect $[\mathrm{As}_{\mathrm{Ga}}]$ is dependent on the sum of the formation energies and not simply $\varepsilon_\mu(\mathrm{As}_{\mathrm{Ga}})$ alone. This should not be surprising, since as stated previously, the individual defect-formation energies are not uniquely defined (also see Appendix). One should be careful not to take the assumption $[V_{\mathrm{As}}]\approx[\mathrm{Ga}_{\mathrm{As}}]$ too seriously. Small errors in the formation energies of these two defects could easily make one of them unimportant. For the extreme case where $[V_{\mathrm{As}}]\gg[\mathrm{Ga}_{\mathrm{As}}]$, the governing reaction would be $\mathrm{As}_{\mathrm{As}}\rightleftarrows 2V_{\mathrm{As}}+\mathrm{As}_{\mathrm{Ga}}$ and the dominant-defect concentration at stoichiometry becomes $[V_{\mathrm{As}}]=e^{-\beta\varepsilon_0}$ with

$$\varepsilon_0(p\text{-type})\approx(\tfrac{1}{3})[2\varepsilon_\mu(V_{\mathrm{As}})+\varepsilon_\mu(\mathrm{As}_{\mathrm{Ga}})-\beta^{-1}\ln 2]\ .$$

Alternatively, if $[\mathrm{Ga}_{\mathrm{As}}]\gg[V_{\mathrm{As}}]$, the governing reaction at stoichiometry becomes Eq. (1), and the concentration of the dominant defect is written as $[As_{\mathrm{Ga}}]=[\mathrm{Ga}_{\mathrm{As}}]=e^{-\beta\varepsilon_0}$ with

$$\varepsilon_0(p\text{-type})\approx(\tfrac{1}{2})[\varepsilon_\mu(\mathrm{As}_{\mathrm{Ga}})+\varepsilon_\mu(\mathrm{Ga}_{\mathrm{As}})]\ .$$

In n-type material we can assume the governing reaction at stoichiometry to be $\mathrm{Ga}_{\mathrm{Ga}}\leftrightarrows 2V_{\mathrm{Ga}}+\mathrm{Ga}_{\mathrm{As}}$. The stoichiometric constraint equation [Eq. (18)] then reduces to $[\mathrm{Ga}_{\mathrm{As}}]\approx(\frac{1}{2})[V_{\mathrm{Ga}}]$, which leads to the expression

$$\varepsilon_0(n\text{-type})\approx(\tfrac{1}{3})\{2\varepsilon_\mu(V_{\mathrm{Ga}})+\epsilon_\mu(\mathrm{Ga}_{\mathrm{As}})-\beta^{-1}\ln 2\}\ .$$

Again the concentration of the dominant defect $[V_{\mathrm{Ga}})$ at stoichiometry is dependent on a linear sum of formation energies. Bearing these facts in mind, it is clear that the interpretation of the formation energies must be made with caution. Nevertheless, one can see that for the case of GaAs the dominant defects are simply those with the lowest formation energies.

It is interesting to observe that, in general, for stoichiometric material ($S=0$), the constraint [Eq. (11)] alone, forces the material to contain secondary dominant defects with concentrations within a factor of 2 or so from the dominant-defect concentration. One can see from this constraint that if a vacancy is the dominant defect then there need be only one other secondary dominant defect (as is the case for n-type GaAs where the dominant-defect V_{Ga} forces the formation of the antisite defect $\mathrm{Ga}_{\mathrm{As}}$ which absorbs the Ga atoms given up in the formation of V_{Ga} defects). However, if an antisite is dominant at $S=0$, then since antisites are weighted by a factor of 2 in the stoichiometric constraint, there may be two other secondary dominant defects (as we find for *p*-type GaAs where the V_{As} and $\mathrm{Ga}_{\mathrm{As}}$ defects are secondary). Thus, at stoichiometry, the concentrations of different defects must be either identical or nearly equal (within a factor of 2 or so) even though their formation energies may be vastly different. In the limit of extreme nonstoichiometry $S \gg 0$ or $S \ll 0$, the stoichiometric constraint normally requires that the defect with the lowest formation energy of the set of defects which use up the excess atoms be dominant even though this defect may not have the lowest overall formation energy. Once the significance of the stoichiometric constraint is recognized, it becomes clear why the concentrations in Fig. 3 change with stoichiometry as they do and why small changes in the formation energies should not affect the ordering or the crossings observed in the curves. This again illustrates that the trends and orderings of the concentrations in Fig. 3 are much less sensitive to such factors as temperature than one might at first expect.

2. GaP

To present at least one other example of a III-V material we have also predicted native defect concentrations in GaP (see Fig. 4). Since GaP has an indirect gap and

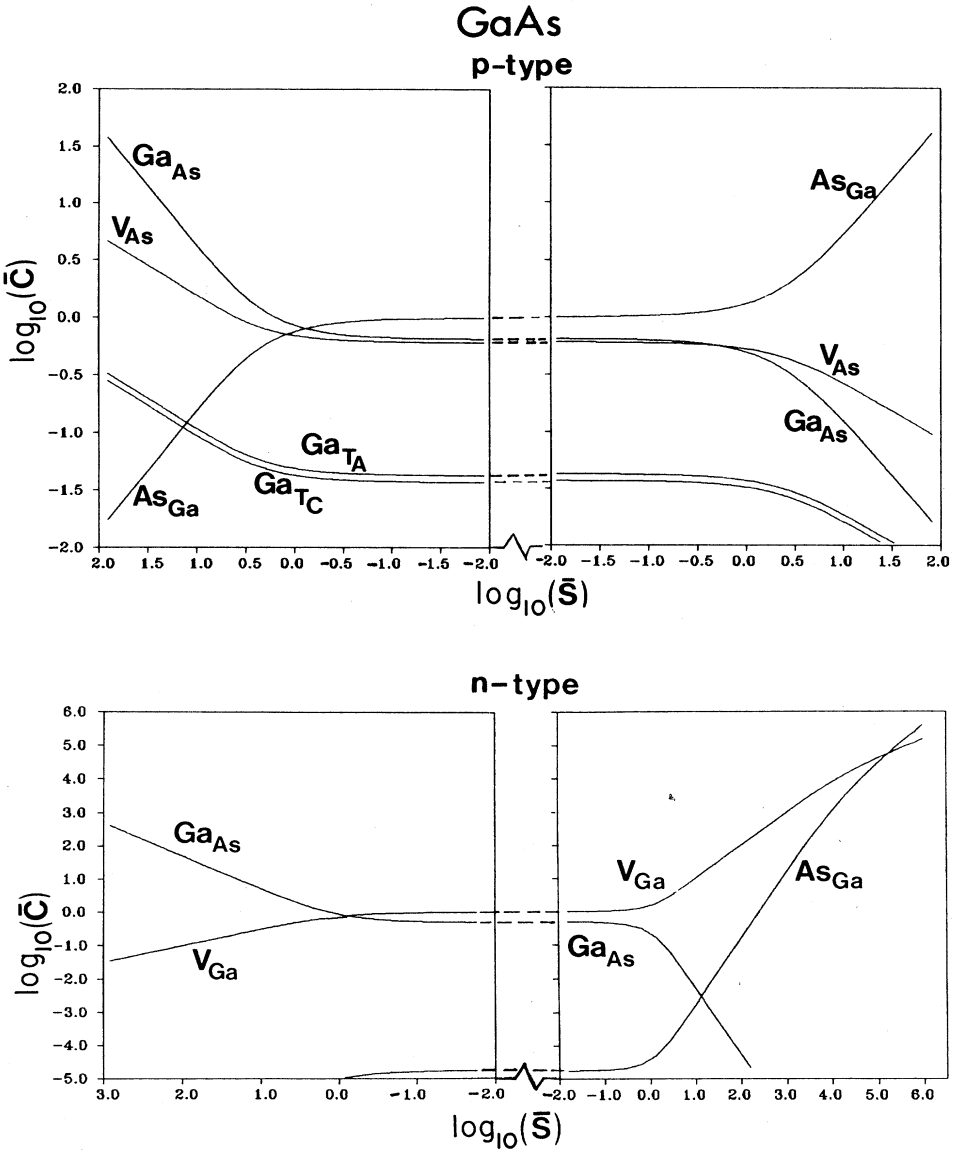


FIG. 3. Native defect concentrations as a function of stoichiometry for *p*- and *n*-type GaAs. The right panel is for positive stoichiometry S (excess As) and the left panel is for negative stoichiometry S (excess Ga). The center point between panels corresponds to perfect stoichiometry. A high temperature (corresponding to $\beta = 7\ \mathrm{eV}^{-1}$) was used, but this has no significant effect on the ordering and acts to make the differences between the defect concentrations less severe. Only the major defects appear in the graph; the remaining defects having concentrations which would appear below the plot window shown.

the density of states near the conduction bands is significantly higher than for the direct-band-gap materials, we do not expect the deep-trap levels and consequently the formation energies and concentrations to be as reliably predicted for this material. The error in the gap is ~ 0.2 eV, and the sp^3 basis does not well represent even the lowest conduction band. Nevertheless, differences in formation energies range over more than 11 eV so that we still expect the ordering and crossings in the levels to be correct. In p-type material we again find the antisite P_{Ga} to be the dominant defect over most of the important range of stoichiometries. A new feature found in GaP is that the V_P becomes dominant at high excess Ga. This was not found to be the case in GaAs where the antisite defect $\mathrm{Ga}_{\mathrm{As}}$ was dominant at high excess Ga. The value of ε_0 found for p-type GaP was 2.4 eV. Similar to what was done for GaAs, if we assume only V_P and P_{Ga} are important defects at stoichiometry, we can approximately write ε_0 as

$$\varepsilon_0(p\text{-type}) \approx (\tfrac{1}{3})[2\varepsilon_\mu(V_{\mathrm{P}}) + \varepsilon_\mu(\mathrm{P}_{\mathrm{Ga}}) - \beta^{-1}\ln 2] \; .$$

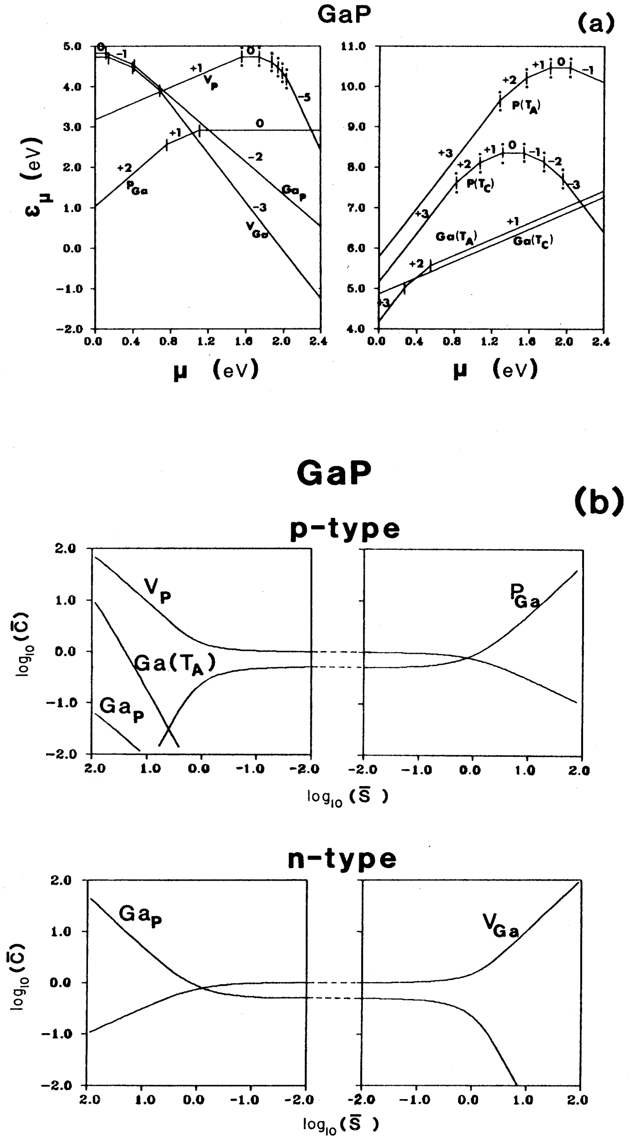


FIG. 4. (a) As in Fig. 2 but for GaP. (b) As in Fig. 3 but for GaP. The unphysically small charge-state splittings for the V_P defect are due to the mixing of the deep-trap level with the conduction bands. The edge of the theoretical conduction band lies about 0.2 eV below the experimental conduction-band edge shown.

In n-type GaP the dominant defects are V_{Ga} and Ga_{P}, which are identical to the corresponding vacancy and antisite defects in GaAs. An unphysical value of -0.68 eV was found for ε_0,

$$\varepsilon_0(n\text{-type}) \approx (\tfrac{1}{3})[2\varepsilon_\mu(V_{\mathrm{Ga}}) + \varepsilon_\mu(\mathrm{Ga}_{\mathrm{P}}) - \beta^{-1}\ln 2]$$

which may be attributed to inaccuracies in the deep-trap levels and inaccuracies in the conduction bands due to the sp^3 basis. Furthermore, the governing reaction in n-type material involves five acceptor states. Thus, small errors in the levels, which are near the valence-band edge, produce significant errors in the formation energies for chemical potentials near the conduction-band edge (i.e., for n-type material). Given that even for an error of 0.2 eV in the deep levels or the conduction-band edge, the formation energies for n-type GaP will shift by 1 eV, a small negative energy ε_0 is not too surprising. Fortunately, *these errors do not upset the ordering of the dominant defects or the shape of the curve shown in Fig. 4(b).* The only factor that could appreciably influence the curves would be the involvement of another defect, but the formation energy of the next important defect (V_P) is ~ 4 eV above the dominant-defect-formation energy, thus restricting this possibility to extreme unlikelihood.

3. *ZnSe*

In the II-VI material ZnSe we find the results for the formation energies [Fig. 5(a)] and related concentrations [Fig. 5(b)] are substantially different from those of the III-V material GaAs (Figs. 2 and 3) and GaP. Remarkably, the dominant defect in p-type ZnSe is the interstitial defect $\mathrm{Zn}(T_A)$. The corresponding interstitial defect in GaAs was found to be unimportant. For p-type, the cation in anion-site antisite defect, which was dominant in stoichiometric GaAs becomes relatively unimportant for ZnSe primarily due to the large formation energy for the $\mathrm{Se}_{\mathrm{Zn}}$ antisite.[22,23] We find interstitial Zn in the large-band-gap material ZnSe produces a deep trap near the conduction-band edge [Fig. 5(a)]. The interstitial formation energy can thus be greatly reduced by transferring the electrons from the deep trap to near the valence-band edge in p-type material. Interstitial Zn in ZnSe has recently been identified experimentally under nonequilibrium conditions by Rong and Watkins[24] and has been studied theoretically in detail along with other interstitial impurities by Jansen *et al.*[25] In n-type material the defects found in ZnSe are identical to the corresponding defects in the III-V compounds. The values of ε_0 are 1.7 eV,

$$\varepsilon_0 \approx (\tfrac{1}{3})\{2\varepsilon_\mu[\mathrm{Zn}(T_A)] + \varepsilon_\mu(\mathrm{Se_{Zn}}) - \beta^{-1}\ln 2\}$$

for p-type ZnSe, and 1.2 eV,

$$\varepsilon_0 \approx (\tfrac{1}{3})[2\varepsilon_\mu(V_{\mathrm{Zn}}) + \varepsilon_\mu(\mathrm{Zn_{Se}}) - \beta^{-1}\ln 2]$$

for n-type ZnSe.

A technologically important problem with ZnSe is the difficulty encountered in producting reliable p-type semiconducting materials.[26,27] The fact that interstitial Zn is an intrinsic dominant defect and a double donor in p-type ZnSe may be related to this difficulty. When the Fermi level approaches the valence-band edge to form p-type material, it becomes energetically favorable for the system to form interstitial Zn defects. Since interstitial Zn is a double donor, this increase in interstitial concentration may act to destroy the p-type. The ZnSe crystal in this way tends to self-compensate any attempt at making it p type. The degree to which this particular mechanism for self-compensation may be important, however, depends strongly on the absolute concentration of interstitial Zn atoms for a given chemical potential—an extremely difficult number to accurately predict theoretically.

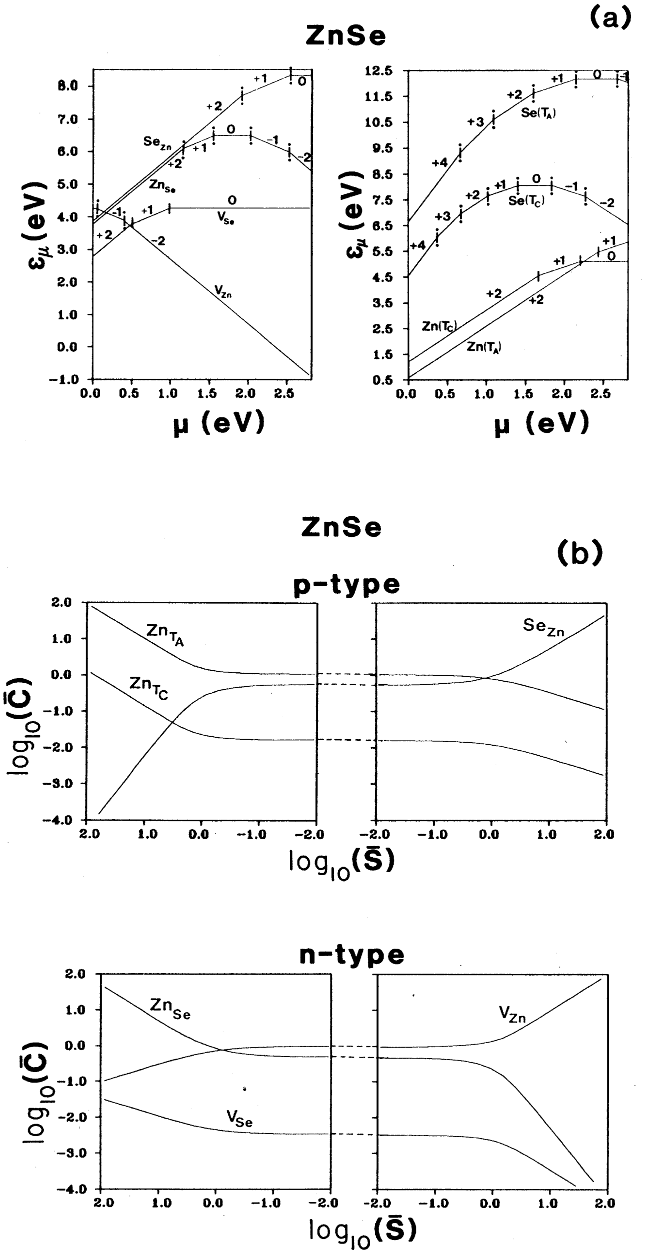


FIG. 5. (a) As in Fig. 2 but for ZnSe. (b) As in Fig. 3 but for ZnSe.

4. *ZnTe*

Further insight into the mechanism of self-compensation in ZnSe might be gained through a comparison with ZnTe. Experimentally it is known that while ZnSe is generally n type, ZnTe tends to form p-type material. This trend may appear unusual since these two materials are chemically similar. To determine what might bring about this trend, let us compare the behavior of the dominant defects in ZnTe with those of ZnSe.

In Fig. 6(a) we show our results for the formation energies of all of the native point defects in ZnTe. Comparing the formation energies of the defects in ZnTe with those in ZnSe [Fig. 5(a)], we find very little qualitative difference between the two materials. The defects with the lowest formation energies for n- and p-type ZnTe are identical to those found for ZnSe. We, therefore, might expect little or no qualitative differences in the dominant-defect concentrations for these two materials. This is in fact found to hold true for n-type material, where the dominant defects for ZnTe are predicted to be identical to those of the corresponding defects in ZnSe [see Fig. 6(b)]. Surprisingly, however, we predict a significant reordering of the major defects in p-type material. For ZnTe the two major defects are found to be Zn interstitials and the Zn-site vacancy defect, whereas in ZnSe the major defects were Zn interstitials and the anion-on-cation antisite defect. What may be surprising is that the V_{Zn} or A_C (anion and cation site) are either completely unimportant or dominant in these two quite similar materials.

To understand the origin of the difference in dominant defects between these two materials, we must again consider the stoichiometric constraints on the crystal. The most energetically favorable (lowest formation energy) defect in both materials is interstitial Zn, Zn_I (which interstitial site $I = T_A$ or T_C is unimportant here). As Zn interstitials form, the stoichiometric constraints on the system force the formation of secondary dominant defects which produce Zn atoms. There are only two defects which give up a Zn atom—the Zn-site vacancy V_{Zn} and the anion on Zn antisite defect A_{Zn}. Except for the case where these two defects have equal or nearly equal formation energies, we may in this situation, approximate the stoichiometric constraint equation [Eq. (11)] near $S = 0$ as

$$S \approx 2[A_{\mathrm{Zn}}] - [\mathrm{Zn}_I] \tag{19a}$$

or else

$$S \approx [V_{\mathrm{Zn}}]-[\mathrm{Zn}_I] \ . \tag{19b}$$

Thus, if interstitial Zn is dominant, then by stoichiometry alone, the concentration of one of the two defects, A_{Zn} or V_{Zn}, must be equal (within a factor of 2 or so) to the Zn_I concentration regardless of how high their formation energies are in comparison to other defects which do not produce Zn atoms. With this constraint in mind, it is easy to understand why the two compounds ZnSe and ZnTe differ in their dominant defects for *p*-type material. The ordering of the A_{Zn} and V_{Zn} formation energies in *p*-type ZnTe is reversed in order from those in *p*-type ZnSe. The dominant defects in *p*-type ZnSe satisfy Eq. (19a), while those in *p*-type ZnTe satisfy Eq. (19b) at stoichiometry.

Let us consider the consequences this difference in the dominant defects between these two *p*-type materials may have on their electronic properties, bearing in mind that ZnSe forms *n*-type while ZnTe forms *p*-type. The two dominant defects in *p*-type ZnTe have opposite-charge states. The Zn_I defect is a double donor, while the V_{Zn} defect is a single or possible double acceptor. The two dominant defects in ZnTe thus tend to cancel each others effect on the Fermi level. The electrons from the Zn_I are

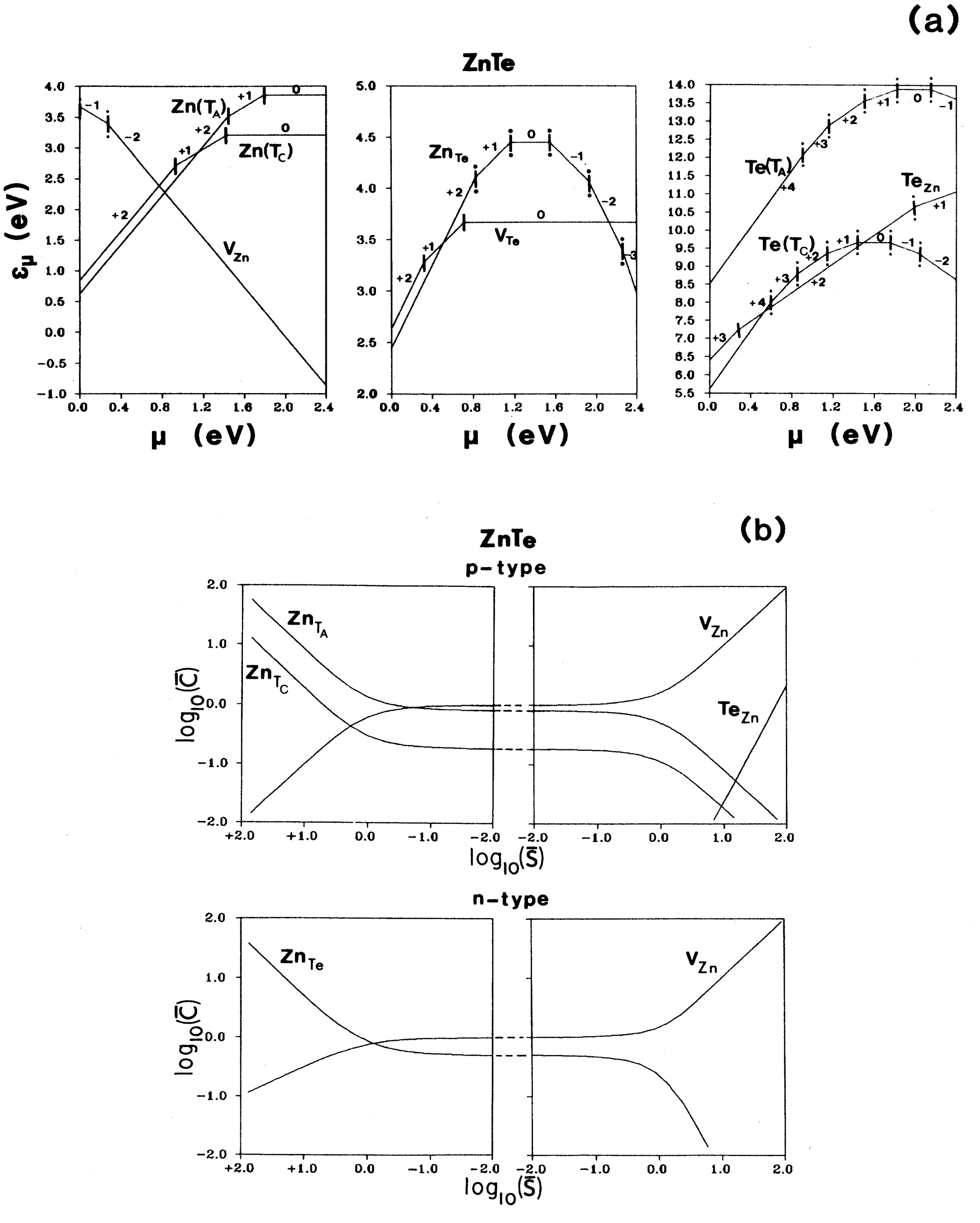


FIG. 6. (a) As in Fig. 2 but for ZnTe. (b) As in Fig. 3 but for ZnTe.

simply transferred to the V_{Zn}. Furthermore, the value of ε_0 for p-type ZnTe was found to be 2.14 eV,

$$\varepsilon_0 \approx (\tfrac{1}{2})[\varepsilon_\mu(V_{\mathrm{Zn}})+\varepsilon_\mu(\mathrm{Zn}(T_A))]$$

which is high enough to make the concentration of any defects relatively small. We, therefore, do not expect any significant self-compensation in p-type ZnTe. This is to be contrasted with the situation predicted for p-type ZnSe where both of the dominant defects were found to be donors and ε_0 was smaller. Both of these effects tend to make p-type ZnSe more self-compensating than p-type ZnTe.

For n-type material the dominant defects are V_{Zn} and Zn_A for both ZnSe and ZnTe and are acceptors. For ZnTe, however, the antisite $\mathrm{Zn}_{\mathrm{Te}}$ is a triple acceptor, while the $\mathrm{Zn}_{\mathrm{Se}}$ antisite defect in ZnSe is only a double acceptor. More importantly, in n-type ZnTe a value of 0.39 eV was found for ε_0,

$$\varepsilon_0 \approx (\tfrac{1}{3})[2\varepsilon_\mu(V_{\mathrm{Zn}})+\varepsilon_\mu(\mathrm{Zn}(T_C))-\beta^{-1}\ln 2] ,$$

which is much lower than was found for ZnSe (1.2 eV) so that the defect abundances (which promote self-compensation) are expected to be higher for n-type ZnTe.

Although there is some degree of uncertainty in these predictions for the two materials due to the limited accuracy in our formation energies, these results do support the notion that the native defects may be important in the understanding of the experimental observation the ZnSe tends to form n-type material, while ZnTe tends to form p-type material.

B. Extrinsic impurities

1. *GaAs:Zn*

Let us now investigate the effects on the defect abundances of adding an extrinsic impurity into a previously perfect compound semiconductor. We chose the impurity X to be Zn as a first example. In Figs. 7(a) and 7(b) we show the results for the formation energies and resulting concentrations, respectively, for Zn in GaAs. The concentrations [Fig. 7(b)] are plotted as a function of total-Zn concentration S_X for both n- and p-type semiconducting material at $S=0$. The concentrations are in units of the concentration of the dominant *native* defect concentration at stoichiometry ($S=S_X=0$). Our calculations predict the dominant extrinsic impurity for both n- and p-type GaAs to be the substitutional Zn on Ga defect $\mathrm{Zn}_{\mathrm{Ga}}$ in agreement with experiment.[28] What is interesting to observe is the influence the Zn has on the intrinsic-defect concentrations. In n-type material the dominant defect at $S_X=0$ (no Zn impurities) is the V_{Ga}, and the secondary dominant defect is the $\mathrm{Ga}_{\mathrm{As}}$ antisite defect. As Zn is introduced into GaAs, Zn begins to fill Ga vacancies causing their concentration to decrease. The Zn also begins to displace the Ga atoms promoting the formation of $\mathrm{Ga}_{\mathrm{As}}$ antisites. When the $\mathrm{Zn}_{\mathrm{Ga}}$ impurity becomes the dominant defect overall we find that the native defects have switched their ordering. Thus the presence of Zn has had a significant effect on the intrinsic-equilibrium concentrations. The Zn acts in a manner very similar to a change in stoichiometry. In p-type material, the influence of Zn on the intrinsic concentrations is much the same as for n-type material except that three native defects are important.

2. *GaAs:Si*

For GaAs we have already investigated an example of a group-II impurity (GaAs:Zn), and to a certain extent a group-III defect (excess Ga in GaAs) and a group-V defect (excess As in GaAs). As a last example of an extrinsic impurity in GaAs, let us now consider the group-IV impurity Si in GaAs. For the case of Zn in GaAs it was fairly obvious that the Zn would prefer the Ga site over the As site since Zn and Ga have similar electronegativities (when compared to Zn and As). For the case of Si (group IV), however, it is difficult to say *a priori* which substitutional site, Ga (group III) or As (group V), is preferred.

The results for the formation energies and corresponding concentrations for Si in GaAs are shown in Figs. 8(a)

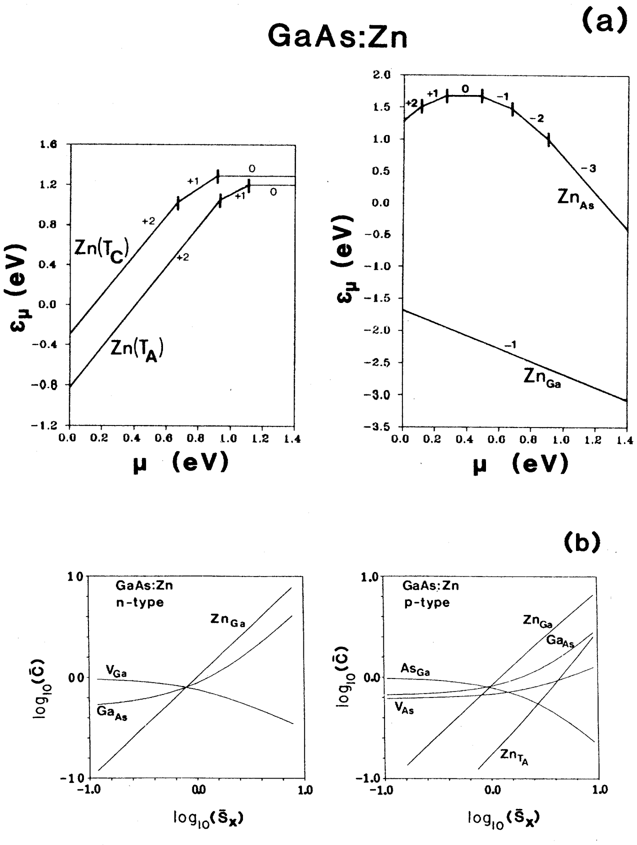


FIG. 7. Formation energies and concentrations for the extrinsic-impurity Zn in GaAs. (a) Point-defect formation energies for Zn in GaAs as a function of chemical potential. (b) Dominant-defect concentrations as a function of Zn concentration in n- and p-type GaAs at $S=0$. All concentrations are in units of the dominant-defect concentration $S=S_X=0$.

and 8(b), respectively. As is intuitively expected, we find that both substitutional impurities $\mathrm{Si_{Ga}}$ and $\mathrm{Si_{As}}$ are important when Si is incorporated into GaAs. The $\mathrm{Si_{As}}$ defect forms an acceptor state and is consequently, more abundant in n-type material, where it can form by absorbing the energy released by the transfer of electrons from near the conduction-band edge to the valence-band resonant $\mathrm{Si_{As}}$ deep trap. The $\mathrm{Si_{As}}$ substitutional defect is, in fact, predicted to be the dominant extrinsic defect in stoichiometric n-type material. This is to be compared with the situation which exists in p-type material, where we predict the Si atoms will prefer to occupy the Ga site. The defect $\mathrm{Si_{Ga}}$ forms a donor state so that the defect energy can be lowered in p-type material by transferring electrons from the conduction-band to near the valence-band edge. Thus the dominant Si-impurity defect for both n- and p-type material is a substitutional defect with the site dependent on the chemical potential.

For intrinsic GaAs host material (μ=midgap) both substitutional Si defects $\mathrm{Si_{As}}$ and $\mathrm{Si_{Ga}}$ are expected to form in nearly the same ratio with $\mathrm{Si_{Ga}}$ slightly favored. Experimentally[29–31] it is believed that the site occupied by a Si impurity in GaAs is the Ga site. This may indicate that the crossing in the two antisite-formation energies shown in Fig. 8(a) should perhaps be higher in the gap. There is also, however, some evidence that the preference for the Ga site may be related to difficulties in sample preparation.[29,30] In fact, Warwick *et al.*[29] have found that in indium-doped samples, the $\mathrm{Si_{As}}$ defect can be made predominant over $\mathrm{Si_{Ga}}$. This might be due to both the filling of Ga-site vacancies by In and the raising of the chemical potential (since the V_{Ga} are acceptors) and gives some support to the notion that a crossing such as that shown in Fig. 8(a) does occur. Further evidence that this crossing occurs comes from GaAs samples grown by liquid-phase epitaxy, where Si is found to be amphoteric.[31] In any case, one must also keep in mind that the defects present in a material grown under nonequilibrium conditions need not be the same as those predicted here where equilibrium is assumed.

The influence the concentration of Si has on the native defect concentrations is similar to that found for the case of Zn in GaAs. Unlike Zn impurities, however, the Si substitutional impurity concentration never becomes greater than any intrinsic-defect concentration in n-type material. Again this is due to the stoichiometric constraints on the system. For every Si atom replacing an As atom to form $\mathrm{Si_{As}}$, there must also exist a vacancy at

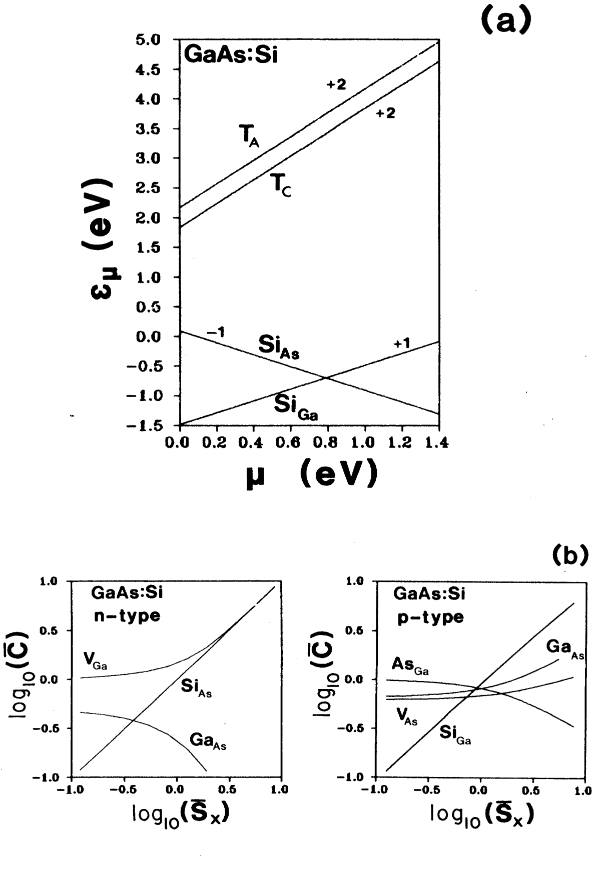


FIG. 8. As in Fig. 7 but for GaAs:Si.

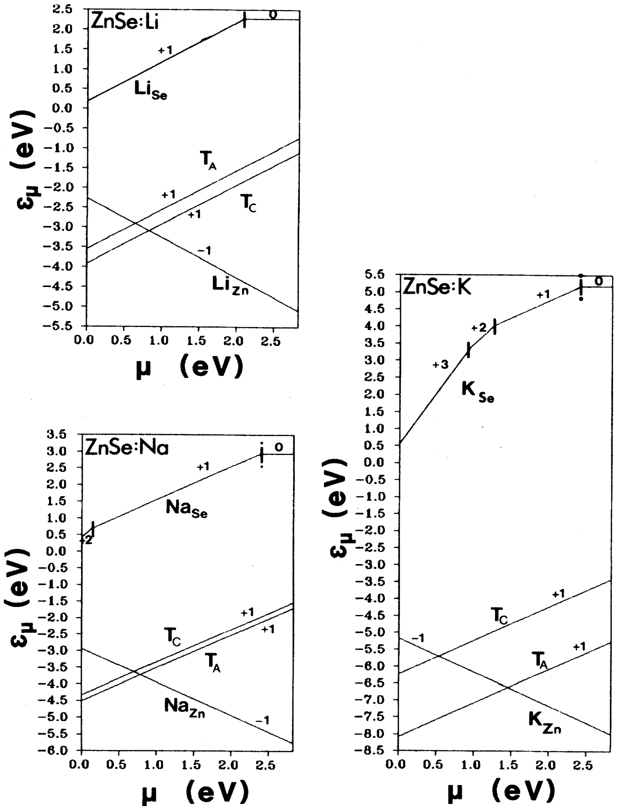


FIG. 9. Formation energies for the extrinsic group-I impurities Li, Na, and K in ZnSe as a function of chemical potential. The vertical single lines represent A_1 transition-state levels and the vertical lines shouldered by dots are T_2 transition-state levels. The charge state of the defects at any chemical potential is also given.

the Ga site if we preserve $N_A = N_C$ (i.e., stoichiometry). The displaced As and Ga atoms combine again at their normal crystal sites expanding the crystal by one unit cell. The influence the Si has on the native defect concentrations in *n*-type material is identical to the effects of an excess amount of As. In *p*-type material the site occupied by Si is the Ga site. This is the same site as was predicted for Zn impurities in *p*-type GaAs. Consequently the effect on the native defects of adding Si impurities in stoichiometric *p*-type GaAs is identical to the effect of adding Zn impurities (unless, of course, the Si donors act to affect the chemical potential).

3. ZnSe:Li,Na,K

As a final example of an important extrinsic system, consider the alkali impurities Li, Na, and K in ZnSe.[32] The formation energies for these three different impuri-

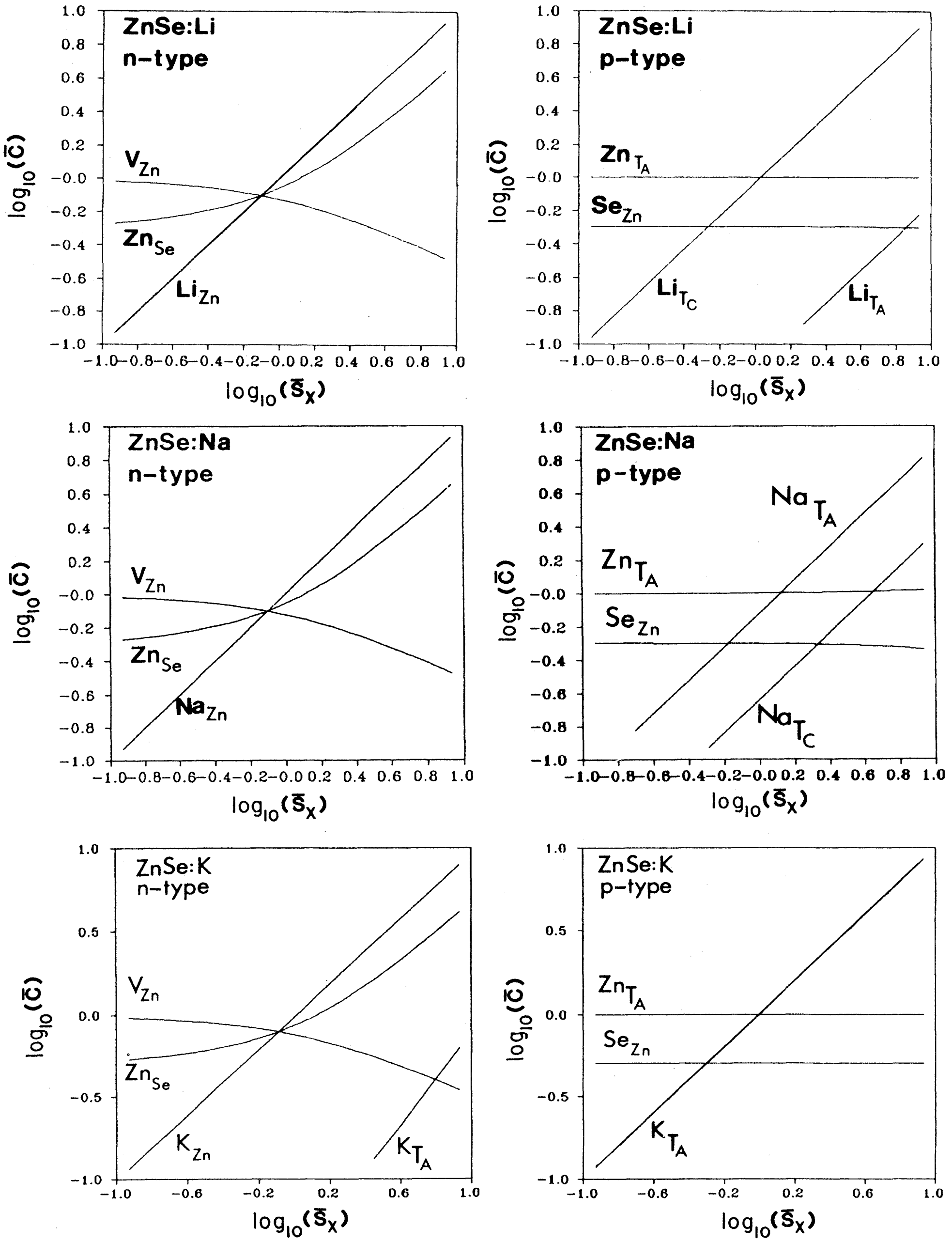


FIG. 10. Dominant defect concentrations as a function of total impurity concentration S_X in *n*- and *p*-type ZnSe at $S = 0$. All concentrations are in units of the dominant-defect concentration at $S = S_X = 0$.

ties are shown in Fig. 9. To a large extent these three elements appear to behave in nearly the same manner when incorporated into ZnSe. The Zn substitutional site and an interstitial site have the lowest formation energies in n- and p-type material, respectively, for all three alkalies. These three impurities are acceptors at the Zn substitutional site and are donors at either of the interstitial sites. The extrinsic defect with the highest formation energy is always the Se-site substitutional impurity defect. As we move from Li to K, the Se-site substitutional defect-formation energy becomes larger. This may be due to the fact that Li is closer to Se in electronegativity than Na or K. Another interesting trend is that for p-type ZnSe the dominant interstitial site switches from T_C for Li impurities to T_A for Na and K. This particular trend is opposite to that predicted from the shell-model calculation of Harding,[33] indicating that the result may be fairly sensitive to charge redistribution.

The defect concentrations for Li, Na, and K (see Fig. 10) like the formation energies from which they are derived are also very similar to each other for both n- and p-type ZnSe. In n-type material the alkali occupies the Zn site and alters the intrinsic-defect concentrations by filling up the V_{Zn} and forcing displaced Zn to occupy Se sites in a manner identical to an excess of Zn atoms in ZnSe and similar to what was found for Zn in n-type GaS. In p-type material the dominant-impurity defects are at interstitial sites. Since these sites are empty in a perfect crystal, occupying them with the extrinsic-alkali impurities does not upset the native defect concentrations as evidenced by the horizontal native defect lines in the figures.

IV. CONCLUSION

An extensive theoretical study has been made on the factors that influence the formation of native defects and extrinsic-impurity incorporation in semiconductors. We have used a complete set of reaction equations within equilibrium statistical mechanics to predict the defect concentrations of intrinsic vacancy, antisite, and tetrahedral T_d-site interstitial defects and extrinsic subtitutional and T_d-site interstitial impurities. Factors that influence the native defect abundances that we have considered are the temperature, stoichiometry, Fermi level, host materials and extrinsic impurities. The host materials we have investigated are GaAs, GaP, ZnSe, and ZnTe. We find the major native defects in the III-V materials GaAs and GaP are the antisite and vacancy defects, whereas interstitial Zn becomes important in the II-VI materials ZnSe and ZnTe. The dominant native defect in stoichiometric p-type ZnSe is predicted to be interstitial Zn which is a double donor and thus may be involved in self-compensation in this material. We find interstitial Zn to be important but not dominant in stoichiometric p-type ZnTe, which tends to form n-type material. To demonstrate how the presence of extrinsic impurity may affect the native defect concentrations, we have also considered a number of extrinsic impurities in GaAs and ZnSe. We find the presence of extrinsic impurities may profoundly alter the native defect abundances, in many cases, acting similar to a change in stoichiometry.

ACKNOWLEDGMENTS

We would like to thank the U.S. Office of Naval Research [Contract Nos. ONR-N00014-89-WX-24036 (R.W.J.), and ONR-N00014-85-K-0442 (O.F.S.)] for support. One of us (R.W.J.) would also like to thank the National Research Council for partial support.

APPENDIX: FORMAL DERIVATION OF CONCENTRATION EQUATIONS

Suppose we have a system of n defects under n_c (stoichiometric) constraints. If there are N_i-type i defects each requiring an energy ε_i to form (i.e., ε_i = formation energy for defect type i), then the total energy of the system can be written as

$$E\{N_1,N_2,\ldots,N_n\}=E_0+\sum_{i=1}^{n} N_i\varepsilon_i \ , \tag{A1}$$

where E_0 is the total energy of the perfect crystal.

There is a rather subtle aspect of Eq. (A1) that should be emphasized. The point is that although we have defined n *formation* energies, there are actually only $n-n_c$ relevant *reaction* energies which may be calculated. Thus n_c of the formation energies ε_i are arbitrary, and n_c of the coefficients N_i are not independent.

To make it obvious that this must be the case consider, for example, the effect of creating an interstitial in a formerly perfect crystal. In order to create the interstitial the system is forced to also contain a vacancy due to the stoichiometric constraints. Thus the energy of forming an interstitial cannot be uniquely defined since it is not an independent process. Therefore Eq. (A1) is only well defined because the weighted *sum* over all formation energies is a unique quantity.

We now define $W\{N_i\}$ to be the number of ways to distribute N_i-type i defects over N_L lattice sites (or interstitial sites). Using the fact that two defects of the same type i are indistinguishable we have

$$W\{N_i\}=\frac{N_L!}{(N_L-N_i)!N_i!} \ . \tag{A2}$$

If the lattice is sufficiently sparsely populated by defects, we may assume that the *distribution* of one type of defect does not influence the distribution of another type. With this assumption, the number of ways to form N_1-type 1 defects, N_2-type 2 defects, . . . , and N-type n defects is simply

$$W\{N_1,N_2,\ldots,N_n\}=W\{N_1\}W\{N_2\}\cdots W\{N_n\}$$
$$=\prod_{i=1}^{n} W\{N_i\} \ . \tag{A3}$$

Equation (A3) represents the number of microstates in the particular macrostate $\{N_1,N_2,\ldots,N_n\}$ for distinguishable defect types but indistinguishable defects of a given type. If we neglect volume and nonconfigurational entropy changes in the Gibb's free energy (since they are normally much smaller than formation energies), the partition function Z is given by

$$Z=\sum_{i=1}^{n}\sum_{N_i}{}' W\{N_1,N_2,\dots,N_n\}e^{-\beta E\{N_1,N_2,\dots,N_n\}}$$

$$=e^{-\beta E_0}\sum_{i=1}^{n}\sum_{N_i}{}'\prod_{j=2}^{n}[W\{N_j\}e^{-\beta N_j\varepsilon_j}]\ , \tag{A4}$$

where $\beta=1/k_B T$, and the prime on the sum reminds us that the sum over all possible microstates is constrained by stoichiometry. The stoichiometric constraints on the defect concentrations can be written in the form [see Eqs. (11) and (12)]

$$f_\mu\{N_1,N_2,\dots,N_n\}=N_L\gamma(\mu)-\sum_j \alpha_j(\mu)N_j=0\ , \tag{A5}$$

where $\alpha_j(\mu)$ is an integer in the range -2–2 depending on the defect, and $\gamma(\mu)$ is the external constraint, such as S (say for $\mu=1$) and S_x (for $\mu=2$).

Replacing the sum in (A4) by the value of its largest term, and minimizing the free-energy subject to the constraints (A5) we then obtain

$$0=-\ln\left[\frac{N_i}{N_L-N_i}\right]-\beta\varepsilon_i+\sum_{\mu=1}^{n_c}\lambda_\mu\alpha_i(\mu)\ , \tag{A6}$$

where $\lambda\mu$ is a constant Lagrange multiplier and the Stirling approximation ($N!\sim N\ln N-N$ for large N) has been employed. Neglecting N_i in comparison to N_L, we thus have for the concentration $[i]$

$$[i]=\frac{N_i}{N_L}=e^{-\beta\varepsilon_i}e^{\chi_i}\ , \tag{A7}$$

where χ_i is defined as $\chi_i=\sum_\mu\lambda_\mu\alpha_i(\mu)$. Notice that, without the constraints χ_i, the concentrations are completely separable and are given by their Boltzmann factors. With the constraints, we must evaluate the χ_i which involves solving for all the unknowns. However, noting that for any reaction the sum of the χ_i's for the defects on the right side equals that on the left, we can form concentration product equations which are independent of χ_i. For example, consider the reaction $0\rightleftarrows V_A+V_C$. In this case $\chi(V_A)=\lambda_1$, and $\chi(V_C)=-\lambda_1$, so that

$$[V_A][V_C]=e^{-\beta\varepsilon_\mu(V_A)}e^{+\lambda_1}e^{-\beta\varepsilon_\mu(V_C)}e^{-\lambda_1}$$

$$=\exp\{-\beta[\varepsilon_\mu(V_A)+\varepsilon_\mu(V_C)]\} \tag{A8}$$

independent of any χ_i's. This can be done for any reaction and is nothing more than the mass-action equation for the reaction. Thus we have established that the law of mass action can be used for these defect systems once the reaction (or formation) energies are determined.

---

[1] R. W. Jansen and O. F. Sankey, Solid State Commun. **64**, 197 (1987); J. Vac. Sci. Technol. B **6**, 1240 (1988).

[2] F. A. Kroger and H. J. Vink, *Solid State Physics,* edited by F. Seitz and D. Turnbull (Academic, New York, 1956), Vol. 3, p. 307; F. A. Kroger, *The Chemistry of Imperfect Solids,* 2nd ed. (North-Holland, Amsterdam, 1973).

[3] L. A. Girifalco, *Statistical Physics of Materials* (Wiley, New York, 1973), Chap. 6.

[4] R. W. Jansen, Ph.D. thesis, Arizona State University, Tempe, 1987.

[5] J. A. Van Vechten, in *Handbook on Semiconductors,* edited by S. P. Keller (North-Holland, Amsterdam, 1980), Vol. 3, p. 1.

[6] D. T. J. Hurle, J. Phys. Chem. Solids **40**, 613 (1979).

[7] D. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[8] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[9] D. M. Ceperley and G. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[10] J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[11] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).

[12] R. W. Jansen and O. F. Sankey, Phys. Rev. B **36**, 6520 (1987).

[13] G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **41**, 892 (1978).

[14] J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. Lett. **41**, 895 (1978).

[15] U. Lindefelt and A. Zunger, Phys. Rev. B **26**, 846 (1982).

[16] Y. Bar-Yam and J. D. Joannopoulos, Phys. Rev. Lett. **52**, 1129 (1984); Phys. Rev. B **30**, 1844 (1984).

[17] R. Car, D. J. Kelly, A. Oshiyama, and S. T. Pantelides, Phys. Rev. Lett. **52**, 1814 (1984).

[18] G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **55**, 1327 (1985); Phys. Rev. B **33**, 7346 (1986).

[19] S. Froyen and A. Zunger, Phys. Rev. B **34**, 7451 (1986).

[20] G. Brouwer, Phillips Res. Rep. **9**, 366 (1954).

[21] W. Walukiewicz, Phys. Rev. B **37**, 4760 (1988).

[22] E. A. Kraut and W. A. Harrison, J. Vac. Sci. Technol. B **2**, 409 (1984).

[23] J. A. Van Vechten, J. Electrochem. Soc. **122**, 423 (1975).

[24] F. Rong and G. D. Watkins, Phys. Rev. Lett. **56**, 2310 (1986).

[25] R. W. Jansen, D. S. Wolde-Kidane, and O. F. Sankey, J. Appl. Phys. **64**, 2415 (1988).

[26] T. Yasuda, I. Mitsuishi, and H. Kukimoto, Appl. Phys. Lett. **52**, 57 (1988), and references therein.

[27] P. J. Dean, *Defects and Radiation Effects in Semiconductors,* 1978 (Inst. Phys. Conf. Ser. 46, 1979), p. 100.

[28] R. K. Watts, *Point Defects in Crystals* (Wiley, New York, 1977), p. 224.

[29] C. A. Warwick, H. Ono, M. Kuzuhara, and J. Matsui, Jpn. J. Appl. Phys. **26**, L1398 (1987).

[30] A. R. Von Neida, S. J. Pearton, M. Stavola, and R. Caruso, in *Gallium Arsenide and Related Compounds 1986, Proceedings of the Thirteenth International Symposium, 1986, Las Vegas,* edited by W. T. Lindley (Institute of Physics and Physical Society, Bristol, England, 1987), p. 57.

[31] A. A. Bergh and P. J. Dean, *Light-Emitting Diodes* (Clarendon, Oxford, 1976), p. 432.

[32] See also G. F. Neumark and C. R. A. Catlow, J. Phys. C **17**, 6087 (1984); H. Chacham, J. L. A. Alvas, and M. L. De Siqueira, Solid State Commun. **64**, 863 (1987); G. F. Neumark, Phys. Rev. B **37**, 4778 (1988).

[33] J. H. Harding, J. Phys. C **14**, 5049 (1981).