Reprint

FESTKÖRPER PROBLEME

ADVANCES IN SOLID STATE PHYSICS 29

VIEWEG

Chemical Binding, Stability and Metastability of Defects in Semiconductors

Matthias Scheffler

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6,
D-1000 Berlin 33, Federal Republic of Germany

Summary: The paper describes recent theoretical studies of the electronic and structural properties of defects in semiconductors. In particular we discuss iron-acceptor pairs and chalcogen pairs in silicon, the As_{Ga}-As_i pair in GaAs and the As_{Ga} antisite defect. The results explain mechanisms which give rise to pair formation and they show possibilities of structural metastabilities.

1 Introduction

In two earlier review papers we discussed the electronic structure [1], the chemical bond between a defect and the crystal and lattice distortions [2] of isolated point defects in semiconductors. The present paper deals with defect-defect interactions with special emphasis on metastabilities. Four different examples of defect systems, for which parameter-free density-functional theory (DFT) calculations have been performed in the last years, are described. The methods and techniques of these studies have been different but are not discussed here, because we focus on the results. The interested reader is referred to the original papers. The following Section 2 sketches the basic theory and defines what we call the configurational coordinate diagram. Section 3 describes the iron-aluminum pair in silicon, which at this time is the best understood metastable defect, Then in Section 4, chalcogen pairing in silicon is considered, to discuss the effects of covalent defect-defect interactions [3]. Section 5 summarizes results for a defect pair in GaAs, namely the distant As-antisite As-interstitial pair [4]. This center has recently attracted much attention, because it was "experimentally identified" [5, 6] as the famous EL2 center. Our theoretical results rise some doubts [4] to this identification and its suggested metastability [7]. In Section 6 we discuss calculations of the isolated As_{Ga} antisite in GaAs. In contrast to what was hitherto expected about this defect, the calculations predict that it can undergo an optically inducible structural transition, ending as a $V_{Ga}As_i$ defect pair [8, 9]. The latter being a metastable configuration. The analysis of the theoretical results suggests that this new type of metastability might be important also for other centers in semiconductors. Section 7 concludes this paper.

2 Basic Aspects

Within the adiabatic approximation for the electrons of the quantum-mechanical, interacting many-atom problem, the positions and dynamics of the nuclei are described by the Hamiltonian [11]

$$H = \sum_i \frac{-\hbar^2}{2M_i} \nabla^2_{\mathbf{R}_i} + U^{static}(\{\mathbf{R}_i\}) , \tag{1}$$

where $U^{static}(\{\mathbf{R}_i\})$ is the electron ground-state total energy, calculated for the interacting many-electron system with a static, external field

$$v_{ext.}(\mathbf{r};\{\mathbf{R}_i\}) = \frac{1}{4\pi\varepsilon_0} \sum_i \frac{e^2 Z_i}{|\mathbf{r} - \mathbf{R}_i|} . \tag{2}$$

Here Z_i and $\mathbf{R}_i$ are the atomic numbers and positions, respectively. If the electron ground-state particle density is $n(\mathbf{r})$ and the magnetization density is $m(\mathbf{r})$, U^{static} is given by

$$U^{static}(\{\mathbf{R}_i\}) = \frac{1}{8\pi\varepsilon_0} \sum_{\substack{i,j \\ i \neq j}} \frac{e^2 Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} + \int v_{ext.}(\mathbf{r};\{\mathbf{R}_i\})\, n(\mathbf{r})\, d^3\mathbf{r}$$

$$+ \frac{1}{8\pi\varepsilon_0} \int\int \frac{e^2 n(\mathbf{r})\, n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' + F[n,m] , \tag{3}$$

where the functional $F[n,m]$ contains the kinetic energy as well as the exchange-correlation interaction of the electrons. $U^{static}(\{\mathbf{R}_i\})$ is the static contribution to the thermodynamic internal energy U. In a complete thermodynamic description the zero-point and temperature-induced vibrations, which follow from eq. (1), must be considered as well [12, 13]. The ground-state densities $n(\mathbf{r})$ and $m(\mathbf{r})$ and $F[n,m]$ can be evaluated using density-functional theory [14-16]. Usually a self-consistent approach is applied to evaluate $n(\mathbf{r})$ and $m(\mathbf{r})$. Approximate densities also give accurate results, provided that the functional $F[n,m]$ is evaluated without destroying the variational principle [17-19]. On the other hand, approximations of U^{static} in terms of a superposition of two- and three-body potentials, of valence-force models, or of an estimate by empirical tight-binding theory work only in a limited region of configuration space $\{\mathbf{R}_i\}$, and it is not clear when such approximations break down.

In our calculations we use the Kohn-Sham Ansatz [15, 16]

$$F[n,m] = T_s[n] + E_{xc}[n,m] , \tag{4}$$

with T_s the kinetic-energy functional of non-interacting Fermions and E_{xc} the exchange-correlation functional. Although the functional T_s is not known explicitly, its correct evaluation is important and can be indeed achieved by solving the Kohn-Sham equation [15, 16]. The main approximation in our calculations of $n(\mathbf{r})$, $m(\mathbf{r})$ and $U^{static}(\{\mathbf{R}_i\})$ is the local-spin-density approximation (LSDA) to E_{xc}, or if $m(\mathbf{r})$ is small and spin splitting negligible, the local-density approximation (LDA). This treatment of exchange and correlation is exact for interacting, many-electron systems of constant densities. Experience over the last years has shown that the jellia-derived exchange-correlation functional is a good approximation also for the description of the solid state. The approximation gets less accurate if densities are very localized, as for example in f-electron systems. In general, a cautious interpretation of theoretical results is necessary if two ground states of different symmetry (or different magnetization) are close in energy. This result will then still be correct, but it may be uncertain which of them is the true ground state.

Eq. (3) should be evaluated under the constraints that the number of electrons

$$N = \int n(\mathbf{r}) \; d^3\mathbf{r} \tag{5}$$

as well is the total spin,

$$S = \int m(\mathbf{r}) \; d^3\mathbf{r} \tag{6}$$

is fixed. A different number of electrons N, i.e. a different charge state, or a different spin S, will give rise to a different total-energy surface $U^{static}(\{\mathbf{R}_i\})$. Each of these total-energy surfaces is an electronic ground state for the constraint given by N and S. If the $\{\mathbf{R}_i\}$ have a point or translation symmetry, then for each irreducible representation of the symmetry group there is a different many-electron wave function. The total-energy surfaces of these states will be different, and in general the absolute minima of these different total-energy surfaces will be at different points in configuration space $\{\mathbf{R}_i\}$. If these points are well separated this gives rise to bi- or multistabilities.

The evaluation of eqs. (3)-(6) requires complicated methods and techniques. This is in particular so for low-symmetry systems, as for example defects, where Bloch's theorem is not valid. Such methods and techniques have been developed only during the last years (see Ref. 1, 20-32 and references therein). Up to date the applicability of these methods and techniques is still limited to special systems and new ideas and improvements of the theory are still important. Several controllable approximations are necessary in an actual first-principles calculation. If carefully applied they will not significantly affect the results for $n(\mathbf{r})$, $m(\mathbf{r})$ and $U^{static}(\{\mathbf{R}_i\})$, but they induce certain numerical inaccuracies. The most accurate method developed so far is the self-consistent pseudopotential Green-function method [1, 20, 21, 24-26] which, if used together with first-principle, norm-conserving pseudopotentials [33-35], is essentially exact. In the LMTO Green-function method [22] the approximation of spherical potentials is introduced. Because of the variational principle in DFT, this is usually not a severe approximation, but it prevents to evaluate defect-induced lattice

distortions. Cluster methods suffer from more severe (sometimes uncontrollable) problems: They impose artificial boundary conditions to the wave-functions and they localize the wave functions and charge densities to the size of the cluster, which can cause a wrong description of covalent binding. The super-cell approach also suffers from these problems of cluster approximations, but the boundary conditions are better, and with modern techniques it is now possible to take a cell size of more than fifty atoms. This allows for a systematic test of cell-size induced inaccuracies. Because of the high complexity of first-principles methods there is always a risk that some aspects may be overlooked in the calculations. It is therefore good to know if two independent groups, hopefully using different techniques, arrived at the same theoretical results.

From the knowledge of $U^{static}(\{\mathbf{R}_i\})$, the forces and force constants with respect to nuclear displacements can be evaluated and in this way it is now also possible to evaluate the vibrational contributions to the internal energy, to determine entropies and to evaluate any desired thermodynamic potential [12, 13, 36]. Two configurations which have similar energy may differ in their configurational, vibrational and electron-hole entropies. Then at $T>0$ K one of them will be favored over the other. Because of space limitations we consider in this paper only low temperature situations. Thus, entropy driven or hindered effects will be neglected.

$U^{static}(\{\mathbf{R}_i\})$ defines an energy surface over the multidimensional space of all atomic coordinates. A configurational coordinate diagram is a section of this total-energy surface along a line of a generalized coordinate which serves to describe a collective nuclear rearrangement. For example, for perfect crystals with one atom per unit cell, a configurational coordinate may be the lattice parameter of a chosen Bravais lattice. The corresponding total-energy curve (configurational coordinate diagram) will usually have only one minimum. For a defect system one interesting configurational coordinate may be the position of the impurity. If the positions of all other nuclei are relaxed according to the impurity position, we move along a valley of the total-energy surface $U^{static}(\{\mathbf{R}_i\})$. In general, the total energy as a function of the impurity position will have several minima. The lowest is the $T=0$ K stable geometry and the others are metastable. If the energy differences between stable and metastable structures are small, it will be possible to excite the defect into one or more metastable minima. In fact, it is also possible that defects are frozen in at metastable configurations, so that the stable one is practically not present. The main purpose of a configurational coordinate diagram is to explain how a system changes from one minimum to another. Usually this is shown in a plot with only one configurational coordinate. One should not forget, however, that the complete picture can be much more complicated, because many ways are possible to go from one point in configuration space to another.

The relative stabilities (or lifetimes) of metastable structures depend on the barriers between them and the stable configuration and on the possible channels to overcome these barriers. Metastable systems can have a considerable lifetime. We remind on the diamond crystal, which is a metastable crystalline structure of carbon, the stable one being graphite. Barriers can be overcome by thermal energy but also by other forms of energy, as for example light, because an excited electronic state and also an ionized system can have a different total-energy

surface, driving the system to a different structure from where the true, stable ground state can be possibly reached more easily (see Section 6 below).

3 Iron-Aluminum Pairs in Silicon

The pairs of substitutional aluminum with interstitial iron are the best understood metastable centers and there is no significant dispute on the mechanisms of pairing and metastability. Therefore these systems serve as the prototype example of iron-acceptor pairs [37-44] and for some general aspects of metastable defects. A too simple generalization of the properties of donor-acceptor pairs to other metastable systems is, however, not possible.

Because of its high solubility and diffusion coefficient [45] iron is often present in Si samples and it is easily introduced during heat treatment. Because of its low migration energy (0.69 eV [45]) it is readily involved in defect reactions [41, 42]. Using EPR on Al doped Si samples van Kooten et al. [43, 44] could identify two different paramagnetic FeAl pairs. One was oriented in the <111> direction the other in the <100> direction. The pairing reaction takes place even at room temperature and heating the samples to 90° C destroys the pairs [43, 44]. In agreement with a DLTS analysis of Chantre and Bois [39], the two pairs are interpreted as interstitial Fe which occupies the nearest and the second-nearest neighbor T_d interstitial site away from a substitutional Al. The <111> oriented pair then has a separation of about 2.35 Å and the <100> oriented pair has a separation of about 2.72 Å.

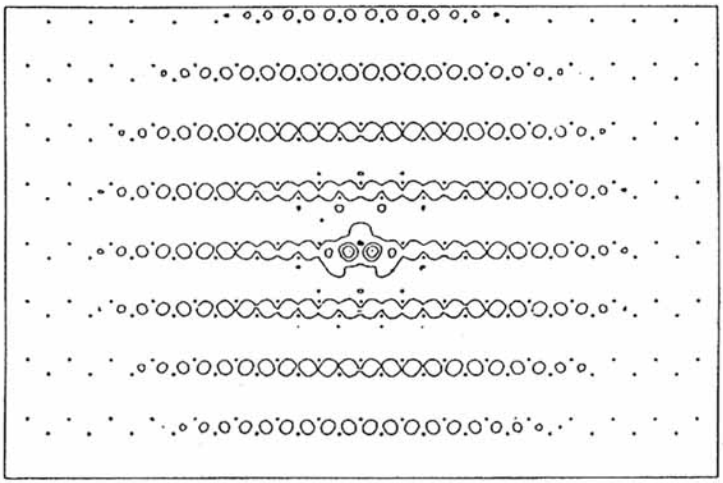

Fig. 1 Wave function (squared) of substitutional aluminum in silicon in the (110) plane. It is calculated using a Green-function approach [46]. Small dots denote the nuclei. Note that the wave function follows essentially the zig-zag bonding chain of the crystal.

We start with a short description of the isolated partners of the pair and then discuss the defect-defect interaction. Isolated substitutional aluminum in silicon is a shallow acceptor, with its energy level at 0.07 eV above the valence-band edge [40]. The wave function is sketched in Fig. 1. It originates from the valence-band states at Γ and extends over more than hundred atoms, essentially following the zig-zag bonding chains of the crystal. The defect-induced potential can be described in the language of a Green-function approach: In a perfect Si crystal one proton of a Si atom is removed, or compensated by a negative charge. This transmutes the Si into an Al nucleus and changes the *external* potential (eq. (2)) by $e^2/4\pi\varepsilon_0 r$. The electrons of the Si valence band will screen this negative charge so that (in an average) the *effective* potential attains a $e^2/4\pi\varepsilon_0\varepsilon r$ behavior. Because we are dealing with a shallow defect (extended wave function) the number of electrons in the shallow level, i.e. the defect charge state, plays only a minor role and affects the potential only at distances larger than about 20-50 Å. At closer distances self-consistent Green-function calculations show that screening in Si is very efficient and that the simple description in terms of a static dielectric constant, ε, is approximately valid already for distances larger than one interatomic spacing [2].

Iron in silicon occupies the tetrahedral interstitial site [37]. The electronic structure of interstitial iron can be understood in terms of the free atom transition-metal 3d orbitals, which are split by the crystal field and some covalent interaction with the crystal valence and conduction band states [24, 25]. Because of the strong localization of the 3d-like states it is important to take the spin-spin interaction into account. Then the iron 3d level gives rise to t_2-up, e-up, t_2-down and e-down states with increasing energy [47, 24] (see Fig. 2).

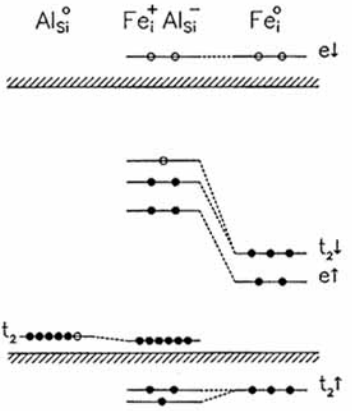

Fig. 2
Single-particle energies and the interaction of substitutional aluminum (left) and interstitial iron (right) to form a nearest-neighbor pair (middle).

The iron-induced deep level is due to the t_2-down state, which in the neutral impurity is filled with three electrons. This electronic structure gives rise to only one transition-state level in the gap, namely the 0/+ transition where the occupancy of t_2-down changes from three to two. The transition level is calculated at 0.25 eV above the valence band [24], which is close to the experimental result of 0.39 eV [40]. When one electron is removed from the t_2-down state, the single-particle energies move down in energy. Therefore, the +/2+ transition is not possible.

Whereas the effect of spin polarization is important for iron, it is irrelevant for the extended aluminum-induced wave functions. In aluminum doped samples, one of the iron 3d electrons will be transferred to the shallow aluminum state (Fig. 1) and the positively charged iron interstitial will be attracted by the aluminum $e^2/4\pi\varepsilon_0\varepsilon r$ potential. From the above description of the wave functions of the two partners, it is clear that there is practically no overlap between the very extended aluminum state, which avoids the interstitial region (Fig. 1), and the very localized interstitial iron 3d-like state. Because of these special conditions and because of the efficient screening of defects in silicon (see above and Ref. 2) the binding mechanism of the pair is largely ionic in character with only little covalent contributions. Recent cluster calculations on a similar pair, namely FeB [48], arrived at a different picture. We believe, that the cluster calculations overestimate the covalent effects, because the acceptor wave function is artificially localized to the cluster size of twenty five Si atoms.

In the pair the iron-like levels are shifted to higher energies because of the repulsive interaction with the screened Al potential (see Fig. 2). Therefore, in difference to isolated Fe impurities, the pair has two iron-like transition levels in the gap: The pair +/0 level, which corresponds to the Fe 2+/+ transition, and the pair 0/− level, which corresponds to the Fe +/0 transition. For the <111>

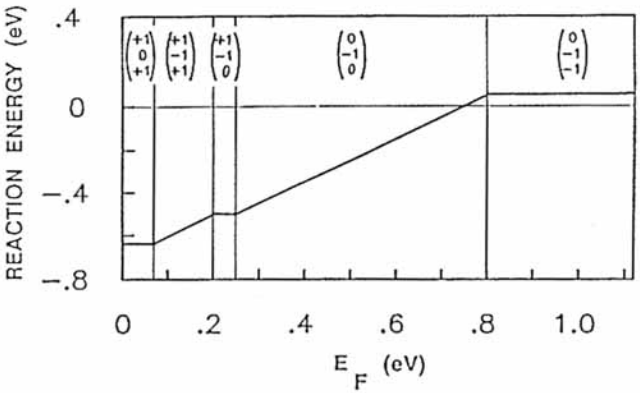

Fig. 3 Reaction energy for the formation of a nearest neighbor Fe_iAl_{Si} impurity pair in silicon (see eq. (7)). The charges of the different species (l, m, n) in eq. (7) are given at the top of the figure. Zero of E_F is the top of the valence band. A negative reaction energy corresponds to an attractive interaction (exothermal process).

pair the two levels are calculated at 0.2 and 0.8 eV above the valence band. These values are obtained without a self-consistent calculation for the pair, but taking only the self-consistent Al defect potential into account as it acts at the nearest neighbor T_d site.

The reaction energy of pair formation is displayed in Fig. 3. The considered reaction reads

$$Fe_i^{(l)} + Al_{Si}^{(m)} + (l+m)\ e^-(E_F) \rightleftharpoons (Fe_i Al_{Si})^{(n)} + n\,e^-(E_F). \tag{7}$$

It is important to note that the Fermi level acts as a reservoir of electrons and electrons can be exchanged in the reaction. $e^-(E_F)$ denotes an electron at the Fermi level. Therefore the binding energy (the energy of the right side of the reaction eq. (7) minus the left side) depends on the Fermi energy. A negative reaction energy means that pairs are energetically more favorable than the dissociated centers. Still, because of the higher configurational entropy of dissociated defects, some unpaired point defects will be present as well. Figure 3 shows that two important aspects can favor or disfavor the formation and stability of such pairs. If the pair is neutral and the two dissociated components are oppositely charged, the pair will be formed because interstitial Fe^+ is attracted by the aluminum $e^2/4\pi\varepsilon_0\varepsilon r$ potential over a long distance, and the Coulomb interaction also holds the pair together. In the ionic model the binding energy of the Fe^+Al^- pair is $e^2/4\pi\varepsilon_0\varepsilon S_1$. For the <111> pair and with $\varepsilon = 11.8$ and $S_1 = 2.35$ Å this gives 0.50 eV. If pair formation is accompanied by an exchange of electrons with the Fermi level, there is an additional and possibly more important (indirect) interaction [3]: If for a given Fermi energy the pair has more electrons than the two dissociated components, pair formation is favored with increasing Fermi energy. If the pair holds less electrons than the dissociated components, pair formation gets less favorable with increasing Fermi energy (see the energy range 0.07-0.2 eV and 0.25-0.8 eV in Fig. 3). In fact, at $E_F > 0.7$ eV the Fe^+Al^- pair becomes unstable. In particular neutral iron is not expected to pair with aluminum.

Using the law of mass action [49, 50], the energies of Fig. 3 can be used to estimate the pair concentration relative to that of the dissociated, isolated point defects, in thermodynamic equilibrium. And from the energy difference of the <111> and <100> pairs the relative concentration of the two pairs can be obtained. For the Fe^+Al^- pair the ionic model gives a difference of the Gibbs free energy between the two pair configurations of $\Delta G = 0.07$ eV - 0.4 $k_B T$. Then at room temperature and in thermodynamic equilibrium the concentration ratio of <111> to <100> pairs should be 14:1. Thus, 7% of the pairs should be <100> oriented. For the $Fe^{2+}Al^-$ pair the difference in Gibbs free energy for a Fe^{2+} ion in the field of the aluminum is changed to 0.14 eV–0.4 $k_B T$ for the two pairs. Then the concentration ratio gets very large, namely 295:1. Thus, <100> oriented pairs are practically absent. If, in an actual experiment, thermodynamic equilibrium is reached or not depends largely on the diffusion constant of the mobile partner and on the capture radius [50-54]. For iron-aluminum pairs the measurements of Chantre and Bois [39] and van Kooten et al. [43, 44] showed that these conditions are such that pairs are formed at room temperature after some hours.

The metastability of the pair was systematically studied by Chantre and Bois [39]. Their analysis is fully in agreement with our microscopic picture, which only adds some details to their explanation. From Fig. 3 and from the above discussion it follows that the Fermi level can be used to control the relative concentration of the stable and metastable structures. Chantre and Bois indeed observed that the charge state of the defect during sample cool-down to lower temperature controls such reversible transmutation behavior between the two pairs. In thermodynamic equilibrium the concentration of $<111>$ FeAl centers should be always higher than that of $<100>$ pairs. Van Kooten et al. [43, 44] noted that this was not obvious from the EPR intensities. The authors argued, however, that a reliable determination of defect concentrations is not possible from EPR.

4 Sulfur Pairs in Silicon

Covalent defect-defect interactions, pair formation and dissociation were studied by Weinert and Scheffler [3] for the chalcogens in silicon. Three different pair structures were considered: Both constituents occupying substitutional sites, both occupying interstitial sites, and the mixed substitutional-interstitial geometry. In all three cases the principal axis of the complex was the threefold-symmetric $<111>$ axis, and the separation was the nearest-neighbor distance (i.e. 2.35 Å).

It had been shown earlier [23] from total-energy calculations that isolated chalcogen point defects in silicon would kick out a Si atom and occupy the substitutional site. In the calculations for the pair the following reaction was therefore considered:

$$S_{Si}^{(m)} + S_{Si}^{(m)} + 2m\,e^-(E_F) \rightleftharpoons S_2^{(n)} + n\,e^-(E_F). \tag{8}$$

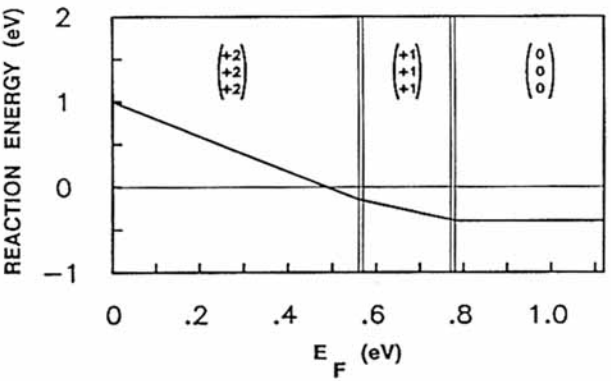

Fig. 4 Reaction energy for the formation of a nearest-neighbor substitutional sulfur impurity pair in silicon (see eq. (8)). The charges of the different species (m, m, n) in eq. (8) are given at the top of the figure. Zero of E_F is the top of the valence band. A negative reaction energy corresponds to an attractive interaction (exothermal process).

The reaction energy for the pair with *both partners at substitutional sites* is shown in Fig. 4 as a function of the Fermi level. For n-type material both the isolated impurities and the pair are neutral. Because the highest occupied defect-induced level is completely filled, the interaction between distant isolated sulfur impurities is practically absent. The attractive interaction starts only at smaller separations. The binding of the nearest-neighbor sulfur pair can be understood in terms of a free S_2 molecule which is placed into a di-vacancy [3]. It is then found that the interaction with the six di-vacancy dangling bonds weakens the S-S bond considerably compared to the gas-phase situation. The calculations show that these pairs are stable only if the Fermi level is in the upper half of the gap. For other Fermi-level positions such pairs are metastable. These results seem to be consistent with EPR and ENDOR measurements which identified sulfur and selenium pairs [55, 56]. The measurements could, however, only determine the symmetry and thus they could not distinguish between the pairs where both partners occupy substitutional sites and those where both occupy interstitial sites. This missing identification step is provided by the total-energy calculations [3] which show that the substitutional pairs should be dominant. Recently it became also possible to evaluate the hyper-fine and ligand hyper-fine fields of isolated point defects and defect pairs [57]. These calculations for the chalcogens show the importance of many-electron screening, and if compared to ENDOR data they prove that these measurements are due to substitutional-substitutional pairs.

A pair where *both partners occupy interstitial sites* is also bound by covalent interactions. The highest occupied wave functions of isolated sulfur interstitials are somehow similar (although more extended) to atomic sulfur 3p orbitals. Therefore the binding in the interstitial-interstitial pair is similar to that in the gas phase S_2 molecule. According to the reaction eq. (8) the relevant process requires that at first two substitutional sulfur impurity atoms go to interstitial positions and then form the pair. The cost to bring the two chalcogens from substitutional to interstitial sites is higher than the pair binding. As a consequence, interstitial-interstitial pairs should not exist in thermodynamic equilibrium in significant concentration.

The electronic structure of the *mixed geometry*, where one sulfur atom occupies a substitutional site and the other an interstitial site is summarized in Fig. 5. It results from the interaction between the a_1 level of the substitutional and the t_2 level of the interstitial [23]. Because the interstitial t_2 level is below the substitutional a_1 level, two electrons are transferred from the more substitutional-type to the more interstitial-type wave function. Thus, a chalcogen interstitial, which is usually believed to act as a donor, can also act as an acceptor and take two additional electrons. As a consequence, the binding of this system has some ionic character. Substitutional sulfur is a deep defect. Therefore a simple treatment in terms of purely Coulomb interactions, as that described in Section 3, is not appropriate and overestimates the binding energy by about a factor of 2. The energy gain due to pairing is smaller than the cost to bring a chalcogen atom from its normal substitutional site to the interstitial position. Only in n-type material are these two energies of similar size and we cannot exclude from the calculations that substitutional-interstitial pairs may exist.

If we compare the mechanisms which yield complex formation in semiconductors to those giving rise to molecule formation in normal chemistry we find that differences are mainly due to the following reasons.
(i) Chemical reactions in solids do not require charge conservation, because the Fermi level can take or give electrons, if needed. Therefore it can depend on the position of the Fermi energy if complex formation is exothermal or endothermal (see eqs. (7) and (8) and Figs. 3 and 4).
(ii) The wave function of the highest occupied level of an isolated defect is usually qualitatively different (spatial distribution and degree of degeneracy) from that of the free atom. To give an example: The electronic structure of a neutral substitutional sulfur impurity in Si has a closed shell and therefore the interaction between distant centers may be better compared to argon than to sulfur atoms.
(iii) Impurities in semiconductors can generally exist at different sites (substitutional and interstitial). This can allow two atoms of the same kind to form a partly ionic bond.
(iv) The impurity-impurity equilibrium distance is largely influenced by the host-crystal structure.
(v) Because of dielectric screening and the larger extent of wave functions the strengths of impurity-impurity ionic and covalent interactions are reduced compared to gas-phase ion-ion or atom-atom interactions. Furthermore, if a deep level shifts upon impurity-impurity interaction and approaches the conduction band (or the valence band) it simultaneously gets more delocalized because of increasing hybridization with the band states.

5 The Distant As_{Ga} -- As_i Pair in GaAs

The distant antisite-interstitial pair was recently proposed to be the defect behind the famous EL2 level [5]. In particular an ODENDOR analysis [6] seems to confirm and to specify this interpretation: Meyer et al. [6] concluded to see a 4.88

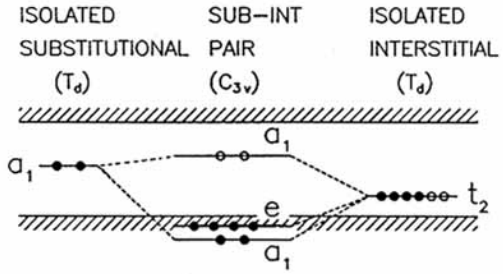

Fig. 5 Single-particle energies and interaction of a substitutional chalcogen (left) and an interstitial chalcogen (right) to form a nearest-neighbor pair (middle).

Å separated <111> oriented As_{Ga}^{+} -- As_i^{+} pair. Thus, the interstitial member is not paramagnetic and placed at the distant tetrahedral site. The authors identified this pair with the paramagnetic state of EL2 (see also Ref. 69).

The antisite-interstitial pair has been investigated theoretically by semi-empirical model calculations [58, 59, 7] and by parameter-free self-consistent studies [4]. We investigated in particular the positive charge state of the pair allowing the interstitial member to move along the [111] axis. In accordance with the results of Baraff et al. [58, 59] we find that the pair has a total-energy minimum when the interstitial is near the hexagonal site, 4 Å away from the antisite (see Fig. 6). The double positive pair is paramagnetic and its total energy minimum is close to the hexagonal interstitial site too. The interaction between the two partners at this quite long distance is small so that the energy levels of the pair are very close to those of the isolated components. We expect a binding at closer separation, similar to that discussed above in Section 4 for the substitutional - interstitial sulfur pair in Si (see Fig. 5). At the distances shown in Fig. 6 and in particular at the ENDOR derived 4.88 Å separation we find that the pair is practically unbound.

With the so far established understanding of pair formation and stabilities (see the above Sections and Ref. 3 and 4) we concluded that [8, 4] the 4.88 Å separated As_{Ga} -- As_i pair cannot exist in significant concentration. We thus question the

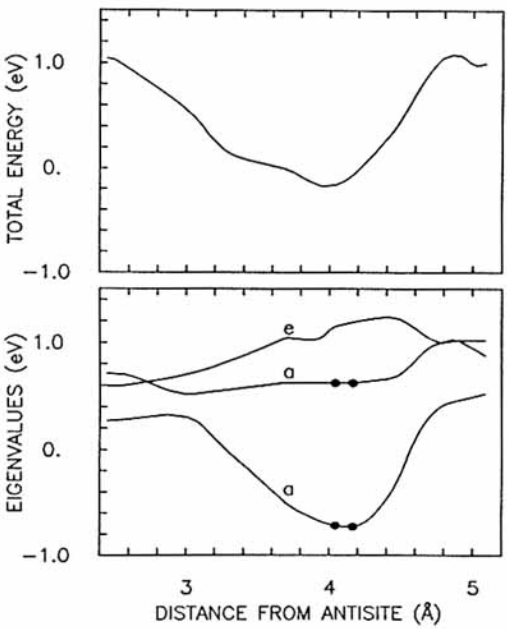

Fig. 6 Single-particle energies (with respect to the top of the valence band) and total energy as functions of the distance between an As-interstitial and an As-antisite. Lattice distortions are neglected. The pair's charge state is singly positive.

statement that this pair is identical with EL2 [5-7] as this would imply that the pair should be the dominant intrinsic deep donor. No mechanism has been suggested so far which brings the two constituents together and no mechanism is known which can be responsible for a strong binding.

It should be noted that a correlation of the paramagnetic EPR and ENDOR with the EL2 absorption does not imply the exact structural identity of these centers. In fact, similarly to what we discussed in Section 3, it is likely that a positively charged As_{Ga} antisite, which undoubtedly is at the core of the EPR [60, 61] and ENDOR [6] centers, may complex differently than the neutral charge state. Independent of the EL2 identification we are left with the question what is seen in paramagnetic resonance [61, 6]. We had speculated earlier [8, 4] that the ENDOR studies [6] are due to an As_{Ga} - X complex where X stands for an acceptor, not just an As_i.

6 The As_{Ga} Defect in GaAs

First studies of the structural properties of the As_{Ga} antisite in GaAs were performed by Bachelet and Scheffler [62, 63] five years ago. Figure 7 shows the electronic structure of the As_{Ga} antisite and how it can be understood in terms of the interaction of a Ga-vacancy and a free As-atom (see also Ref. 64, 8 and 9). In the neutral ground state the electronic configuration of the As_{Ga} antisite is $a_1^2t_2^0$. It was found [62, 63] that the defect behaves rather normal, when in this electronic configuration. For the excited electronic configuration $a_1^1t_2^1$ a structural instability was predicted, and it was pointed out [62, 63] that this instability resembles the properties of the metastable transition of EL2. It was also noted (see the discussion of Fig. 4 in Ref. 62) that the electronic structure of Fig. 7 and the instability of the electronic $a_1^1t_2^1$ configuration should be a general property of deep substitutional donors, not just of the As_{Ga} antisite. A careful total-energy calculation of low-symmetry defects was not possible at that time, but was recently achieved by Dabrowski and Scheffler [8, 9] and Chadi and Chang [10] and will be described below.

Figure 8 summarizes results [8, 9] for the As_{Ga}-antisite, when the central As-nucleus is displaced in the <111> direction. Displaced geometries thus correspond to a $V_{Ga}As_i$ defect pair with varied separation. The defect symmetry is now C_{3v} and the T_d t_2 state of Fig. 7 splits into two states now labeled as 2a and 1e. Figure 8 shows the single-particle energies of the 1a, 2a and 1e states (top) and three total-energy curves obtained for the three electronic configurations $1a^22a^0$ (labeled as F, which stands for fundamental), $1a^12a^1$ (labeled as E, which stands for excited) and $1a^02a^2$ (labeled as M, which stands for metastable). For the *E* total-energy curve, the Jahn-Teller effect *starts* the distortion, but already after a small displacement we find that the different wave functions (namely the 1a and the 2a single-particle states) will mix. It is most likely that the excited system falls back down to the F-curve, the ground-state total energy. Then the system ends again as a tetrahedral As_{Ga} antisite. However, Fig. 8 shows that for displaced geometries also another electronic configuration, namely $1a^02a^2$ becomes possible. Thus, once excited to the E-curve, the system has a certain probability to change to the M-curve. Then the arsenic defect atom will end

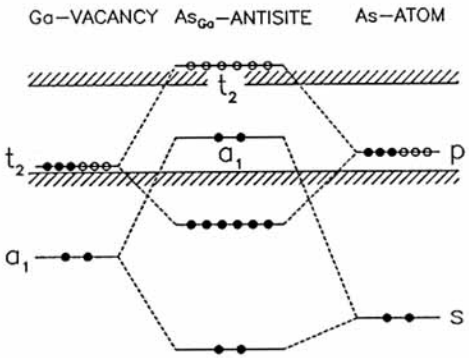

Fig. 7 Single-particle energies and interaction of a neutral Ga vacancy in GaAs (left) and a free As atom (right) which results in a As_{Ga}-antisite defect (middle).

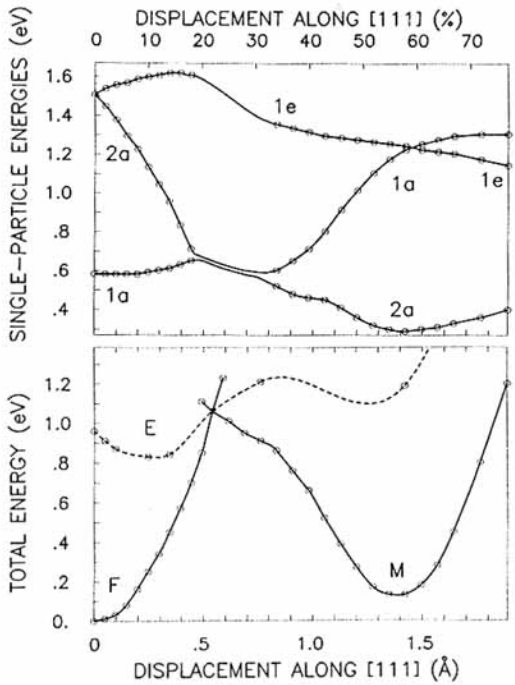

Fig. 8 Single-particle energies with respect to the valence-band edge (top), and total energies of the S=0 ground states (curves F and M) as functions of the position of the arsenic defect atom (bottom). Zero displacement refers to the tetrahedral As-antisite configuration. The total-energy curve labelled E is an electronic excited state with electronic configuration $1a^1 2a^1$. These calculations were performed with a basis set of E_{cut} = 8 Ryd, keeping all neighbors of the displaced arsenic atom at their perfect crystal positions (see Ref. 8).

considerably far away (about 1.4 Å) from the initial, central gallium site. We refer to this metastable atomic configuration as the gallium vacancy - arsenic interstitial pair. The As-interstitial is about 1 Å away from the T_d interstitial site. It is therefore chemically bound to only three arsenic atoms.

As the arsenic defect atom leaves the gallium site, its bond with one arsenic neighbor which is left behind is stretched and it almost breaks when the defect enters the barrier region. This is shown in Fig. 9. The barrier of the structural transition is reached when the arsenic atom passes through the (111) plane of three As neighbors. In the metastable configuration the arsenic defect atom (now

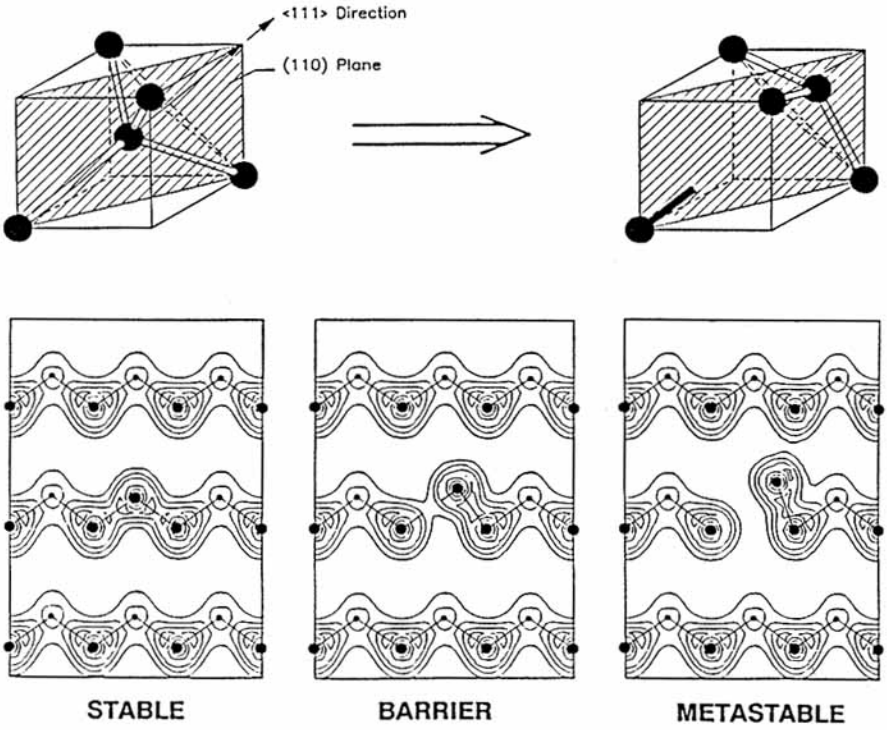

Fig. 9 The metastability of the As-antisite defect. Top: atomic structure, and bottom: The electron density in the (110) plane. Thick dots represent As atoms. Small dots represent Ga atoms. The left side shows the fundamental configuration, where the arsenic defect atom is bound to four nearest neighbors (only two are in the displayed (110) plane). The middle panel shows the barrier region. The right pictures correspond to the metastable situation (the $V_{Ga}As_i$ defect pair). Here the arsenic defect is bound to three arsenic neighbors (only one of them is in the (110) plane). The black "dangling bond" in the top right picture indicates the vacancy state (labeled 2a in Fig. 8), which is responsible for the barrier between the metastable and the fundamental configuration.

an interstitial) binds to these three atoms, similar to the bonding in crystalline grey arsenic. In the vacancy region there is one broken bond, which is filled with two electrons. The calculations of Fig. 8 give a barrier for the neutral ground state of 0.92 eV between the minimum of the metastable configuration (the $V_{Ga}As_i$ pair) and the fundamental configuration (the As_{Ga} antisite). We note that these calculations are performed with a smaller basis set and without allowing the nearest neighbor atoms to relax [8, 9]. If these two constraints are removed we obtain a barrier of about 0.4 eV.

According to the calculations [8-10] the four-fold coordinated As_{Ga} antisite and the metastable configuration with the threefold coordinated As_i have similar total energy. This result is indeed plausible for a group V element. The origin of the barrier between the two configurations is, however, not obvious. It may be understood by the fact that the covalent radius of an As atom is 1.2 Å. Therefore the As_i is too "thick" to pass easily through the (111) plane of the three As atoms. This argument is not complete and it cannot explain why for other charge states the barrier will in fact disappear (see Ref. 8, 9). The main reason for the barrier is the filled vacancy-like dangling bond schematically shown in Fig. 9. This state is antibonding with respect to the arsenic interstitial and its energy (the 2a level in Fig. 8) increases if the As-interstitial is moved from the metastable configuration towards the vacancy. The occupied vacancy dangling orbital therefore contributes to the repulsion between the constituents of the metastable pair. It is therefore obvious that the barrier will change if one electron is removed from this level. For a negatively charged $V_{Ga}As_i$ pair the barrier was found to be significantly reduced [8, 9] and an electron-induced regeneration $V_{Ga}As_i \rightarrow As_{Ga}$ should be likely. For a positively charged As_{Ga} the calculations [8, 9] predict that it should not exhibit metastable behavior.

The optically inducible transition to the metastable state competes with two other processes: At small displacements it is possible that the excited electron of the As_{Ga} goes to the conduction band ($1a^12a^1 \rightarrow 1a^12a^0 + e^-$) or that it goes back to the 1a state ($1a^12a^1 \rightarrow 1a^22a^0$). These two processes bring the arsenic defect atom back to the fundamental configuration, i.e. the tetrahedral As_{Ga} antisite. Because of these competitors the probability of the metastable transition will be small and it should be sensitive to local stress and other perturbations. It also depends sensitively on the conduction-band structure.

There is no clear, unambiguous experimental verification of the above described As_{Ga} metastability. However, the comparison of the theoretical results for the isolated arsenic antisite and the $As_{Ga} \rightleftharpoons V_{Ga}As_i$ metastability with a carefully compiled list of 17 experimentally established properties of the EL2 defect (see Ref. 8 and 9) revealed clear similarities. We therefore identified the mechanism of the EL2 metastability with that of the As_{Ga} antisite.

Although the described type of metastability should be present in principle for all substitutional deep donors, we cannot predict without a full calculation when it really will be observable: The height of the barrier and the energy difference of the metastable and stable configurations will be different for different systems. A very similar system as the As_{Ga}-antisite is a negatively charged Si_{Ga} defect in GaAs. The number of electrons is the same as in the neutral As_{Ga}. Indeed, according to calculations of Chadi and Chang [65], this defect shows a $Si_{Ga} \rightleftharpoons$

$V_{Ga}Si_i$ metastable behavior. The properties of the Si_{Ga} or $V_{Ga}Si_i$ defect seem to explain the properties of the famous DX centers in GaAlAs and in GaAs under pressure [65, 66]. The $As_{Ga} \rightleftharpoons V_{Ga}As_i$ type of process may be also important for defect diffusion. With this respect we note a similar phenomenon, found in group IV A elements (e.g. Ti and Zr), which are the transition-metal counterparts of the group IV B elements (e.g. C and Si). Group IV A elements show a transition from the bcc to the hexagonal structure, the so called ω-phase, in which the atoms are trigonally bonded. In the bcc phase a so called ω-embryo can be formed, which explains certain self-diffusion properties in these materials [67, 68].

7 Summary

In this paper we described defect-defect interactions and mechanisms of defect pair formation. Only point defects and defect pairs have been considered, because larger complexes have not been studied so far by first-principles calculations. Three classes of stable $\rightleftharpoons$ metastable configurations were discussed:
1) The bound pair versus the dissociated, isolated point defects,
2) bound pairs with different separations, and,
3) as a special case of No. 2, the $As_{Ga} \rightleftharpoons V_{Ga}As_i$ process.
The described calculations show that defect metastabilities can be much more important than this was often expected. It is also found that the pairing or dissociation is largely determined by the defect charge state and the crystal Fermi level. The concentrations of different atomic configurations are given by the law of mass action, assuming that the system is in thermodynamic equilibrium. However, often thermodynamic equilibrium will be not reached, because metastable configurations can have significant lifetimes. Then the concentrations can be affected by the history of the sample, and a higher temperature thermodynamic-equilibrium state can be frozen in. Furthermore, for compound semiconductors, as for example GaAs, the crystal environment can be crucial for the existing defect structures, because the concentration of intrinsic defects and their complexes depend on the Ga or As chemical potentials. These are controlled by the crystal environment, i.e. if Ga metal is present or if the sample is in an As_2 or As_4 gas atmosphere. Both, the electron chemical potential (i.e. the Fermi energy) and the atomic chemical potential, can affect defect-formation energies by more than one eV [12].

The $As_{Ga} \rightleftharpoons V_{Ga}As_i$ process, discussed in Section 6 was only recently found to be a likely reaction. In fact, this new type of metastability should be common to many substitutional defects. It resembles properties of the diamond $\rightleftharpoons$ graphite metastability.

We feel that with the theoretical results discussed in this paper we are just at the beginning of a microscopic understanding of defect complex formation in semiconductors and of defect metastabilities.

Acknowledgements

I gratefully acknowledge the collaboration with F. Beeler, C.M. Weinert and J. Dabrowski on some of the above discussed subjects. I thank H. Overhof, U. Scherz and R. Gillert for reading the manuscript.

References and Footnotes

[1] *M. Scheffler*, Festkörperprobleme XXII ed. by P. Grosse, (Vieweg, Braunschweig, 1982), p. 115.

[2] *M. Scheffler*, Physica **146** B, 176 (1987).

[3] *C. M. Weinert*, and *M. Scheffler*, Phys. Rev. Letters **58**, 1456 (1987).

[4] *J. Dabrowski*, and *M. Scheffler*, Materials Science Forum **38-41**, 51 (1989).

[5] *H. J. von Bardeleben*, *D. Stievenard*, *D. Deresmes*, *A. Huber*, and *J. C. Bourgoin*, Phys. Rev. B **34**, 7192 (1986).

[6] *B. K. Meyer*, *D. M. Hofmann*, *J. R. Niklas*, and *J.-M. Spaeth*, Phys. Rev. B **36**, 1332 (1987).

[7] *C. Delerue*, *M. Lannoo*, and *D. Stievenard*, Phys. Rev. Letters **59**, 2875 (1987).

[8] *J. Dabrowski*, and *M. Scheffler*, Phys. Rev. Letters **60**, 2183 (1988).

[9] *J. Dabrowski*, and *M. Scheffler*, submitted to Phys. Rev. B (1989).

[10] *J.D. Chadi*, and *K.J. Chang*, Phys. Rev. Letters **60**, 2187 (1988)

[11] In the adiabatic approximation the wave functions are written as $\psi(\{r_i\},\{R_j\}) = \phi(\{r_i\},\{R_j\})\ X(\{R_j\})$, with $\phi(\{r_i\},\{R_j\})$ beeing the ground state of the electronic Hamiltonian of a fixed nuclear structure. If electron-phonon interactions are neglected, the Hamiltonian of eq. (1) has to be solved only for $X(\{R_i\})$.

[12] *M. Scheffler*, and *J. Dabrowski*, Phil. Mag. A **58**, 107 (1988).

[13] *S. Biernacki*, and *M. Scheffler*, submitted to Phys. Rev. (1989).

[14] *P. Hohenberg*, and *W. Kohn*, Phys. Rev. **136**, B 864 (1964).

[15] *W. Kohn*, and *L.J. Sham*, Phys. Rev. **140**, A 1133 (1965).

[16] The Inhomogeneous Electron Gas, ed. by *N.H. March* and *S. Lundqvist*, (Plenum, New York, 1984).

[17] *J.R. Chelikowsky* and *S.G. Louie*, Phys. Rev. B **29**, 3470 (1984).

[18] *J. Harris*, Phys. Rev. B **31**, 1770 (1985).

[19] *H.M. Polatoglou*, and *M. Methfessel*, Phys. Rev. B **37**, 10403 (1989).

[20] *M. Scheffler*, *J.P. Vigneron*, and *G. B. Bachelet*, Phys. Rev. Letters **49**, 1765 (1982); Phys. Rev. B **31**, 6541 (1985).

[21] *G.A. Baraff*, and *M. Schlüter*, Phys. Rev. B **30**, 1853 (1985).

[22] *O. Gunnarsson*, *O. Jepsen*, and *O.K. Andersen*, Phys. Rev. B **27**, 7144 (1983).

[23] *F. Beeler*, *M. Scheffler*, *O. Jepsen*, and *O.K. Andersen*, Phys. Rev. Letters **54**, 2525 (1985).

[24] *F. Beeler*, *O.K. Andersen*, and *M. Scheffler*, Phys. Rev. Letters **55**, 1498 (1985); Phys. Rev. B, in print.

[25] *J. Bernholc*, *N.O. Lipari*, and *S.T. Pantelides*, Phys. Rev. Letters **41**, 895 (1978); Phys. Rev. B **21**, 3545 (1980).

[26] *R. Car, P.J. Kelly, A. Oshiyama,* and *S.T. Pantelides,* Phys. Rev. Letters **52,** 1814 (1984); **54,** 360 (1985).
[27] *Y. Bar-Yam,* and *J.D. Joannopoulos,* Phys. Rev. Letters **52,** 1129 (1984); Phys. Rev. B **30,** 1844 (1984).
[28] *R. Car,* and *M. Parrinello,* Phys Rev. Letters **55,** 2471 (1985).
[29] *G.L. Chiarotti, F. Buda, R. Car,* and *M. Parrinello,* to be published.
[30] *Ch.G. Van de Walle, Y. Bar-Yam,* and *S.T. Pantelides,* Phys. Rev. Letters **60,** 2761 (1988).
[31] *K.C. Pandey,* Phys. Rev. Letters **57,** 2287 (1986).
[32] *S. Froyen,* and *A. Zunger,* Phys. Rev. B **34,** 7451 (1986).
[33] *G.P. Kerker,* J. Phys. C **13,** L189 (1980).
[34] *D.R. Hamann, M. Schlüter,* and *C. Chiang,* Phys. Rev. Letters **43,** 1494 (1979).
[35] *G.B. Bachelet, D.R. Hamann,* and *M. Schlüter,* Phys. Rev. B **26,** 4199 (1982).
[36] *U. Scherz, D. Weider,* and *M. Scheffler,* to be published.
[37] *G.W. Ludwig* and *H.H. Woodbury,* in Solid State Physics **13,** ed. by *F. Seitz* and *D. Turnbull* (Academic, New York, 1962), p. 223.
[38] *P.J. Dean,* Prog. in Sol. St. Chem. **8,** 1 (1973).
[39] *A. Chantre,* and *D. Bois,* Phys. Rev. B **31,** 7979 (1985).
[40] *H. Feichtinger, J. Oswald, R. Czaputa, P. Vogl,* and *K. Wünstel,* Proceedings of the 13th International Conference on Defects in Semiconductors, ed. by *L.C. Kimerling* and *J.M. Parsey,* Jr. (The metallurgical Society of AIME, 1984), p. 855.
[41] *K. Wünstel,* and *P. Wagner,* Appl. Phys. A **27,** 207 (1982).
[42] *K. Graff* and *H. Pieper,* J. Electrochem. Soc. **128,** 669 (1981).
[43] *J.J. van Kooten, G.A. Weller,* and *C.A.J. Ammerlaan,* Phys. Rev. B **30,** 4564 (1984).
[44] *J.J. van Kooten,* Ph.D. thesis, Amsterdam (1987).
[45] *E.R. Weber,* Appl. Phys. A **30,** 1 (1980).
[46] *J.P. Vigneron* et al., unpublished results.
[47] *H. Katayama-Yoshida,* and *A. Zunger,* Phys. Rev. Letters **53,** 1256 (1984); Phys. Rev. B **31,** 7877 (1985).
[48] *L.V.C. Assali,* and *J.R. Leite,* Phys. Rev. B **36,** 1296 (1987).
[49] *O. Madelung,* Festkörpertheorie Vol. III (Springer, Berlin 1972).
[50] *F.A. Kröger,* The Chemistry of Imperfect Crystals (North-Holland, Amsterdam 1974).
[51] *H. Reiss, C.S. Fuller,* and *F.J. Morin,* Bell Syst. Tech. J. **35,** 535 (1956).
[52] *W.H. Shepherd,* and *J.A. Turner,* J. Phys. Chem. Solids **23,** 1697 (1962).
[53] *L.C. Kimerling, J.L. Benton,* and *J.J. Rubin,* in Defects and Radiation Effects in Semiconductors (Inst. Phys. Conf. Ser. 59, 1981), p. 217.
[54] *L.C. Kimerling,* and *J.L. Benton,* Physica **116B,** 297 (1983).
[55] *R. Wörner,* and *O.F. Schirmer,* Solid State Commun. **51,** 665 (1984).
[56] *S. Greulich-Weber, J.R. Niklas, J.-M. Spaeth,* J. Phys. Cond. Matter **1,** 35 (1989).
[57] *H. Overhof, M. Scheffler,* and *C.M. Weinert,* Materials Science Forum **38-41,** 293 (1989); and to be published (1989).
[58] *G.A. Baraff,* and *M. Lannoo,* Revue Phys. Appl. **23,** 817 (1988).
[59] *G.A. Baraff, M. Lannoo,* and *M. Schlüter,* Phys. Rev. B **38,** 6003 (1988).

[60] *E.R. Weber, H. Ennen, U. Kaufmann, J. Windscheif, J. Schneider,* and *T. Wosinski,* J. Appl. Phys. **53,** 6140 (1982).

[61] *M. Wattenbach, J. Kröger, C. Kisielowski-Kemmerich,* and *H. Alexander,* Materials Science Forum **38-41,** 73 (1989).

[62] *M. Scheffler, F. Beeler, O. Jepsen, O. Gunnarsson, O.K. Anderesen,* and *G.B. Bachelet,* Proceedings of the 13th International Conference on Defects in Semiconductors, ed. by *L.C. Kimerling and J.M. Parsey,* Jr. (The metallurgical Society of AIME, 1984), p. 45.

[63] *G.B. Bachelet,* and *M. Scheffler,* Proceedings of the 17th Int. Conf. on the Physics of Semiconductors, ed. by J.D. Chadi and W.A. Harrison (Springer, New York, 1985), p. 755.

[64] *G.B. Bachelet, M. Schlüter,* and *G.A. Baraff,* Phys. Rev. B **27,** 2545 (1983).

[65] *J.D. Chadi,* and *K.J. Chang,* Phys. Rev. Letters **61,** 873 (1988).

[66] *P. Mooney,* this volume.

[67] *J.M. Sanchez,* and *D. de Fontaine,* Phys. Rev. Letters **35,** 227 (1975).

[68] *U. Köhler,* and *Ch. Herzig,* Phil. Mag. A **58,** 769 (1988).

[69] At this coference *B. Meyer* told that the 4.88 Å separation reported in Ref. 6 may be not very accurate. He emphasized, however, that the C_{3v} symmetry of the complex and the existance of an As-interstitial close to the As_{Ga} can be safely identified from the ENDOR data.