Philosophical Magazine A
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713396797

Parameter-free calculations of total energies, interatomic forces and vibrational entropies of defects in semiconductors

Matthias Scheffler [a]; Jaroslaw Dabrowski [a]
[a] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, F.R. Germany

Online Publication Date: 01 July 1988

To cite this Article Scheffler, Matthias and Dabrowski, Jaroslaw(1988)'Parameter-free calculations of total energies, interatomic forces and vibrational entropies of defects in semiconductors',Philosophical Magazine A,58:1,107 — 121
To link to this Article: DOI: 10.1080/01418618808205178
URL: http://dx.doi.org/10.1080/01418618808205178

PHILOSOPHICAL MAGAZINE A, 1988, VOL 58, NO. 1, 107–121

Parameter-free calculations of total energies, interatomic forces and vibrational entropies of defects in semiconductors

By MATTHIAS SCHEFFLER and JAROSLAW DABROWSKI

Fritz-Haber-Institut der Max-Planck-Gesellschaft,
Faradayweg 4–6, D-1000 Berlin 33, F.R. Germany

ABSTRACT

We discuss calculations from first-principles (using the local-density approximation for exchange and correlation) of defect total energies, vibrational modes, internal energies and entropies. Results are presented for the defect-induced distortion field of an arsenic impurity in silicon and for the vibrational entropy of a silicon vacancy. We also discuss the important role of electron and atom chemical potentials, presenting results for the Ga vacancy in the GaAs bulk and at the (111) surface.

§ 1. INTRODUCTION

The requirement of semiconductor technology for accurate control of defect densities and profiles, atomic structure at interfaces and suchlike has stimulated extensive experimental and theoretical research. An important step forward in the theory of defects was achieved by the development of first-principles methods to calculate the electronic and magnetic structure, and the total energy (Scheffler, Vigneron and Bachelet 1982, 1985, Baraff, Schlüter and Allan 1983, Baraff and Schlüter 1984, Car, Kelly, Oshiyama and Pantelides 1984, 1985, Beeler, Scheffler, Jepsen and Gunnarsson 1985, Bar-Yam and Joannopoulos 1984, Froyen and Zunger 1986, Pandey 1986). These calculations allowed the determination of defect-induced lattice distortions, force constants, reaction energies and vibrational modes (Scheffler 1987, Scheffler and Scherz 1986). In many cases knowledge of the total energy is not sufficient; it may be important to consider the appropriate thermodynamic function for the experimental conditions (the Gibbs free energy, for example). If complete thermodynamic equilibrium is possible (at sufficiently high temperature) the requirement of finding the minimum of the appropriate thermodynamic function refers to the entire system (namely, bulk, surfaces and outer region). Difficulties may arise because in many (maybe most) practical situations thermodynamic equilibrium is not attained. For example, at or below room temperature chemical reactions in the bulk may not be in equilibrium with the surface. Then it is often assumed that a 'partial equilibrium' exists and thermodynamics can be applied only to certain defect reactions (see § 5 below).

For optical experiments the sample cannot adjust its volume in the time of the optical transition. Therefore these experiments should be considered as performed at constant volume and entropy. The proper thermodynamic function to use is the internal energy. On the other hand, for defect structure, atomic transport and reactions the time scale is usually slow so that electrons *and* nuclei can relax. The usual experimental conditions are those of constant temperature and pressure. Then the

structure of the crystal, its surfaces and its defects are determined by the minimum of the Gibbs free energy,

$$G(N_A, N_B, \ldots, T, P) = U - TS + PV, \tag{1}$$

where N_A, $N_B, \ldots, T, P, U, S$ and V are the particle numbers of particles of types A, B, ..., temperature, pressure, internal energy, entropy and crystal volume respectively. Particle numbers, temperature and pressure are determined by the experimental situation. U, S and V follow from the atomic configuration, which must be varied to find the minimum of G among all possibilities.

In § 2 we summarize the basic equations which must be evaluated to obtain the Gibbs free energy. In § 3 we describe the methods used to calculate those quantities from first-principles which enter the evaluation of G. In § 4 we present results for the distortion field induced by a defect, using as an example the As impurity in silicon. A short discussion of calculations of the vacancy-induced change in the vibrational entropy is also given. In § 5 we discuss the role of atom and electron chemical potentials in defect formation, using as an example the Ga vacancy in GaAs. Finally, in § 6 we summarize the results and give a critical discussion of the capabilities and limitations of our methods.

§ 2. Calculation of the Gibbs free energy: basic equations

In the following we assume that the system has only one stable configuration and that metastable structures have a significantly higher Gibbs free energy. In order to calculate G one has to evaluate the internal energy U and the entropy S. The internal energy can be written as

$$U = U^{\text{static}} + U^{\text{vib.}} \tag{2}$$

where U^{static} is often referred to as either the static or the structural or as the *total* energy (see § 3). In the harmonic approximation the vibrational contribution to the internal energy is

$$U^{\text{vib.}} = \sum_{i=1}^{3N} \left\{ \frac{\hbar\omega_i}{\exp(\hbar\omega_i/k_B T) - 1} + \frac{1}{2}\hbar\omega_i \right\}, \tag{3}$$

which for high temperatures approaches the form $3Nk_B T$. The energies $\hbar\omega_i$ are for the normal vibrational modes. In § 4 we describe a tractable method for their calculation from first principles.

The entropy can be written as

$$S = S^{\text{config.}} + S^{\text{vib.}} + S^{\text{e-h}}, \tag{4}$$

where the three terms are respectively the configurational, vibrational and electron–hole pair contributions. For covalent semiconductors at or below room temperatures $S^{\text{e-h}}$ is negligible. The configurational entropy of a defect is given by the number of possible configurations in which the defect can exist (Madelung 1972, Kröger 1974). The vibrational entropy is

$$S^{\text{vib.}} = k_B \sum_{i=1}^{3N} \left\{ \frac{\hbar\omega_i}{k_B T} \left[\exp\left(\frac{\hbar\omega_i}{k_B T}\right) - 1 \right]^{-1} - \ln\left[1 - \exp\left(\frac{-\hbar\omega_i}{k_B T}\right) \right] \right\}, \tag{5}$$

which for high temperatures approaches the form

$$k_B \sum_{i=1}^{3N} \left\{ 1 - \ln\left(\frac{\hbar\omega_i}{k_B T}\right) \right\}.$$

The vibrational entropy (eqn. (5)) follows from eqn. (3) by solving it for $1/T$ (which equals $\partial S^{\text{vib.}}/\partial U$) and then integrating this expression over U. It is therefore often important to take into account that the entropy depends on the variation of the internal energy with volume. For example, Harding (1985) has shown for vacancy formation that the defect-induced entropy change $\Delta S_V^{\text{vib.}}$ may even have a different sign than $\Delta S_P^{\text{vib.}}$. Here $\Delta S_V^{\text{vib.}}$ is the entropy change calculated under the assumption that defect formation happens at constant volume. $\Delta S_P^{\text{vib.}}$ would be the result if the defect is formed under the condition of constant pressure.

§ 3. First-principles methods for the evaluation of forces and total energies

From eqns. (2)–(5) it follows that the Gibbs free energy can be evaluated if the total energy U^{static} and the normal vibrational-mode energies $\hbar\omega_i$ are known. These quantities can be calculated in a parameter-free way by using the density-functional theory (DFT) (March and Lundqvist 1984). This gives the self-consistent electron density and from this the total energy of the many-electron system is obtained. Forces, force constants and vibrational frequencies follow from calculations of the total-energy gradient with respect to nuclear positions. These forces can be calculated directly (see Scheffler *et al.* (1982, 1985)), or from a numerical derivative of total energies calculated for different atomic configurations.

The main approximations of recent density-functional theory calculations are the Born–Oppenheimer approximation, the frozen-core approximation and the local-density approximation for the exchange-correlation functional (see for example, March and Lundqvist (1984)). We currently use two different methods in order to solve the Kohn–Sham equation of DFT for the problem of a point defect in a bulk crystal. One of the methods is the self-consistent Green-function method (Bernholc, Lipari and Pantelides (1978), Baraff and Schlüter (1978), for a review see Scheffler (1982)). This method has many of the advantages of cluster calculations without their inaccuracies. The Green-function method may be viewed as a special type of cluster calculation, where the cluster is embedded correctly in an infinite crystal. This embedding is achieved by solving the Dyson equation. Therefore (in contrast to cluster calculations) the Green-function method does not place a restriction on the extent of wavefunctions.

An alternative to the self-consistent Green-function method is the super-cell approach (Louie, Schlüter, Chelikowsky and Cohen 1976, Bar-Yam and Joannopoulos 1984, Froyen and Zunger 1986, Pandey 1986). Here a cluster, usually of between 8 and 54 atoms, is taken as the unit cell of an artificial crystal. With this method it is possible to calculate the total energy (U^{static}) for a perfect crystal, for a defect and for a finite number of layers plus vacuum region, in order to simulate a surface. Such periodically repeated clusters, each containing one defect, introduce a spurious 'defect band structure'. Therefore, single-particle energies must be used with caution. However, if the size of the super-cell is not too small, the total energies of a super-cell and from the Green-function method agree.

The results presented below were obtained by using the super-cell approach with a sixteen-atom unit cell. Test calculations have also been performed with a 54-atom cell. The ionic (frozen core) potentials were replaced by the first-principles pseudo-

potentials of Bachelet, Hamann and Schlüter (1982). We used a plane-wave basis with an energy cut-off of 16 Ry. The **k**-integration is performed by using two special points in the irreducible part of the super-cell Brillouin zone. For the exchange-correlation functional we used the results of Ceperley and Alder (1980). A comparison with our self-consistent Green-function calculations is given for those cases where such calculations had been performed earlier.

§4. Defect-induced lattice distortions and vibrational frequencies

4.1. *Method*

It is possible to evaluate the long-range distortion field around a defect, as well as the dynamical matrix and the defect's vibrations *directly* from self-consistent total energies. However, we believe that this is unnecessarily rigorous and instead suggest taking advantage of knowledge from elasticity theory. We therefore apply an analytic formula which describes *changes* of the total energy *for small deviations* from the equilibrium geometry and combine this analytic expression with numerical results from first-principles calculations. The valence-force approach (Martin 1970, Torres and Stoneham 1985) appears to be appropriate. We use Keating's formulation (Keating 1966). The change of the total energy of an N-atom cluster is then

$$E(\{\mathbf{R}_i\}) = \sum_{i=1}^{3N} \sum_{\substack{j \\ \text{bonds at} \\ \text{site } i}} \left\{ \frac{3}{16(\mathbf{R}_i^0 - \mathbf{R}_j^0)^2} \alpha_j^i [(\mathbf{R}_i^0 - \mathbf{R}_j^0)^2 - (\mathbf{R}_i - \mathbf{R}_j)^2]^2 \right.$$

$$\left. + \sum_{\substack{k \\ \text{adjacent} \\ \text{bonds}}} \frac{3}{8|\mathbf{R}_i^0 - \mathbf{R}_j^0| \cdot |\mathbf{R}_i^0 - \mathbf{R}_k^0|} \beta_{jk}^i [(\mathbf{R}_i^0 - \mathbf{R}_j^0)(\mathbf{R}_i^0 - \mathbf{R}_k^0) - (\mathbf{R}_i - \mathbf{R}_j)(\mathbf{R}_i - \mathbf{R}_k)]^2 \right\}. \quad (6)$$

Here $\mathbf{R}_i^0$ are the positions of the atoms at (or close to) the equilibrium geometry, while $\mathbf{R}_i$ are the positions of the displaced atoms and α_j^i and β_{jk}^i are so-called bond-stretching and bond-bending parameters. We note that $\alpha_j^i = \alpha_i^j$ and $\beta_{jk}^i = \beta_{kj}^i$. Equation (6) is quite simple, thus a relatively large cluster can be used in order to evaluate the long-range lattice distortions and lattice vibrations. For the latter, long wavelength vibrations are suppressed when the wavelength becomes comparable to the cluster diameter. In the calculations discussed below we use a cluster of 524 atoms. For defects in silicon, vibrations with $\hbar\omega_i \lesssim 20$ meV are suppressed (see also Scherz and Scheffler (1988)). We believe that this suppression of long wavelength vibrations is not a severe approximation as we are interested in defect-induced *changes* of the entropy. (The entropy is given by the sum over all vibrations and defect vibrations are localized.) Calculations for the Si vacancy support this point of view (Biernacki, Scherz and Scheffler 1988).

The Keating approach has been used in semi-empirical calculations in order to evaluate perfect crystal vibrations when the parameters α and β were determined by fitting eqn. (6) to experimental results (see for example Martin (1970)). In contrast to this semi-empirical approach we calculate the parameters from first-principles using DFT as described in § 2 (see also Scheffler and Scherz (1986, 1988)). We note that a defect in a covalent semiconductor changes the electron density of the perfect crystal significantly only in a small region of the crystal, for instance in a sphere with a radius of 1–2 interatomic distances (see fig. 2). Therefore, the parameters $\mathbf{R}_i^0$, α and β which belong to more distant atoms are taken to have their values as in the perfect crystal. First we

determine the perfect-crystal parameters. From density-functional theory we calculate for the perfect crystal the total energy as a function of the lattice parameter. This numerical 'equation of state' gives the equilibrium geometry, i.e. the positions $\mathbf{R}_i^0$, and the bulk modulus B^0. The latter is related to α and β by

$$B^0 = (1/a^0)(\alpha + \tfrac{1}{3}\beta). \tag{7}$$

We have dropped the indices of α and β because in a perfect Si crystal there is only one type of atom, a^0 is the crystal lattice constant. As we are interested in defect-induced changes of geometry and vibrations we consider a situation where the nearest neighbours of one atom are displaced in a breathing-mode fashion but all other atoms are kept at their perfect-crystal positions. The force constant of this distortion is related to α and β by

$$k_b = 16\alpha + 42\beta. \tag{8}$$

Then with eqns. (7) and (8) and the theoretical values for B^0 (0·92 Mbar) and k_b (61 eV Å^{-2}) we obtain the values for α and β for the perfect crystal.

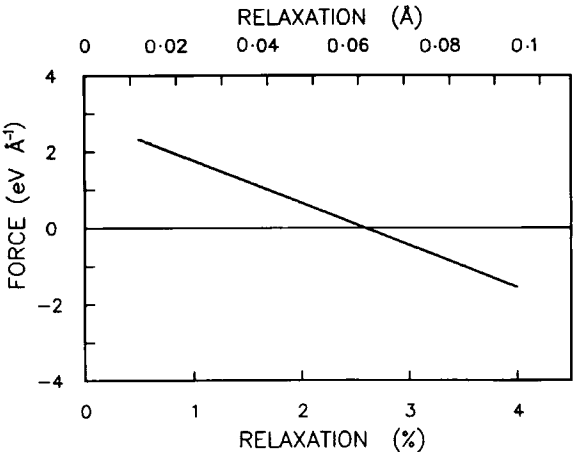
As already noted, for a defect system the parameters close to the defect must be changed. These changes are determined in the following way. We perform first-principles force calculations for the defect system for several geometries under the constraint that only the defect and its nearest neighbours move, keeping all other atoms at their perfect crystal positions. Then we perform the same calculations using eqn. (6) together with the perfect crystal parameters and adjust the parameters $\mathbf{R}_i^0$, α_j^i and β_{jk}^i close to the defect so that the first-principle and analytic total energies agree. Thus, eqn. (6) gives essentially the same results as DFT under the mentioned constraint. For tetrahedral substitutional defects we expect that only five parameters will differ significantly from the perfect crystal values: namely (i) the distance between the defect and its nearest neighbours ($d = |\mathbf{R}_1^0 - \mathbf{R}_i^0|$, with $i = 2$–5), (ii) the defect to nearest-neighbour bond-stretching force constants (α_i^1, with $i = 2$–5), (iii) the defect to nearest-neighbour bond-bending force constant (β_{jk}^1, with $j, k = 2$–5), (iv) the nearest-neighbour to next-nearest-neighbour bond-bending force constant ($\beta_{jk}^i, j = 1, i = 2$–$5, k = 6$–17) and (v) the defect nearest-neighbour to next-nearest-neighbour bond-stretching force constants (α_j^i, with $i = 2$–$5, j = 6$–17). Only five self-consistent defect calculations of total energies and forces are required to determine these numbers.

4.2. *Lattice distortions at* T = 0

We discuss here the results for a substitutional As impurity in Si. The minimum of $E(\{\mathbf{R}_i\})$ is determined using a cluster of 524 atoms. In these calculations we considered only (as a first approximation) defect-induced changes of the parameters $\mathbf{R}_j^0$, α_j^1 and β_{jk}^1, with $j, k = 2$–5. All other parameters were taken the same as in the perfect Si crystal.

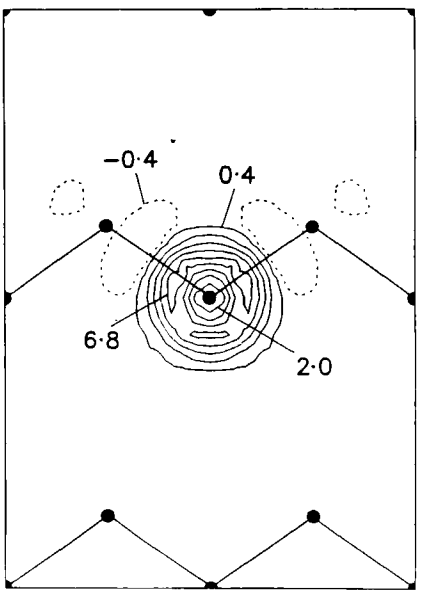
Figure 1 shows the calculated forces acting on the nearest neighbour Si nuclei for the single donor As in Si. The results are practically identical to those published earlier using the self-consistent Green-function method (Becker and Scheffler 1984, Scheffler 1987). We note that substitutional As is a shallow donor, namely the donor electron is bound in an effective-mass wavefunction derived from the conduction band X-point and thus this electron is distributed over a large volume. As a consequence, the charge density at the impurity is practically the same whether the shallow level is occupied or not, and the atomic relaxation close to the impurity does not significantly depend on the defect charge state. The calculations shown in fig. 1 were therefore performed for the system Si: As$^+$. What is the mechanism which drives the neighbours of an As$^+$ defect

Fig. 1



Force acting on the four nearest neighbours of an As impurity in Si, as calculated from eqn. (6). Zero relaxation corresponds to the geometry of the perfect Si crystal. The calculations show that the Si atoms undergo a breathing-mode relaxation, moving away from the impurity.

Fig. 2



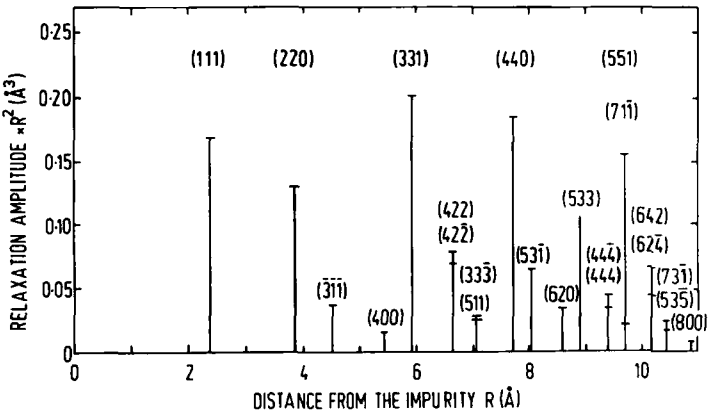
The change in electron density induced by an arsenic impurity in silicon. Full lines describe an increase in density, dashed lines a decrease. The figure shows contours along the (110) plane.

away from the impurity? A heterovalent impurity gives rise to two modifications which act on the neighbouring nuclei and which make them move away from their perfect-crystal positions. First, there is a change in the ion core (the nucleus and core electrons) at the impurity site. Second, there is a change in the electronic charge density, because the crystal valence electrons rearrange themselves in order to screen the modified ion core, that is, the additional nuclear charge. Figure 2 shows this screening-charge density for $Si:As^+$ in the (110) plane of the Si crystal. Close to the impurity there is a strong increase in the electron density. In fact, there are slightly more electrons close to the impurity than actually needed for the screening. This overscreening is corrected farther away in an oscillating way. These calculations show the following.

(1) The screening in Si is very efficient, and happens within a very short distance.
(2) Although we introduced an As^+ defect, the centre looks locally more like a neutral atom.
(3) The distortion direction follows from the electrostatic interaction between this anisotropic screening charge density together with the change in nuclear charge and the ionic potential of the impurity neighbours. In the particular example of $Si:As^+$ the neighbouring Si nuclei experience repulsion and move away from the centre.

As a result of the movement of the four neighbours of the impurity, the more distant nuclei also relax. For an outwards distortion of the nearest neighbours by 0·03 Å, the distortion pattern of fig. 3 is obtained. Note that the distortion amplitudes in fig. 3 are multiplied by a factor R^2, where R is the distance to the impurity. Figure 3 reveals that the amplitude of the distortion decreases rather rapidly with increasing distance from the centre and that the distortion pattern is quite anisotropic: the distortion propagates in particular along the zigzag bonding chains of the (110) planes (see also fig. 2). With increasing distance from the impurity, which is at the centre (denoted by (000)), the Si nuclei along these bonding chains are characterized by the prototype vectors (111) (nearest neighbours), (220) (second-nearest neighbours), (331) (fifth-nearest neighbours), (440) (tenth-nearest neighbours), (551), (seventeenth-nearest neighbours),

Fig. 3



Relaxation amplitudes (multiplied by R^2) induced by an arsenic impurity in silicon. The impurity sits at $R=0$. All atoms move away from the centre.

etc. Their relaxation amplitudes are rather large compared to other nuclei closer to the impurity. Corresponding results concerning this anisotropy of the distortion pattern and the electron charge density had been obtained and discussed also by Kane (1985) and Fleszar and Resta (1986).

The results of fig. 3 for the relaxation of the first- and second-nearest neighbours at Si : As^+ agree well with a recent analysis of extended absorption fine structure (EXAFS) measurements by Erbil, Weber, Cargill and Boehme (1986). These authors find that the nearest neighbours of an As^+ impurity move away from the defect by $0{\cdot}06 \pm 0{\cdot}02$ Å. Our theory gives a value between 0·02 and 0·04 Å. Concerning the second-nearest neighbours the theoretical relaxation amplitude is between $+0{\cdot}005$ and $+0{\cdot}012$ Å. Here the EXAFS analysis and the theoretical results agree perfectly.

4.3. *Defect-induced vibrations and vibrational entropy*

Vibrational frequencies are determined by

$$\det(\mathbf{D} - \mathbf{M}\omega_i^2) = 0. \tag{9}$$

Here **M** is the matrix of atomic masses and **D** is the dynamical matrix,

$$\mathbf{D} = \left.\frac{\partial^2 U^{\text{static}}}{\partial X_\mu \partial X_\nu}\right|_{\{X_i^0\}}, \tag{10}$$

where X_μ denote the $3N$ coordinates of the N-atom system and $\{X_i^0\}$ is the equilibrium geometry.

The dynamical-matrix elements may be obtained from the total energies of the equilibrium structure and distorted geometries. This 'direct approach' requires several hundred self-consistent calculations. Instead, Scheffler and Scherz (1986) have adopted a different approach using eqn. (6) together with first-principles calculations to determine firstly the lattice distortions and then the dynamical matrix. Some results for vibrations of defects in GaAs had already been published (Scheffler and Scherz 1986). Here we discuss our first attempt to evaluate the vibrational entropy for the doubly positive charge state of the Si vacancy using eqn. (5). At room temperature we obtain a value for $\Delta S_V^{\text{vib.}}$ of about $3\,k_B$ (Biernacki, Scherz and Scheffler 1988). This implies that the entropy part of the defect-induced change of the Gibbs free energy, $T\,\Delta S_V^{\text{vib.}}$, is indeed negligibly small ($\approx 0{\cdot}06$ eV at room temperature) if comparedd to the total-energy part. The latter (i.e. the vacancy formation energy) is about 4 eV. We also mention that Bachelet *et al.* (1987) have recently calculated the vacancy entropy using an eight-atom-super-cell and the direct approach. Their results, as well as an empirical valence-force model calculation of Lannoo and Allan (1982, 1986), agree well with ours.

§5. Formation of a Ga vacancy in bulk GaAs and at the (111) surface

We give here a short discussion of calculations for the formation of defects in crystal bulk and at the surface. In particular we will stress the important role of atom and electron chemical potentials in defect reactions. We consider small defect concentrations so that defect–defect interactions can be neglected and the defect distribution can be assumed to be at random. Furthermore, as a result of §4.3, we neglect contributions from the vibrational entropy.

Defect reactions are conveniently described in terms of chemical reactions between so-called structure elements (Kröger 1974). We consider the perfect, but finite, crystal as the reference state and label it in the reaction equation by zero. With the exception of

the immediate surface region the crystal is considered as built from identical layers parallel to the surface of interest. This layer is built from crystal unit cells and is of minimal thickness. The layers in the immediate surface region may differ from other layers only due to adsorption or reconstruction, if present. Defects in different charge states or in a different chemical bonding state (in the bulk or at the surface), the electrons at the Fermi level, as well as the different atoms in their various charge states are treated as distinct chemical species: A, B, C.... Their chemical potentials are:

$$\mu(\mathrm{A}) \equiv \partial G / \partial N_{\mathrm{A}}|_{P,T,N_{\mathrm{B}},N_{\mathrm{C}},\ldots}. \tag{11}$$

In thermodynamic equilibrium we have

$$\mu(\mathrm{A}) + \mu(\mathrm{B}) + \ldots = 0. \tag{12}$$

In the following we investigate the formation of a Ga vacancy in GaAs. This reaction reads:

$$0 \rightleftarrows \mathrm{V}^{(m)}_{\mathrm{Ga},\beta} + \tilde{\mathrm{G}}\mathrm{a} + m \cdot \mathrm{e}^{-}\ (E_{\mathrm{F}}), \tag{13}$$

where $\mathrm{V}^{(m)}_{\mathrm{Ga},\beta}$ denotes a vacancy on a Ga site in the βth layer. For the surface layer $\beta = 1$ and for the crystal bulk $\beta = \infty$. The defect charge state is indicated by the exponent (m), which gives it in terms of positive elemental charges. As the vacancy in the reaction of eqn. (13) is in the mth charge state, m electrons e^{-} are transferred to (or from, if $m < 0$) the electron reservoir, the Fermi level. We assume that this is determined by the background doping of the crystal. $\tilde{\mathrm{G}}\mathrm{a}$ in eqn. (13) denotes a neutral Ga atom in the Ga reservoir. Its chemical potential is $\mu(\tilde{\mathrm{G}}\mathrm{a})$ and the chemical potential of the vacancies depend on the concentration as:

$$\mu(\mathrm{V}^{(m)}_{\mathrm{Ga},\beta}) = \mu^{0}(\mathrm{V}^{(m)}_{\mathrm{Ga},\beta}) + k_{\mathrm{B}} T \ln (N_{\mathrm{V}^{(m)}_{\mathrm{Ga},\beta}}/N). \tag{14}$$

$N_{\mathrm{V}^{(m)}_{\mathrm{Ga},\beta}}$ is the number of m-charged Ga vacancies and N is the number of Ga lattice sites per layer. Equations (13) and (14) hold quite generally because the specific experimental conditions are contained in the values of the Ga and electron chemical potentials.

Surface vacancies ($\beta = 1$) differ from bulk vacancies because of a different electronic structure (see the discussion of fig. 4 below). However, from the second or third layer on, the main difference between a near surface and a bulk vacancy is due to a different electrostatic field caused by ionized defects (space-charge region), if such charged defects are present. For $\beta = 1$ or 2 a distinction between electrostatic and chemical effects appears to be difficult. We may write:

$$\mu(\mathrm{V}^{(m)}_{\mathrm{Ga},\beta}) = \mu(\mathrm{V}^{(m)}_{\mathrm{Ga},\infty}) + m \cdot \mathrm{e}^{+} \cdot \varphi_{\beta}, \quad \text{if } \beta > 2. \tag{15}$$

φ_{β} is an average electrostatic potential in the βth layer and we used the convention that $\varphi_{\infty} = 0$. Thus for $\beta > 2$ $\mu(\mathrm{V}^{(m)}_{\mathrm{Ga},\beta})$ equals the electrochemical potential of vacancies.

For defect formation in the bulk at or below room temperature the defects may not be in equilibrium with the surface. Then one should consider a 'partial equilibrium', neglecting the existence of surface and crystal environment. The removed Ga atom will then go to bulk interstitial positions. Thus, the Ga reservoir ('R1') for the reaction of eqn. (13) consists of bulk interstitials. Their chemical potential is

$$\mu(\tilde{\mathrm{G}}\mathrm{a}) \equiv \mu(\mathrm{Ga}^{(0)}_{i,\infty}) = \mu^{0}(\mathrm{Ga}^{(n)}_{i,\infty}) + n \cdot E_{\mathrm{F}} + \ln (N_{\mathrm{Ga}^{(n)}_{i,\infty}}/\alpha N). \tag{16}$$

Here $\mathrm{Ga}^{(n)}_{i,\infty}$ denotes a bulk ($\beta = \infty$) Ga-interstitial in charge state n. $N_{\mathrm{Ga}^{(n)}_{i,\infty}}$ is their number per layer. α is an integer giving the number of available interstitial sites per

lattice site. Strictly speaking, these Ga interstitials are not a true reservoir, but this is not important here. Assuming that the crystal bulk is perfect and the number of Ga interstitials equals that of Ga vacancies, the Ga vacancy concentration then follows from eqns. (14) and (16) as

$$\frac{N_{V_{Ga,\infty}^{(m)}}}{N} = \sqrt{\alpha} \exp\{[-\mu^0(V_{Ga,\infty}^{(m)}) - \mu^0(Ga_{i,\infty}^{(n)}) - (m+n)\cdot E_F]/2k_BT\}. \quad (17)$$

The energy $\mu^0(V_{Ga,\infty}^{(m)}) + \mu^0(Ga_{i,\infty}^{(n)}) + (m+n)E_F$ has been calculated by Baraff and Schlüter (1985). It varies between 6·2 and 4·5 eV, depending on the Fermi level. Note that it is multiplied in eqn. (17) with a factor $\frac{1}{2}$, due to the fact that two defects (vacancy and interstitial) are created simultaneously.

On the other hand, the formation of defects close to a surface (in the second or third atomic layer, for example) or directly at the surface differs by three important aspects from the bulk situation. First, the electronic structure of a surface defect differs from the bulk defect (see the discussion of fig. 4, below). Second, the electrostatic field at the surface may differ from the average field in the bulk, due to the electron dipole layer and due to charged defects (space-charge region). Third, the Ga chemical potential is not only determined by the crystal bulk, but also by the environment. We consider two extreme possibilities:

(1) The reservoir 'R2' which consists of droplets of Ga metal on the surface, which may occur in Ga-rich environment, and
(2) the reservoir 'R3', which consists of As_2 gas, which can, together with the removed Ga atom, form a new GaAs unit cell:

$$Ga \leftrightarrows GaAs - \tfrac{1}{2}As_2. \quad (18)$$

The actual chemical potential can have any value between those of R2 and R3, depending on the partial pressure and composition of the As gas. Taking the energy of a neutral free Ga atom as zero, the energy gain in bringing this Ga atom to reservoir R2 is 2·8 eV (the Ga metal cohesive energy). For reservoir R3 the energy gain is 4·8 eV: breaking a bond between two As atoms in As_2 costs 1·98 eV per atom and forming a new GaAs unit cell brings a gain of 6·8 eV. Thus we find for the Ga chemical potential:

$$-4{\cdot}8\,\text{eV} \leqslant \mu(\tilde{\text{Ga}}) \leqslant -2{\cdot}8\,\text{eV}. \quad (19)$$

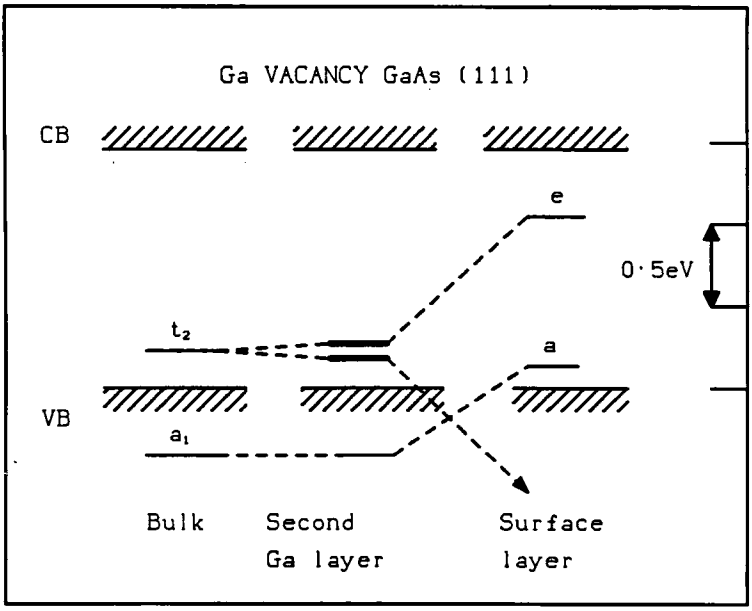
The values of Ga metal and GaAs cohesive energy and the binding energy of As_2 are taken from experiment (Hultgren *et al.* 1973, Weast 1980). Of course, they could also be calculated. The vacancy concentration in the βth layer is

$$\frac{N_{V_{Ga,\beta}^{(m)}}}{N} = \exp\{[-\mu^0(V_{Ga,\beta}^{(m)}) - \mu(\tilde{\text{Ga}}) - m\cdot E_F]/k_BT\}. \quad (20)$$

It is obvious from eqn. (20) that the vacancy concentration is controlled by E_F and the crystal environment, that is, by $\mu(\tilde{\text{Ga}})$. A more accurate treatment of the different defect charge states would introduce a correction in the exponentials of eqns. (17) and (20) which is usually not important.

We now focus our study further, considering the unreconstructed Ga-terminated (111) surface. This surface is built entirely from Ga atoms, each having a dangling bond normal to the surface. These dangling bonds are not occupied by electrons. Our above introduced notation of layers β, which build the crystal, refers here to atomic double layers consisting of a pure Ga and a pure As planar layer.

Fig. 4



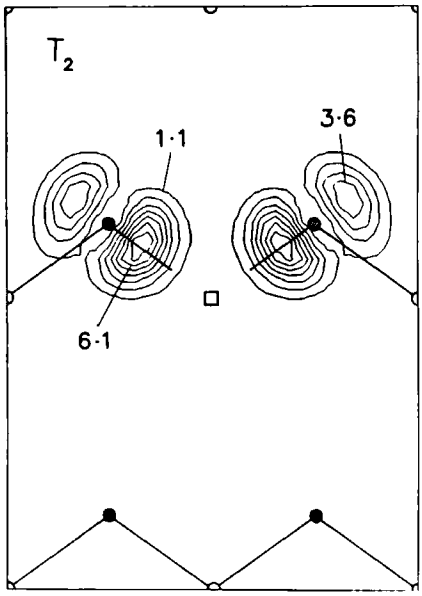
The energy levels of a Ga vacancy in the bulk and at the unreconstructed Ga-terminated GaAs (111) surface.

Figure 4 shows the important energy levels of a Ga vacancy in the bulk and at the unreconstructed Ga-terminated GaAs (111) surface. In the bulk each atom has four nearest neighbours, located at tetrahedral positions. At the surface, which is considered here, the Ga surface atom is bound only to three As atoms, which lie in the second layer.

The results of our super-cell calculation for the bulk vacancy are practically identical to those of earlier self-consistent Green-function calculations (Bachelet, Baraff and Schlüter 1981, Scheffler and Scherz 1986). The electronic structure of a Ga vacancy in the bulk can be understood in terms of four dangling orbitals which are centred at the vacancy's nearest neighbours, namely at the As atoms. Neglecting lattice distortions, these four orbitals give rise to two one-electron states, a singlet a_1 state and triplet t_2 state (see fig. 4). The result of our self-consistent calculation for the t_2 state is shown in fig. 5. The orbital character is mainly arsenic p-like. This picture also holds qualitatively when the vacancy sits close to the surface, as in the second Ga layer ($\beta = 2$), which is the third atomic layer from the surface. Only when the Ga vacancy is directly in the surface layer ($\beta = 1$) is the situation changed significantly, since the Ga vacancy has only three As neighbours. As a consequence the bulk t_2 level splits (see fig. 4).

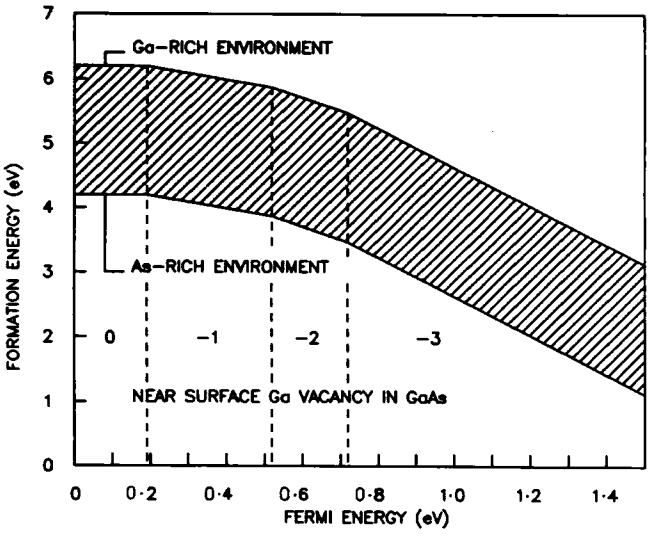
Figure 6 shows the energy $\mu^0 (V^{(m)}_{Ga,\beta}) + \mu(\tilde{Ga}) + mE_F$ calculated for a near-surface Ga vacancy. Defect-induced distortions are neglected as well as electrostatic fields due to defects (space-charge). The latter would change the energies by $m \cdot e^+ \cdot \varphi_\beta$. The near-surface vacancy can exist in four charge states: neutral, negative, double negative and triple negative, corresponding to an occupancy of the t_2 level with three, four, five and six electrons. Because the total system must be neutral, the additional electrons are taken from the crystal Fermi energy E_F. As a consequence, the formation energy depends on the Fermi-level position, that is, on the electron chemical potential. Furthermore, the formation energy also depends on the Ga chemical potential. In fig. 6

Fig. 5



Electron density of the t_2 level wavefunction of a Ga vacancy in bulk GaAs along the (110) plane.

Fig. 6



The formation energy calculated for a near-surface Ga vacancy.

we assume that this is determined by the crystal environment, namely the reservoirs R2 and R3. As mentioned above, any value between these limits is also possible (see eqn. (19)).

The calculations behind fig. 6 are performed in two steps. At first we remove a neutral Ga atom (creating a neutral vacancy) and take it to infinity. Then, in the second step, we bring the neutral atom back from infinity to the reservoir. The Fermi-energy dependence of fig. 6 follows from the theoretical energy-level structure (compare the middle part of fig. 4) and the calculated electron–electron interaction energy of the t_2 level.

The important result of this study is that for n-type material vacancy formation is particularly favourable, if thermodynamic equilibrium with the crystal environment is possible. If the Ga vacancy is formed *directly* at the surface the formation energy is lowered (even more) so that the formation energy becomes negative. This means that, at an unreconstructed surface, vacancies form automatically. In other words, the unreconstructed Ga-terminated (111) surface is not stable. This result is consistent with a recent study of Kaxiras *et al.* (1986). These authors investigated the $c(2\times 2)$ reconstruction of GaAs (111). They showed that in a Ga-rich environment a Ga-vacancy structure (every fourth surface atom missing) on the Ga-terminated surface has the lowest total energy. For an As-rich environment an As-triangle $c(2\times 2)$ structure should have even lower energy.

§ 6. Conclusion

In recent years it has become possible to perform parameter-free calculations of the total energy, forces and atomic vibrations of defects in semiconductors. As a consequence, it is now also possible to evaluate the entropy and thermodynamic potentials, such as the Gibbs free energy, from first principles. The latter two steps are only now being considered in such calculations (see § 4.3).

Total-energy and force calculations had already facilitated the understanding of defect formation, lattice distortions, reactions and migration in cases where experimental methods were impractical or gave incomplete information (see for example Beeler *et al.* (1985) and Scheffler (1987)). Whereas in earlier investigations we had applied the self-consistent Green-function method, the calculations discussed in this paper were based on the super-cell approach. It turns out that this approach gives essentially the same results as the Green-function method, although the spurious 'defect band structure' may cause some difficulties if the super-cell is too small. Often a cell size of between 16 and 54 atoms seems appropriate. For defect-induced distortions it is also important to ensure that the basis set is sufficiently complete. It is our impression that this aspect was sometimes not taken seriously in earlier super-cell calculations. Bigger unit cells (54 atoms, for example) together with good basis sets have become tractable only recently, because of the availability of big computers (such as the Cray X-MP and Cray 2) and because of new developments in the theory (such as the Car–Parrinello method: Car and Parrinello (1985)).

The results presented here show that the theory is able to describe the physics behind defect-induced lattice distortions and to give quantitative predictions of relaxation amplitudes (see § 4.2 above and Scheffler (1987)). Our first calculation of a defect-induced change in vibrational entropy presented in § 4.3 indicates that this term

may not be very important for simple defects. Thus at or below room temperature the entropy part of thermodynamic potentials may be negligible compared to the internal energy.

The super-cell approach also allows one to model surfaces. In this paper we presented results for the bulk and near-surface Ga vacancy in GaAs. One result of this investigation is the sensitivity of the formation energy to the electron and atom chemical potentials. A change in either of them can affect the formation energy by several electron volts.

We believe that first-principles theoretical studies can help elucidate defect properties (as well as other solid-state properties) on a quantitative microscopic level. With recent developments and refinements of theoretical methods the impact of theory on the better understanding of chemical bonding, the structure of matter, metastabilities and so on will continue to increase. Nevertheless, there is still room for significant improvements of the theory. Accurate calculations are at present only possible for simple point defects and for periodic structures with small unit cells. Defect complexes of more than three atoms and low-symmetry surface structures or steps have not yet been investigated by first-principles calculations. Maybe the main challenge at present is to go beyond the Born–Oppenheimer approximation. An understanding of high-temperature phenomena and of the instabilities of defects in the bulk and at the surface will probably remain incomplete until this important next step has been taken.

Acknowledgments

We gratefully acknowledge discussions with K.-P. Charlé.

References

Bachelet, G. B., Baraff, G. A., and Schlüter, M., 1981, *Phys. Rev.* B, **24,** 915.
Bachelet, G. B., Hamann, D. R., and Schlüter, M., 1982, *Phys. Rev.* B, **26,** 4199.
Bachelet, G. B., Jacucci, G., Car, R., and Parrinello, M., 1987, *Proceedings of the Eighteenth International Conference on the Physics of Semiconductors* (Singapore: World Scientific), p. 801.
Baraff, G. A., and Schlüter, M., 1978, *Phys. Rev. Lett.*, **41,** 892; 1984, *Phys. Rev.* B, **30,** 1853; 1985, *Phys. Rev. Lett.*, **55,** 1327.
Baraff, G. A., Schlüter, M., and Allan, G., 1983, *Phys. Rev. Lett.*, **50,** 739.
Bar-Yam, Y., and Joannopoulos, J. D., 1984 a, *Phys. Rev. Lett.*, **52,** 1129; 1984 b, *Phy. Rev.* B, **30,** 1844.
Becker, P., and Scheffler, M., 1984, *Acta crystallogr.* A, **40,** Suppl. C-341.
Beeler, F., Scheffler, M., Jepsen, O., and Gunnarsson, O., 1985, *Phys. Rev. Lett.*, **54,** 2525.
Bernholc, J., Lipari, N. O., and Pantelides, S. T., 1978, *Phys. Rev. Lett.*, **41,** 895.
Biernacki, S., Scherz, U., and Scheffler, M., 1988, (to be published).
Car, R., Kelly, P. J., Oshiyama, A., and Pantelides, S. T., 1984, *Phys. Rev. Lett.*, **52,** 1814; 1985, *Ibid.*, **54,** 360.
Car, R., and Parrinello, M., 1985, *Phys. Rev. Lett.*, **55,** 2471.
Ceperley, D. M., and Alder, B. J., 1980, *Phys. Rev. Lett.*, **45,** 566.
Erbil, A., Weber, W., Cargill III, G. S., and Boehme, R. F., 1986, *Phys. Rev.* B, **34,** 1392.
Fleszar, A., and Resta, R., 1986, *Phys. Rev.* B, **34,** 7140.
Froyen, S., and Zunger, A., 1986, *Phys. Rev.* B, **34,** 7451.
Harding, J. H., 1985, *Physica* B, **131,** 13.
Hultgren, R., *et al.*, 1973, *Selected Values of the Thermodynamic Properties of the Elements* (Ohio: American Society for Metals).
Kane, E. O., 1985, *Phys. Rev.* B, **31,** 5199, 7865.
Kaxiras, E., Pandey, K. C., Bar-Yam, T., and Joannopolous, J. D., 1986, *Phys. Rev. Lett.*, **56,** 2819.
Keating, P. N., 1966, *Phys. Rev.*, **145,** 637.

KRAUT, E. A., and HARRISON, W. A., 1985, *J. vac. Sci. Technol.* B, **3,** 1267.
KRÖGER, F. A., 1974, *The Chemistry of Imperfect Crystals* (Amsterdam: North-Holland).
LANNOO, M., and ALLAN, G., 1982, *Phys. Rev.* B, **25,** 4089; 1986, *Ibid.*, **33,** 8789.
LOUIE, S. G., SCHLÜTER, M., CHELIKOWSKY, J. R., and COHEN, M. L., 1976, *Phys. Rev.* B, **13,** 1654.
MADELUNG, O., 1972, *Festkörpertheorie Vol. III* (Berlin: Springer-Verlag).
MARCH, N., and LUNDQVIST, S., (editors) 1984, *The Inhomogeneous Electron Gas* (New York: Plenum).
MARTIN, R. M., 1970, *Phys. Rev.* B, **1,** 4005.
PANDEY, K. C., 1986, *Phys. Rev. Lett.*, **57,** 2287.
SCHEFFLER, M., 1982, *Festkörperprobleme Vol. XXII*, edited by P. Grosse (Braunschweig: Vieweg), p. 115; 1987, *Physica* B, **146,** 176.
SCHEFFLER, M., and SCHERZ, U., 1986, *Mater. Sci. Forum*, **10,** 353.
SCHEFFLER, M., VIGNERON, J. P., and BACHELET, G. B., 1982, *Phys. Rev. Lett.*, **49,** 1765; 1985, *Phys. Rev.* B, **31,** 6541.
SCHLÜTER, M., CHELIKOWSKY, J. R., LOUIE, S. G., and COHEN, M. L., 1975, *Phys. Rev.* B, **12,** 4200.
TORRES, V. J. B., and STONEHAM, A. M., 1985, *Handbook of Interatomic Potentials III: Semiconductors*, AERE Report R-11947.
WEAST, C. R. (editor), 1980, *Handbook of Chemistry and Physics*. Sixtieth edition. (Boca Raton, Florida: Chemical Rubber Company), F-220.