ACCEPTED MANUSCRIPT

Characterisation of negative-U defects in semiconductors

To cite this article before publication: Jose Coutinho *et al* 2020 *J. Phys.: Condens. Matter* in press https://doi.org/10.1088/1361-648X/ab8091

Topical Review

Characterisation of negative-U defects in semiconductors

José Coutinho

i3N, Department of Physics, University of Aveiro, Campus Santiago, 3810-193 Aveiro, Portugal

E-mail: `jose.coutinho@ua.pt`

Vladimir P Markevich and Anthony R Peaker

Photon Science Institute, School of Electrical and Electronic Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom

Abstract. This review aims at providing a retrospective, as well as a description of the state-of-the-art and future prospects regarding the theoretical and experimental characterisation of negative-U defects in semiconductors. This is done by complementing the account with a description of the work that resulted in some of the most detailed, and yet more complex defect models in semiconductors. The essential physics underlying the negative-U behaviour is presented, including electronic correlation, electron-phonon coupling, disproportionation, defect transition levels and rates. Techniques for the analysis of the experimental data and modelling are also introduced, namely defect statistics, kinetics of carrier capture and emission, defect transformation, configuration coordinate diagrams and other tools. We finally include a showcase of several works that led to the identification of some of the most impacting negative-U defects in group-IV and compound semiconductors.

Contents

Negative-U defects in semiconductors   3

1. Preface

The presence of defects and contaminants in crystalline solids, and in particular crystalline semiconductors, is unavoidable. Defects, if not deliberately introduced, necessarily occur due to a fundamental reason: the lowering of free energy by increasing configurational entropy. Contamination, on the other hand, essentially depends on the contact of the sample with alien species, and therefore on the specificities of growth and processing technologies.

Defects and impurities in semiconductors can be intentionally introduced in order to confer the material with customised mechanical, electrical, optical or magnetic functionality. However, even when intentionally present, they may also turn out to be critical as the semiconductor can acquire other prominent, but unwanted properties. For instance, while a certain amount of dissolved oxygen is highly beneficial in terms of the mechanical stability of silicon wafers, under certain conditions the oxygen impurities precipitate and form clusters which act as shallow donors and so increase the availability of electrons for electrical transport at room temperature. This effect alters the electrical specifications of the silicon, and in p-type material type-inversion can occur. Another example is the creation of silicon vacancies in silicon carbide which can act as single-photon emitters for quantum communications. However, they are also responsible for the introduction of several deep electron traps which are highly detrimental in terms of compensation effects in n-type material. Clearly, understanding defect properties, and if possible, their quantum mechanics, is paramount for technologies that range from solar power conversion to quantum computing.

Defects can be classified in many ways, but regarding their geometric properties they are usually distinguished as point-like and extended. Examples of point-defects are vacancies, interstitials, antisites (misplaced atoms in compound crystals) and atomic-scale complexes. On the other hand, extended defects include dislocations, surfaces and interfaces. Regardless of their size and shape, they change the potential in the Hamiltonian of the pristine material. Consequently, that often leads to the appearance of defect states within a spectral range which would be forbidden to the otherwise perfect crystal.

Unlike crystalline states, defect states are localised in real space. They may capture charge carriers depending whether the stabilisation energy (for instance due to bond formation or exchange interactions) overcomes the repulsive energy between the captured charge and other charges already present. In some cases, defects may actually capture more than one carrier of the same type, but now the stabilisation energy has to exceed the repulsion from previously captured

charges as well. This extra cost is often referred to as correlation energy (U). We may find an analogy of this effect in the sequential ionisation of an atom, for instance,

$$\text{He}^0 + I_1 \rightarrow \text{He}^+ + \text{e}^-; \tag{1}$$
$$\text{He}^+ + \text{e}^- + I_2 \rightarrow \text{He}^{++} + 2\text{e}^-, \tag{2}$$
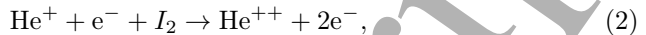
where I_n is the n-th ionization energy of the He atom. Here $I_2 > I_1$, and the difference $U = I_2 - I_1$, which essentially accounts for the electron-electron repulsion and nuclear screening, is obviously positive. Alternatively, U may be defined as

$$U = E(\text{He}^0) + E(\text{He}^{++}) - 2E(\text{He}^+), \tag{2}$$

where $E(\text{He}^q)$ is the energy of the He atom in the q charge state. The fact that $U > 0$, makes a pair of He^+ ions more stable than $\text{He}^0 + \text{He}^{++}$, so that a He^+ gas will not decompose into a mix of He^0 and He^{++} ensembles.

As in the case of isolated atoms, most defects in semiconductors show positive U-values. However, we know today, that many effectively show a negative-U correlation between sequential charging events in apparent defiance of electrostatics. It is as if they became less 'eager' for electrons (holes) immediately after electron (hole) emission!

In the early 1970's, Anderson introduced the idea of negative-U to explain the absence of paramagnetism in glassy-semiconductors doped with n- and p-type impurities [1]. The material was envisaged as a random network of three-state bonds whose potential energy was given by $V = 1/2\,c\,x^2 - \lambda_{\text{ep}}(n_\uparrow + n_\downarrow)x + n_\uparrow n_\downarrow U$. Here x represented a bond coordinate, c its respective harmonic coefficient, and $n_\uparrow$ and $n_\downarrow$ were spin-up and spin-down bond occupancies ($n_\sigma = \{0, 1\}$). The last term accounted for an electronic Hubbard correlation energy [2], with impact only when both electrons pair to form the bond ($n_\uparrow = n_\downarrow = 1$). The striking ingredient of the model was the introduction of an electron-phonon coupling constant λ_{ep} connecting the bond displacement with its occupancy. Accordingly, considering atomic relaxations and the definition of Eq. 2, the net effective correlation involving the three states ($n_\uparrow + n_\downarrow = \{0, 1, 2\}$) is given by $U_{\text{eff}} = U - \lambda_{\text{ep}}^2/c$. It was then noted that for sufficiently strong coupling, U_{eff} becomes negative, thus providing an explanation for why in many glasses and polymers, paramagnetic $n_\uparrow + n_\downarrow = 1$ states are unstable against diamagnetic $n_\uparrow + n_\downarrow = \{0, 2\}$ states.

Although the original idea of Anderson was to describe the formation of bipolarons (where pairing of two electrons and two holes in a localised region is favoured against single polaron formation), the analogy to the negative-U effect in defects is evident. In fact,

Negative-U defects in semiconductors   4

the concept was extended to defects in semiconducting glasses (like As_2Se_3) by Street and Mott [3]. Kastner and co-workers [4] analysed the applicability of the model to several other compounds, emphasising that the dominant contribution to the negative correlation energy is of chemical origin, or in other words, is driven by rebonding of atoms.

In 1979, Baraff, Kane and Schülter anticipated numerically that neutral, positive and double-positively charged vacancies in crystalline silicon (V_{Si}^0, V_{Si}^+ and V_{Si}^{++}, respectively) formed an 'Anderson negative-U' system [5]. A Jahn-Teller distortion in the neutral charge state was claimed to be responsible for the effect, being sufficiently stabilising as to render V_{Si}^+ unstable against V_{Si}^0 and V_{Si}^{++}, irrespective of the position of the Fermi energy. A few months later, these predictions were confirmed experimentally by Watkins and Troxell [6]. The following arguments provided the grounds for the claim: (i) the paramagnetic state V_{Si}^+, as monitored by electron paramagnetic resonance (EPR), was only observed at cryogenic temperatures under photo-excitation. This was an indication that V_{Si}^+ is metastable; (ii) Upon turning off the light, the intensity of the EPR signal bleached at a rate limited by a barrier of 0.05 eV. This figure was interpreted as the activation barrier for hole emission during $V_{Si}^+ \rightarrow V_{Si}^0 + h^+$; (iii) From deep level transient spectroscopy (DLTS), a peak with an activation barrier for hole emission of 0.13 eV, with twice the intensity expected for a single hole emission, could be connected to the following two-hole emission sequence,

$$V_{Si}^{++} \rightarrow V_{Si}^+ + h^+ \rightarrow V_{Si}^0 + 2h^+. \tag{3}$$
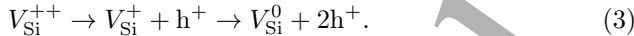
It was then argued that for a negative-U double donor, the hole involved in the first ionisation was bound more strongly (0.13 eV) than the second one (0.05 eV). Thus, at a temperature where the first hole is emitted, the second hole should follow immediately, naturally accounting for the double intensity of the peak.

Baraff, Watkins and others, inspired an entire community towards the characterisation of many other negative-U defects in semiconductors. This account is about their work and the observations and concepts that followed. It aims at reviewing the literature and presenting the most recent results regarding experimental and theoretical methods involved in the characterisation of negative-U defects in semiconductors.

The main text provides a survey regarding several experimental and theoretical reports on negative-U defects (Section 2), an introduction to the physical concepts involved in the description of negative-U defects (Section 3), details regarding experimental and theoretical methods that are used for their characterisation (Section 4), a tabular showcase of a selection of negative-U defects in semiconductors, accompanied by a detailed description of some of

the most interesting and technologically relevant ones (Section 5), and finally a revision of the main concepts, results and challenges for the future of the topic (Section 6).

2. Introduction

Multi-stability is among the most fascinating properties of point defects in semiconductors [7]. In the present context, we refer to a multi-stable defect as one which can be found in several inequivalent atomic configurations for a particular charge state. Along the same lines, a bi-stable defect has two non-degenerate atomic structures in the same charge state. The relative stability as well as formation/annihilation rates of different configurations of multi-stable defects are often sensitive to the application of external stimuli (*e.g.* temperature, electro-magnetic fields or mechanical stress). Hence, under favourable conditions, their populations can be shifted from those observed under equilibrium.

Upon changing the atomistic structure of defects, a change of its electronic structure also takes place. It is therefore expected that stimuli-induced defect reconfigurations may affect significantly the properties of the host material. Exhibition of persistent photo-conductivity, photo-induced capacitance quenching, or temperature-dependent carrier trapping are symptoms that are commonly connected to multi-stability. The magnitude and timescale of these effects depend not only on the properties of the defects themselves, like transformation barriers, cross-sections and transition dipole moments, but also on sample properties like defect concentration and distribution, its thermal history, and of course the measurement conditions.

Negative-U defects are intimately tied to multistability and metastability. These defects have at least three charge states, say with electron occupancy $|N-1\rangle$, $|N\rangle$, and $|N+1\rangle$. Here N is an arbitrary reference and the $|N-m\rangle$ state also includes m electrons at a reservoir with energy E_F per electron (Fermi level). The distinction of negative-U defects is that the intermediate $|N\rangle$ state is metastable and not found under equilibrium conditions. Figure 1 shows a schematic phase diagram of such a defect, where transition boundaries are drawn as a function of the effective correlation energy ($U_{eff} = I_2 - I_1$) and E_F. Here $I_m = E^{(N+1-m)} - E^{(N+2-m)}$ is the m-th ionization energy, with $E^{(N)}$ standing for the energy of state $|N\rangle$. When $U_{eff} > 0$, all three charge states can be populated under equilibrium conditions, depending on the location of the Fermi energy with respect to the transition levels at $E_c - I_1$ and $E_c - I_2$, where E_c is the conduction band bottom. However, for a negative-U defect, only $|N-1\rangle$ or $|N+1\rangle$ states are observed, with
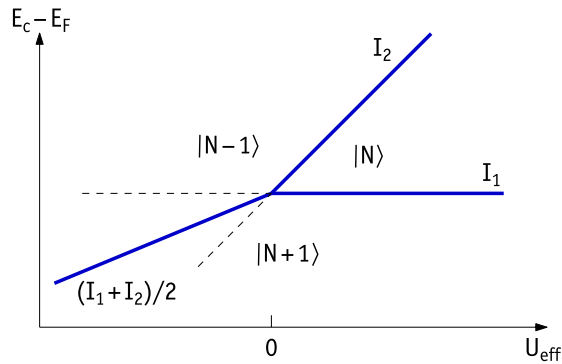
Negative-U defects in semiconductors　　　　　　　　　　5

Figure 1. Phase diagram of a three-charge-state defect, with electron occupancy $|N{-}1\rangle$, $|N\rangle$, and $|N{+}1\rangle$. The horizontal axis is the effective correlation energy ($U_{\mathrm{eff}} = I_2 - I_1$), whereas the vertical axis is the Fermi energy with respect to the conduction band bottom ($E_{\mathrm{c}} - E_{\mathrm{F}}$). First and second ionisation energies of the defect are I_1 and I_2, respectively. Phase boundaries (blue lines) define the stability borders ($U_{\mathrm{eff}}, E_{\mathrm{F}}$) between different charge states. (Adapted from Ref. [8]).

equal populations being found when the Fermi level is located at $E_{\mathrm{c}} - (I_1 + I_2)/2$.

Negative-U defects show strong lattice relaxation effects upon capture and emission of carriers. For that reason, charge state transitions may involve considerable barriers and large Franck-Condon shifts, often making optical absorption and luminescence data hardly comparable to measurements carried out at equilibrium conditions (*e.g.* Hall effect). The metastable $|N\rangle$ state of negative-U defects has a short lifetime, being formed only when sample conditions are away from equilibrium. Although reactions like $|N{-}1\rangle \to |N\rangle + \mathrm{h}^+$ or $|N{+}1\rangle \to |N\rangle + \mathrm{e}^-$ can in principle be stimulated with the application of external perturbations, *e.g.* injection of current or light pulsing, the population of $|N\rangle$ critically depends on a delicate balance between the time of the measurement, the forward and backward reaction (decay) rates, the latter being thermodynamically favoured.

Another difficulty for detection of negative-U defects is related to the fact that the stability of $|N{-}1\rangle$ and $|N{+}1\rangle$ states is usually connected to the formation of closed-shell diamagnetic states, and therefore they are undetectable by EPR [1]. These are among the many features which make the characterisation of negative-U defects a challenge, not only in the laboratory, but also from the perspective of theory and modelling.

The decade of the 1980's was particularly prolific regarding the design and conception of experiments capable of probing the properties of negative-U defects. Watkins and co-workers showed that in a DLTS experiment, after applying a filling pulse to a Si diode irradiated with electrons, carriers trapped at single vacancies (V_{Si}) and boron interstitials (B_{i}) in Si,

were subsequently emitted in pairs under reverse bias [6, 9]. This is a particular signature of the negative-U property, which follows from the relatively fast rate of the second emission compared to the slow rate of the first one. To large extent, this is determined by the condition $I_1 > I_2$. Hence, a conventional DLTS experiment can only measure a single transient, with a decay characteristic of the first (slower) emission. For the case of B_{i}, the measured transient corresponded to electron emission from the negatively charged B_{i}^- (the deeper $|N{+}1\rangle$ state) [9]. Direct detection of the shallower emission $B_{\mathrm{i}}^0 \to B_{\mathrm{i}}^+ + \mathrm{e}^-$ (from the $|N\rangle$ state) was not possible. The series of electrical trap-filling pulses required for DLTS measurements, quickly filled all traps into the $|N{+}1\rangle$ state, no matter how short the applied pulses were. This was later solved by replacing the biased injection of electrons by optical injection into boron-doped diodes with a mesa structure [10, 11].

A prominent example of a multi-stable defect which shows a negative-U ordering of electronic transitions is the M-centre in InP. This defect is observed in electron irradiated (1 MeV) undoped, nominally n-type InP, and was found in at least two main forms (A and B), depending on the bias conditions of the sample [12, 13, 14]. A clear indication that the A-form shows a negative-U ordering of transition levels was found from thermally stimulated capacitance (TSCAP) measurements. While most transitions indicated an ordinary heating-induced loss of a single charge accumulated in a trap, two of them (labelled A2 and A3), occurred simultaneously, thus pointing to a very specific feature of negative-U defects: a two-electron emission event. Despite the detailed experimental data already reported for the M-centre, the only atomistic model available in the literature is phenomenological. It consists of an indium-vacancy-phosphorous-antisite ($V_{\mathrm{In}}P_{\mathrm{In}}$) and phosphorous-vacancy-phosphorous-antisite-pair (V_{P}–$2P_{\mathrm{In}}$), which were assigned to A and B forms, respectively [15].

Another defect showcasing the negative-U effect was firstly reported by Henry and Lang in order to explain the observation of persistent photoconductivity in chalcogen-doped $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ alloys [16, 17]. There was no doubt that the defect involved a shallow donor species (D). However, because it also had a deep state at $\sim E_{\mathrm{c}} - 0.1$ eV, postulated at the time as due to complexing with an undetermined constituent (X), the defect was labelled 'DX' centre. Based on first-principles pseudopotential calculations, Chadi and Chang came up with a model of a deep DX acceptor state localised on a broken S_{As}-Ga (or Al) bond [18]. Accordingly, when the Ga (Al) atom connected to neutral S_{As} (referred to as d^0) moves away along the $\langle 111\rangle$ crystallographic axis towards the interstitial site, the total energy increases as expected for a stable

Negative-U defects in semiconductors   6

structure, but the donor state becomes progressively deeper. When the level has dropped by more than the local repulsive correlation, the total energy is now $E_{\text{cap}} \sim 0.4$ eV above the minimum energy of d^0. At this point a free electron (donated by another S_{As} dopant in the sample) is captured. The subsequent atomistic relaxation stabilises the negative state by $E_{\text{e}} = -0.5$ eV, with DX^- landing at $E_{\text{cap}} - E_{\text{e}} = 0.1$ eV below the d^0 state. The authors underlined that the disproportionation reaction $2d^0 \to d^+ + DX^-$ was exothermal and therefore indicative of a negative-U defect. This provided a natural explanation for the lack of EPR involving the DX^- state.

A refinement of the model was subsequently proposed by Dobaczewski and Kaczor [19], who carried out a detailed analysis of the photoionisation of DX in Te-doped $Al_xGa_{1-x}As$. In their work, the intermediate metastable state could better describe the ionisation kinetics if assumed to be a localised DX^0 defect (as opposed to the neutral shallow donor proposed by Chadi and Chang [18]). It is however consensual that the observed persistent photoconductivity can be accounted for as resulting from the optical ionisation of an equilibrated population of DX^- (in the $|N+1\rangle$ state) into metastable $DX^0 + e^-$ ($|N\rangle$ state), which quickly converts into $d^+ + 2e^-$ ($|N-1\rangle$ state). The latter persists due to a large capture barrier hindering the recovery of DX^-.

Many other reports of negative-U defects followed the above examples, including defects with huge impact on the electronic properties of semiconductors and devices. Prominent examples are the early members of the thermal double donor family of defects in Si and Ge [20, 21], interstitial hydrogen which can be involved in a multitude of solid-state reactions with defects and dopants in several semiconductors [22, 23], the carbon vacancy in 4H- and 6H-SiC which decisively limits the life-time of minority carriers in n-type material [24, 25], or boron-oxygen complexes involved in the light induced degradation of solar Si [26].

The characterisation of negative-U defects involves the experimental monitoring of a two-carrier emission/capture reaction $|N+1\rangle \rightleftarrows |N\rangle + e^- \rightleftarrows |N-1\rangle + 2e^-$. This is a multi-step process involving both electronic transitions and geometric transformations. The kinetics is usually limited by a slow step, effectively 'masking' any subsequent fast steps. Under these circumstances, special tricks have been introduced in order to access the individual steps experimentally. Examples are the combination of junction spectroscopy with light excitation, or the judicious control of the amount of free carriers available for capture by using very short and limited injection pulses. From the above, it is also clear that theory, in particular first-principles modelling of the electronic struc-

ture, has played a huge role in unveiling the workings of negative-U defects and providing guidance for experiments. It is in this spirit of collaboration that we intend to introduce the reader to the topic of negative-U defects in semiconductors. Besides presenting a survey regarding what has been achieved so far, we dedicated much of the content to the concepts and techniques (both experimental and theoretical) involved in the characterisation of this class of defects. We end up with the identification of several problems, from minor loose ends to completely obscure issues, in the hope of stirring up those who may feel challenged by the topic.

3. Negative-U defects in semiconductors

3.1. Electronic correlation

Electronic correlation can be thought of as a measure of how much entangled two or more electrons are, or alternatively, how hard electron motion is within the field of other electrons. It is instructive to look at the concept from the perspective of the Hartree-Fock (HF) method [27]. Accordingly, the wave function is represented by a Slater determinant of M spin-orbitals $\phi_i(\mathbf{x}) = \psi_i(\mathbf{r})\sigma(s_i)$ with space- and spin-dependence $\psi_i(\mathbf{r})$ and $s_i = \{\uparrow, \downarrow\}$, respectively, where $i = 1, \ldots, M$ is a state-index. The manifold $\mathbf{x}$ is a composite of space and spin coordinates. The total energy has contributions from single-electron and two-electron components [28],

$$E_{\text{HF}} = \sum_i H_i + \sum_{i>j} (J_{ij} - K_{ij}), \tag{4}$$

where *core integrals*,

$$H_i = \int d\mathbf{r}^3 \, \psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla_i^2 - \sum_\alpha \frac{Z_\alpha}{|\mathbf{r} - \mathbf{R}_\alpha|} \right) \psi_i(\mathbf{r}) \tag{5}$$

describe a set of independent electrons orbiting in the field of nuclei located at coordinates $\mathbf{R}_\alpha$. Here, the Born-Oppenheimer approximation holds, so that $\mathbf{R}_\alpha$ are fixed (the nuclear kinetic energy vanishes, whereas the nuclear-nuclear repulsion terms add up to an obvious constant which can be included in the result after computing the electronic energy). Interactions between electrons, namely their mutual repulsion and exchange contributions are given by the *Coulomb* and *exchange* integrals,

$$J_{ij} = \iint d\mathbf{r}^3 d\mathbf{r}'^3 \frac{|\psi_i(\mathbf{r})|^2 |\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \tag{6}$$

and

$$K_{ij} = \iint d\mathbf{r}^3 d\mathbf{r}'^3 \, \psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r}') \frac{\langle \sigma(s_i)|\sigma(s_j)\rangle}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}')\psi_j(\mathbf{r}), \tag{7}$$

respectively, and they obey to the relation $J_{ij} \geq K_{ij} \geq 0$. Unlike J_{ij} elements, exchange terms involve

Negative-U defects in semiconductors

the interaction between states with parallel spin only. The scalar product $\langle \sigma(s_i)|\sigma(s_j)\rangle = 0$ when $s_i \neq s_j$, *i.e.* when the interacting spin-orbitals have opposite spin. This follows from the anti-symmetry of the wave function with respect to the exchange of any electronic coordinate.

Equations 6 and 7 translate that each electron effectively interacts with an average charge distribution due to other electrons. This mean-field approximation ignores dynamical correlation, *i.e.*, that electron coordinates are inexorably interdependent, so that they instantaneously avoid one another to reduce their mutual repulsion as much as possible. Nondynamical correlation is also disregarded since the wave function is approximated by a single Slater determinant. Both effects were referred to by Löwdin [29] as total correlation energy, $E_{\text{corr}} = E - E_{\text{HF}}$, where E is the exact non-relativistic energy of the Schrödinger equation. E_{corr} is a very difficult quantity to obtain – a prohibitively large number (millions) of Slater determinants are needed in order to obtain a wave function approaching the exact solution, thus making the problem intractable, even for today's largest supercomputers [30].

If we are solely interested in the ground state, significant progress can be made with the use of density functional theory (DFT) [31, 32, 33], which replaces the HF total energy (a functional of the wave function) by a functional of the electron density ρ [34],

$$E[\rho(\mathbf{r})] = T_0[\rho(\mathbf{r})] + \int \mathrm{d}\mathbf{r}^3 \, v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) +$$
$$+ \frac{1}{2} \iint \mathrm{d}\mathbf{r}^3 \mathrm{d}\mathbf{r}'^3 \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{\text{xc}}[\rho(\mathbf{r})]. \quad (8)$$

Analogous to the core integrals H_i of Eq. 4, the energy functional accounts for the kinetic energy of an arbitrary number of non-interacting electrons (T_0), subject to an external potential (v_{ext}), due to nuclear and other non-electronic fields. The electron density is obtained from summation over occupied orbitals, $\rho(\mathbf{r}) = \sum_{i=\text{occ}} |\psi_i(\mathbf{r})|^2$, which satisfy a set of partial differential equations analogous to single-particle Schrödinger equations. These are solved self-consistently in order to obtain the ground state density (which depends solely on the external potential) [34]. The electron-electron Coulomb repulsion term in Eq. 8) is analogous to the contribution of J_{ij} integrals in HF, although it now includes an unphysical self-interaction energy (electrons are repelled by themselves). In a similar fashion to the exchange integrals K_{ij}, the exchange and correlation functional E_{xc} describes the electronic exchange interactions, but also incorporates all remaining effects, including correlation (absent in HF) and the neutralisation of the above-mentioned spurious self-interactions.

Unfortunately, the mathematical form of E_{xc} is not known (except for an homogeneous electron gas [35, 32, 36]). However, several approximations have been proposed, varying in physical detail and accuracy, and of course in computational load (see for instance Ref. [33] and references therein).

A very simple and intuitive method of treating correlation is that proposed by Hubbard for narrow electronic bands [2]. It nicely describes electronic motion within systems whose electrons are strongly localised on atomic orbitals like Mott insulators. Within the second-quantisation formalism, the Hubbard Hamiltonian is given by [37],

$$\hat{H}_{\text{Hubb}} = t \sum_{\langle ij \rangle, \sigma} \hat{c}_{i,\sigma}^\dagger \hat{c}_{i,\sigma} + U \sum_i \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow} \quad (9)$$

where $\langle ij \rangle$ means that the summation runs over nearest-neighbour atomic sites i and j, while $\hat{c}_{i,\sigma}^\dagger$, $\hat{c}_{i,\sigma}$ and $\hat{n}_{i,\sigma}$ stand for creation, annihilation and number operators for electrons of spin σ on site i. For strongly localised states, electronic motion proceeds via hopping at a rate proportional to the jump integral t, which relates to the band width. This is described by the first term of Eq. 9 and involves electron transfer between neighbouring sites $\langle ij \rangle$. The second term accounts for the fact that each site is capable of accommodating up to two electrons of opposite spin. However, when a site is fully occupied $(\hat{n}_{i,\uparrow} = \hat{n}_{i,\downarrow} = 1)$, both electrons interact with an energy U. Hence, according to the Hubbard model, correlation is the energy raise during an electron transfer event involving two neighbouring sites with identical occupancy (one electron each).

Let us look at the above concept using a pair of oxygen atoms as an example. We know that the energy of a generic atomic system is a piecewise linear function of the electron occupancy (see for instance Refs. [38]). This is illustrated by the Frost diagram of Figure 2 (red line), which shows the change in the total energy for the sequential reduction of atomic oxygen in contact with an electron reservoir (read the diagram from right to left along the horizontal axis). Oxygen is a highly electronegative species with positive first electron affinity $(A_1 = 1.46 \text{ eV})$, and because of that, O^- is more stable than the neutral atom. However, capture of an additional electron is not favourable due to accumulated Coulomb repulsion. This is exhibited by a negative second affinity, $A_2 = -7.71 \text{ eV}$. From the diagram, it is also clear that any (even fractional) charge transfer $0 < \delta \leq 1$ between a pair of O^- anions, $2O^- \rightarrow O^{-1-\delta} + O^{-1+\delta}$, raises the total energy by a correlation energy $\Delta E^{(N+\delta)} = (A_1 - A_2)\delta$. For the transfer of a whole electron $(\delta = 1)$, we obtain a general expression for the total correlation of a reference state with N electrons.

$$U^{(N)} = \Delta^{(N+\delta)} = E^{(N+1)} + E^{(N-1)} - 2E^{(N)}, \quad (10)$$

Negative-U defects in semiconductors   8
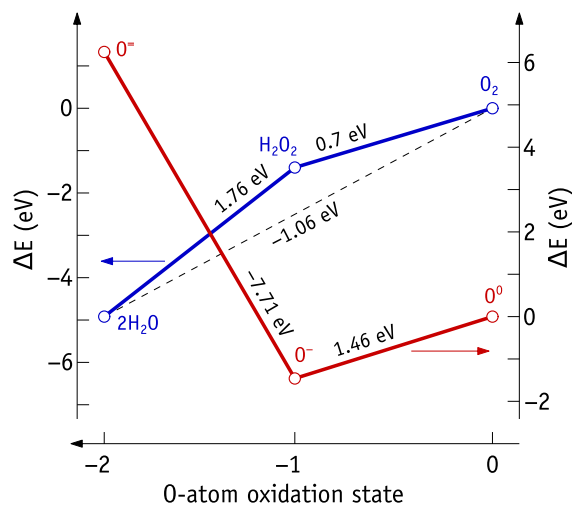

Figure 2. Frost diagram showing the first and second electron affinities of atomic oxygen (red) [39], along with disproportionation energetics of hydrogen peroxide (H_2O_2) in aqueous solution (blue) [40]. Numbers edging solid and dashed line segments represent first- and second-order variations, respectively, of the free-energy per oxygen atom upon oxidation.

where the energies are indexed to the total number of electrons on each state. For the particular case of the O^- atomic state, the correlation energy is $U = 1.46 + 7.71 = 9.17$ eV.

Strictly speaking, correlation is exclusively an electronic property, and that is unambiguously defined for a single atom by Eq. 10. However, for a more complex moiety (*e.g.* a molecule, atom in solution, defect in a solid, *etc.*), the atomistic geometry, vibrational modes and electronic structure of ionised $|N+1\rangle$ and $|N-1\rangle$ states may differ substantially from the those of the reference $|N\rangle$-electron state. It is convenient to recall at this point the concept of effective correlation energy, which besides the electronic contribution, includes a relaxation energy,

$$U_{\text{eff}}^{(N)} = U^{(N)} - \left(\Delta U_{\text{rel}}^{(N+1)} + \Delta U_{\text{rel}}^{(N-1)} \right)$$
$$= U^{(N)} - \Delta U_{\text{rel}}^{(N)}. \tag{11}$$

In the above, $U^{(N)}$ is the electronic correlation as defined at the geometry of the N-electron reference state. On the other hand, $\Delta U_{\text{rel}}^{(N+1)}$ and $\Delta U_{\text{rel}}^{(N-1)}$ are positive relaxation energies, respectively affecting the $|N+1\rangle$ and $|N-1\rangle$ states just after transitioning from the N-electron state. These quantities reflect ionisation-induced atomic reconfigurations, entropy changes, *etc.*, and they are analogous to the Franck-Condon energy dissipated by a molecule or defect following absorption or luminescent transitions [41].

3.2. Disproportionation

An electron transfer event involving two identical moieties with N electrons can be broken into two steps: (i) a sudden electronic excitation where an electron-hole pair is created – this raises the energy by $U^{(N)}$, and (ii) a subsequent electronic/atomistic relaxation that lowers the energy by $\Delta U_{\text{rel}}^{(N)}$. According to Eqs. 10 and 11, for a relaxation as large as $\Delta U_{\text{rel}}^{(N)} > U^{(N)}$, we have $2E^{(N)} > E^{(N+1)} + E^{(N-1)}$. In that case, a close pair of N-electron moieties becomes metastable and disproportionates into $N-1$ and $N+1$ electron species. Disproportionation is a well-known effect in electrochemistry, either taking place spontaneously or via thermal activation, eventually with the help of a catalyst [42]. Although disproportionation necessarily involves electron transfer, it is essentially driven by atomic rearrangement. An eloquent example is given by the reaction

$$2\text{Cu}^+ \, (\text{aq}) \rightarrow \text{Cu} \, (\text{s}) + \text{Cu}^{++} \, (\text{aq}) \tag{12}$$
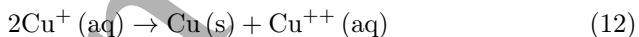
where univalent copper (in solution) precipitates into solid Cu and cupric ions. Being exothermic, Reaction 12 gives $2E(\text{Cu}^+) > E(\text{Cu}^0) + E(\text{Cu}^{++})$, with $E(\text{Cu}^q)$ being the free energy per Cu species in the oxidation state q in its respective phase. In this reaction, the relaxation energy ΔU_{rel} essentially results from the formation of Cu bonds in the copper metal.

Another textbook example of a disproportionation reaction is the decomposition of hydrogen peroxide accelerated at a metallic surface [43],

$$2\text{H}_2\text{O}_2^- \, (\text{aq}) \xrightarrow{\text{cat}} 2\text{H}_2\text{O}^= \, (\text{l}) + \text{O}_2 \, (\text{g}). \tag{13}$$
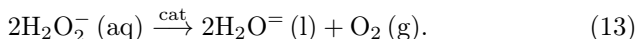
The superscripts in the componds above stand for the oxidation state of oxygen. According to the reaction, for every O^--pair in H_2O_2, one oxygen atom is reduced (to form water), while another becomes oxidised (leading to formation of molecular oxygen). The Frost diagram for Reaction 13 is represented in Figure 2 (blue line). Using Eq. 10, we find that the correlation energy per oxygen atom in H_2O_2 is

$$U_{\text{eff}} = \left[E(2\text{H}_2\text{O}) + E(\text{O}_2) - 2E(\text{H}_2\text{O}_2) \right] / 2 \tag{14}$$
$$= -1.06 \, \text{eV} \tag{15}$$

The factor of $1/2$ arises due to the fact that each compound has two oxygen atoms. The negative effective correlation energy tells us that O^- ions in hydrogen peroxide are metastable against disproportionation into $O^=$ an O^0 in water and molecular oxygen, respectively. In Figure 2, U_{eff} is the free energy difference per oxygen atom, between the H_2O_2 (aq) state and a mix of H_2O (l) and O_2 (g) represented by the dashed line. Like in the copper ion redox Reaction 12, the effective correlation of O^-

ions in H_2O_2 is negative due to atomic reconfiguration, more specifically due to proton transfer and formation of oxygen double bonds after electron transfer.

Many other examples are found in the literature, including the solid state disproportionation of tin(II) oxide (SnO) into tin(VI) and tin dioxide (SnO_2), again involving dramatic structural transformations [44], or the loss of bond length translational order in several perovskites due to random disproportionation of the metallic ions in the lattice [45].

3.3. Electronic transition levels of defects

The term *electronic transition level* of a defect has been given different meanings depending on the context. It has been interpreted differently, depending whether one refers to single-particle or many-body calculations, if it embodies electron-phonon coupling or not, or perhaps if it relates to measurements based on thermodynamic or kinetic quantities. In the present case, an electronic transition level of a defect (or simply *defect level*), is denoted as [46],

$$E(N+1\,/\,N) = E^{(N+1)} - E^{(N)}, \tag{16}$$

and corresponds to the Fermi energy of a material, E_F, above which the ground state of $N+1$ electrons bound to a defect is more stable than any N-electron state (plus one electron at the Fermi reservoir). Analogously, one could say that when E_F drops below a $E(N+1\,/\,N)$ level, the $(N+1)$-electron state becomes unstable against the N-electron state (plus one electron at the Fermi reservoir). When the Fermi level coincides with the defect level, N- and $(N+1)$-state populations are identical. Under such conditions, the following isothermic conversion between $D^{(N)}$ and $D^{(N+1)}$,

$$D^{(N)} + e^- \rightleftarrows D^{(N+1)}, \tag{17}$$

shows identical forward and backward rates and occurs via exchange of electrons (e^-) with a reservoir with Fermi energy $E_F = E(N+1\,/\,N)$.

Let us assume that we have a concentration [D] of non-interacting and identical defects in a sample at temperature T. We also postulate that for the allowed range of E_F values, the defects may occur in up to two charge states, with respective concentrations $[D^{(N)}]$ and $[D^{(N+1)}]$, thus possessing a single electronic transition level $E(N+1\,/\,N)$. Under equilibrium, the fraction of defects with N bound electrons is given by [46, 47, 48],

$$f^{(N)} = Z^{-1} g^{(N)} \exp\left(-E_f^{(N)}/k_B T\right), \tag{18}$$

where k_B is the Boltzmann constant, Z the partition function,

$$Z = \sum_m g^{(m)} \exp\left(-E_f^{(m)}/k_B T\right), \tag{19}$$
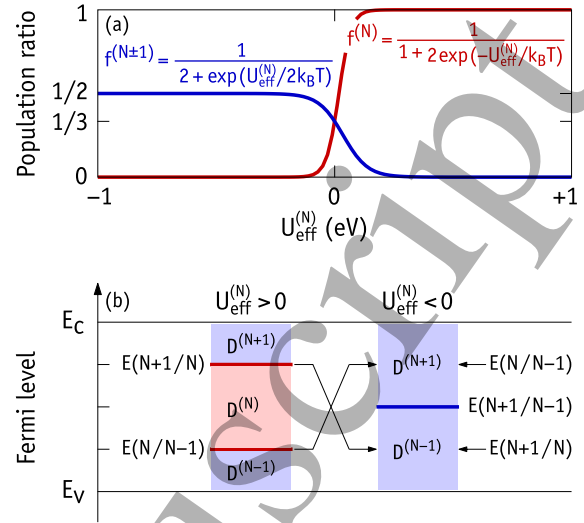
Figure 3. (a) Population fraction of $|N\rangle$ and $|N \pm 1\rangle$ defect states as a function of the effective correlation $U_{eff}^{(N)} = E(N+1\,/\,N) - E(N\,/\,N-1)$. The Fermi level is located at half-way between $E(N+1\,/\,N)$ and $E(N\,/\,N-1)$ transition levels. Other assumptions are a band gap width of $E_g = 1$ eV, $T = 300$ K, and $g^{(N-1)} = g^{(N)} = g^{(N+1)} = 1$. (b) Transition level diagram for positive- and negative-U defects (left and right, respectively). Transition levels are shown as horizontal thick lines. Labels in the shaded areas indicate the most abundant (stable) defect state under equilibrium conditions for the different values of E_F.

and $g^{(m)}$ is a degeneracy factor [47]. In Eqs. 18 and 19, $E_f^{(m)}$ is the formation energy of a stable m-electron state,

$$E_f^{(m)} = E^{(m)} - mE_F \tag{20}$$

with the second term on the right of Eq. 20 representing the chemical potential of electrons in a reservoir with Fermi energy E_F. For our two-state defect this results in the well-known distribution function,

$$f^{(N+1)} = \left[1 + \left(\frac{g^{(N)}}{g^{(N+1)}}\right) \exp\left(\frac{E(N+1\,/\,N) - E_F}{k_B T}\right)\right]^{-1}. \tag{21}$$

Hence, if E_F is lowered below $E(N+1\,/\,N)$, the backward rate of Reaction 17 exceeds the forward rate, and most defects bind N electrons. Conversely, if E_F is raised above $E(N+1\,/\,N)$, occupation of the next available bound state becomes energetically favourable and the reaction proceeds to the right. Defined in this way, an electronic level depends on ground state energies $E^{(N)}$ and $E^{(N+1)}$ only. It is a thermodynamic quantity, analogous to a critical chemical potential for which a phase transition takes place.

Next we consider the case of a defect with two transition levels in the band gap. These define the borders between three charge states, say $D^{(N-1)}$, $D^{(N)}$ and $D^{(N+1)}$, in phase space. Let us inspect the populations of the three charge states by varying the level positions, but keeping the Fermi level locked at

midway between $E(N\,/\,N-1)$ and $E(N+1\,/\,N)$, *i.e.*, $E_\mathrm{F} = [E^{(N+1)} - E^{(N-1)}]/2$. Using Eq. 18, we readily arrive at,

$$f^{(N)} = \left[1 + \frac{g^{(N+1)} + g^{(N-1)}}{g^{(N)}} \exp\left(-\frac{U_\mathrm{eff}^{(N)}}{k_\mathrm{B}T}\right)\right]^{-1} \quad (22)$$

$$f^{(N\pm 1)} = \left[1 + \frac{g^{(N\mp 1)}}{g^{(N\pm 1)}} + \frac{g^{(N)}}{g^{(N\pm 1)}} \exp\left(\frac{U_\mathrm{eff}^{(N)}}{2k_\mathrm{B}T}\right)\right]^{-1} \quad (23)$$

with

$$U_\mathrm{eff}^{(N)} = E(N+1\,/\,N) - E(N\,/\,N-1). \quad (24)$$

The above fractions are plotted in Figure 3(a), where we assume that $g^{(N-1)} = g^{(N)} = g^{(N+1)} = 1$. Under these conditions $f^{(N+1)} = f^{(N-1)}$, thus being represented by a common function $f^{(N\pm 1)}$. The above degeneracy factors do not influence the location of the transition levels when extrapolated to $T \to 0$. In the graph of Figure 3(a) we also assume that the band gap width is $E_\mathrm{g} = 1$ eV and $T = 300$ K.

For a positive-U defect ($U_\mathrm{eff}^{(N)} > 0$), we have $f^{(N)} \approx 1$ (see right-hand side of Figure 3(a)) and the population of the other two states is negligible. On the other hand, for a negative-U defect ($U_\mathrm{eff}^{(N)} < 0$), $\mathrm{D}^{(N)}$ is metastable as seen by the negligible probability of finding this state (left-hand side of Figure 3(a)). Under these conditions, the reaction

$$2\mathrm{D}^{(N)} \rightleftarrows \mathrm{D}^{(N+1)} + \mathrm{D}^{(N-1)}, \quad (25)$$
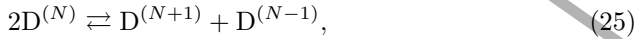
becomes an exothermic disproportionation process with an energy balance,

$$\Delta E_\mathrm{r} = E(N+1\,/\,N) - E(N\,/\,N-1) = U_\mathrm{eff}^{(N)}. \quad (26)$$

When the Fermi level is in the middle of the two levels with negative-U ordering, $\mathrm{D}^{(N-1)}$ and $\mathrm{D}^{(N+1)}$ show identical populations $f^{(N\pm 1)} = 1/2$, implying the existence of a $E(N+1\,/\,N-1)$ transition level at

$$E(N+1\,/\,N-1) = [E(N+1\,/\,N) + E(N\,/\,N-1)]/2, \quad (27)$$

which is represented by the thick blue line in the right-hand side of Figure 3(b). The $E(N\,/\,N-1)$ and $E(N+1\,/\,N)$ levels involve the metastable $\mathrm{D}^{(N)}$ state, so that they cannot correspond to thermodynamic transition levels. Their location is indicated by the arrows in Figure 3(b). Instead, they determine the position of the thermodynamic $E(N+1\,/\,N-1)$ transition, which involves the exchange of two electrons with the Fermi reservoir. Finally, for the peculiar case of $U_\mathrm{eff}^{(N)} = 0$, Figure 3(a) indicates that $f^{(N)} = f^{(N\pm 1)} = 1/3$ so that a triple-point with all states equally populated is attained. Under these conditions, all states in Reaction 25 are equilibrated under isothermic conditions ($\Delta E_\mathrm{r} = 0$).

3.4. Formation energy diagrams

In the dilute limit, where defect-defect interactions are negligible, the equilibrium concentration of a defect on a charge state q depends on its formation energy $E_\mathrm{f}^{(q)}$ as,

$$[\mathrm{D}^{(q)}] = [\mathrm{D}_0^{(q)}] \exp(-E_\mathrm{f}^{(q)}/k_\mathrm{B}T), \quad (28)$$

with $[\mathrm{D}_0^{(q)}]$ being the density of degenerate configurations (lattice sites and orientations per unit volume) that the defect has in the sample. In the above, we now use the charge state q to label the electronic state of the defect (instead of the number of electrons or holes). The formation energy in Eq. 28 expresses the cost to create an isolated defect by trading electrons and atomic species with electronic and atomic reservoirs with chemical potentials E_F and μ_i, respectively (see Ref. [49] and references therein),

$$E_\mathrm{f}^{(q)}(\mu_i, E_\mathrm{F}) = E^{(q)} - \sum_i n_i \mu_i + qE_\mathrm{F}, \quad (29)$$

were $E^{(q)}$ is the energy per defect in charge state q, surrounded by a sufficiently large volume of host material, enclosing n_i atoms of species i (dilute limit). We note that in Eq. 28 we did not consider the contribution of entropy to the formation energy. To do so, $E_\mathrm{f}^{(q)}$ would have to be replaced by a free energy of formation, and $E^{(q)}$ in Eq. 29 would become $H^{(q)} - TS^{(q)}$, with $H^{(q)}$ and $S^{(q)}$ being the enthalpy and entropy per defect in charge state q, the later accounting for vibrational, electronic and magnetic entropy terms. While the formation entropy can be of significant importance, in particular at high temperatures [50], for defects with deep states in the gap, electronic excitations depend exponentially on large activation energies, so that electronic and magnetic entropy can often be neglected. More care has to be taken regarding the vibrational entropy, particularly for negative-U defects where distinct structures with rather different vibrational spectra can occur. Although being usually small when compared to the electronic transitions of deep defects, vibrational entropy can be easily incorporated in Eq. 29 (see for instance Estreicher et al. [51]).

Defect formation energies are in principle positive quantities, otherwise the host material would become unstable against spontaneous defect creation. The graphical representation of the formation energy of a defect as a function of the Fermi level (for a fixed set of atomic chemical potentials) is a common procedure in defect physics and chemistry. It allows us to compare the relative binding energy of electrons to different defects (with variable charge and stoichiometry) with respect to a common reference – the Fermi level.

On the left-hand side of Figure 4 we depict a formation energy diagram of the carbon vacancy

(V_C) located on the pseudo-cubic site of 4H-SiC, for E_F values in the upper part of the band gap (n-type material). The diagram was constructed solely based on experimental observations. The origin of the horizontal axis (right limit of the axis) corresponds to the conduction band bottom. As we move to the left, E_F lowers within the band gap.

The V_C defect at the pseudo-cubic site is a double acceptor that shows negative-U ordering of levels. These were measured by conventional DLTS and Laplace-DLTS at $E_c - E(= /-) = 0.64$ eV and $E_c - E(-/0) = 0.41$ eV [24, 53]. From Eq. 27 and taking the singly negative state as reference, the acceptor states show an effective negative correlation energy of $U_{eff} = -0.23$ eV, and that implies the existence of a thermodynamic transition level at,

$$E(= /0) = [E(= /-) + E(-/0)]/2 = E_c - 0.53\,\text{eV}, (30)$$

while $E(= /-)$ and $E(-/0)$ are metastable. The $E(= /0)$ level establishes that under equilibrium conditions, only V_C^0 or $V_C^=$ states can be found in n-type 4H-SiC. This is highlighted in Figure 4 as solid and dashed lines for the representation of the formation energy of stable and metastable states, respectively. That explains why observation of the V_C^- paramagnetic state requires optical excitation [54]. All three relevant transitions are indicated in the horizontal axis of Figure 4 by vertical lines at the crossing points of the formation energy segments. The slope of each segment is qE_F, clearly distinguishing the charge state of the respective V_C^q defect. The charge neutralisation condition holds via trading of $(2 - q)\text{e}^-$ electrons with the Fermi reservoir. The last piece of information needed to construct the formation energy diagram of Figure 4 is the formation energy of the neutral species (V_C^0). This was measured by means of high temperature annealing/quenching experiments as $E_f^{(0)} = 4.8$ eV [55].

3.5. Configuration coordinate diagrams

While formation energy diagrams capture defect thermodynamics, they are not so useful when it comes to representing the kinetics of transitions which are actually measured by many techniques, including optical spectroscopy and DLTS. In essence, a typical experiment involves monitoring the response of a sample subjected to the application/removal of an external perturbation (e.g. applied field or temperature), leading to a displacement/recovery of the equilibrium conditions. The measured excitation/relaxation rates of defects in semiconductors may involve absorption/emission of light, the exchange of energy and momentum with phonons and electrons within the sample, or a combination of all these processes. Although being applicable to any type of electronic transitions,

we introduce the concept of a configuration coordinate (CC) diagram to illustrate the vibronic nature of non-radiative capture and thermal emission of carriers, which due to the conservation of energy, involve a trade of phonons with the host lattice [16, 56]. This is the most common type of transitions for negative-U defects, where strong electron-phonon coupling effects are normally found.

The process of non-radiative phonon assisted transitions is depicted in the CC diagram of Figure 5, where we assume that the whole system has a single effective vibrational degree of freedom. The $|N\rangle$ ground state finds its minimum at the generalised atomic coordinate $Q^{(N)}$. The upper parabola, which is displaced up in energy by the band gap width (E_g), represents an excited state comprising a $|N\rangle$ defect plus an uncorrelated electron-hole pair (eventually created after illumination of the sample with above-bandgap light). In the middle of the CC diagram we represent a $|N-1\rangle$ state, which due to electron-phonon coupling, has a minimum-energy coordinate shifted to $Q^{(N-1)}$. To avoid unnecessary complications, we assume that the curvature of both $|N\rangle$ and $|N-1\rangle$ parabola corresponds to the same vibrational frequency with quanta $\hbar\omega$.

Thermal activated conversion of $|N\rangle$ into $|N-1\rangle$ can be achieved in two ways: either (i) upon capture of a free-hole followed by multi-phonon emission, or (ii) via emission of a bound electron into the conduction band assisted by multi-phonon capture [57]. In the first case the hole capture coefficient is given by

$$C_p = \sigma_p \langle v_v \rangle_{th} p, (31)$$

where σ_p is the apparent capture cross-section for holes traveling in the valence band top with thermal-average velocity $\langle v_v \rangle_{th}$ and density p. In the second case, the emission rate is

$$e_n = \sigma_n \langle v_c \rangle_{th} N_c \frac{g^{(N-1)}}{g^{(N)}} \exp\left(\frac{E_c - E(N / N-1)}{k_B T}\right), (32)$$

with N_c being the density of available states in the conduction band and $g^{(N-1)}/g^{(N)}$ the degeneracy ratio of the final to the initial state. The quantities σ_n and $\langle v_c \rangle_{th}$ are electron-analogues of σ_p and $\langle v_v \rangle_{th}$.

It is important to note that the above transitions occur close to the crossing points of the energy curves. Henry and Lang [16] have shown that the uncertainty principle implies that the adiabatic approximation (upon which the wave function is instantly consistent with the potential of the oscillating nuclei), breaks down close to these points, so that transitions may in fact occur for Q values in a range where $|E^{(N)} - E^{(N-1)}| \lesssim 60$ meV. This feature is incorporated in the temperature dependence of the effective capture cross-section,

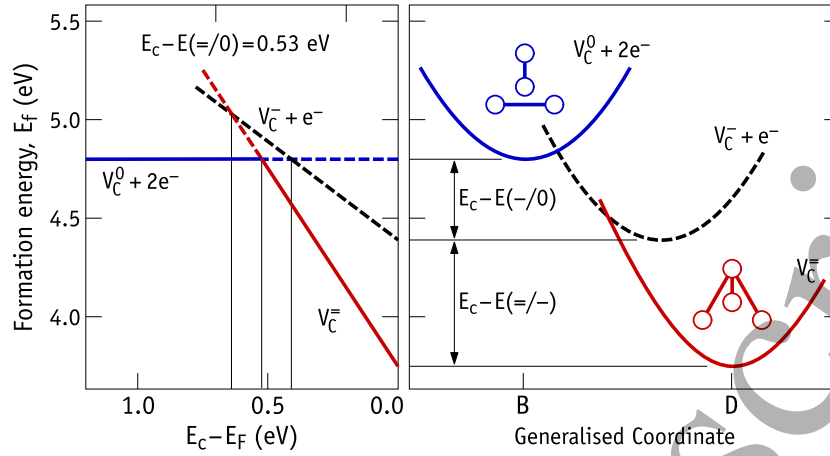$$\sigma = \sigma_\infty \exp(-E_\sigma/k_B T), (33)$$

Figure 4. Formation energy diagram (left) along with a configuration coordinate diagram (right) for the carbon vacancy at the pseudo-cubic site in n-type 4H-SiC. Structures B and D are represented as four Si atoms (white circles) forming an approximate tetrahedron viewed along the main crystallographic axis. The vacant site lies below the Si atom at the centre. Atoms connected by segments are separated by a shorter distance in comparison to equivalent atoms in bulk [52].
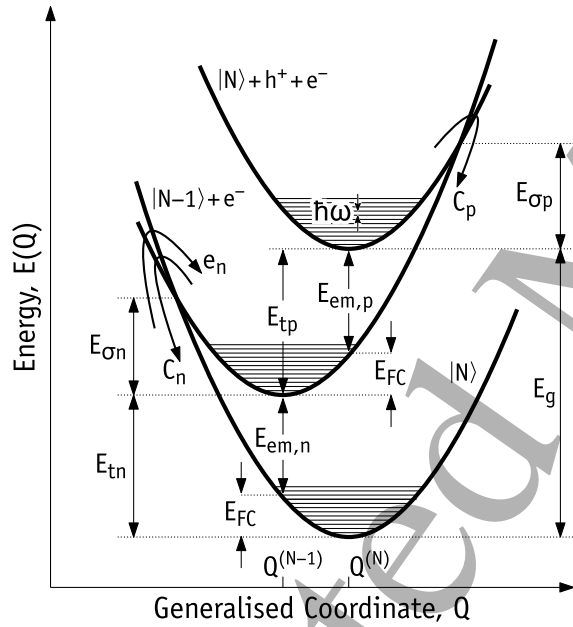
Figure 5. Configuration coordinate diagram of a defect with a transition level responsible for an electron trap located at $E_c - E(N/N-1)$ and a hole trap at $E(N/N-1) - E_v$. Each parabola represents a vibronic state that depends on a generalised coordinate of atoms. Closely spaced horizontal segments represent effective phonon quanta of energy. At the bottom, we show the $|N\rangle$ state, which can emit an electron and become $|N-1\rangle + e^-$, where e^- is a free electron at the conduction band bottom. At the top, we show the $|N\rangle$ state plus an uncorrelated electron-hole pair. $C_{n/p}$ are capture coefficients for electrons/holes, whereas e_n are emission rates for electrons. Electron/hole capture barriers are $E_{\sigma n/p}$. (Reprinted from Ref. [57], with the permission of AIP Publishing).

with σ_∞ being the geometrical (T-independent) capture cross section. Capture rates are therefore thermally activated (see Section 4.2.2 for further details), and according to Figure 5, the capture barrier for electrons or holes is

$$E_{\sigma n/p} = \frac{(E_{tn/p} - E_{FC})^2}{E_{FC}}, \tag{34}$$

where the depth of the trap for electrons is

$$E_{tn} = E_c - E(N/N-1) = E_{em,n} + E_{FC}, \tag{35}$$

or for holes,

$$E_{tp} = E(N/N-1) - E_v = E_{em,p} + E_{FC}. \tag{36}$$

In the above, trap energies are divided into two components: a *vertical emission* and a *Franck-Condon relaxation*, $E_{em,n/p}$ and E_{FC}, respectively (see Figure 5). Within a single-mode approximation, the latter relates to the vibrational frequency by the Huang-Rhys factor as $E_{FC} = S_{HR}\hbar\omega$, which for strongly coupled transitions discloses the dissipation of many phonons ($S_{HR} \gg 1$) [58].

We can now use our example of the carbon vacancy (double acceptor) in 4H-SiC, and combine its CC and formation energy diagrams to obtain a consistent and insightful picture of the measurements. The early DLTS experiments by Hemmingsson *et al.* [24] have found that the conspicuous $Z_{1/2}$ peak corresponds to the superposition of two nearly identical Z_1 and Z_2 negative-U defects, differing only on the sublattice location. These measurements were recently refined by Koizumi *et al.* [59] in 6H-SiC and by Capan *et al.* [53] in 4H-SiC, using Laplace-DLTS. The negative-U ordering of levels implies that during filling of the defect, the binding energy of the second electron is higher than that of the first one. On the other hand,

Negative-U defects in semiconductors 13

for the reverse process, the thermal emission of the first electron immediately induces a second emission.

The above is better perceived with help of Figure 4. Accordingly, before pulsing in the majority carriers (n-type material), the diode is under reverse bias and the Fermi level is lower than the acceptor levels. The formation energy diagram of Figure 4 indicates that the stable state is the neutral one (V_C^0). After applying a filling pulse (zero or forward bias), the Fermi level edges the conduction band minimum and the CC diagram on the right-hand side of Figure 4 applies. In these conditions the double negative state ($V_C^=$) is more stable and the kinetics of the reaction $V_C^0 + 2e^- \rightarrow V_C^=$ is essentially limited by the second capture event [24, 53] – the negative charge accumulated on the defect after the first capture makes it less effective for a second capture. Interestingly, what makes V_C a negative-U defect are strong pseudo-Jahn-Teller relaxations that mainly affect the close shell states (neutral and double negative) [52], driving V_C^- to become metastable (see dashed black lines in Figure 4). The relevant structures (labelled B and D), which were found by first-principles calculations [52], are indicated in the CC diagram, where atoms connected by segments show considerably shorter distances than those in bulk SiC.

When reverse bias is restored, $V_C^=$ becomes unstable, and if the temperature is high enough, two electrons are emitted into the conduction band and swept away from the depletion region. We note that the electron capture barriers for V_C^0 and V_C^- (at the pseudo-cubic site) were found to be as low as $E_{\sigma n}^{(0)} \sim 0$ eV and $E_{\sigma n}^{(-)} \sim 0.03$ eV, respectively [53]. Considering the relative depth of both acceptor levels with respect to E_c, electron emission from $V_C^=$ has a higher barrier than that from V_C^- so that $e_n^{(=)} < e_n^{(-)}$ (where $e_n^{(q)}$ is the electron emission rate from V_C^q). Hence, conventional DLTS is only able to monitor a transient whose decay reflects the first (slower) emission only. This feature is clearly shown in Figure 4, illustrating a major difficulty regarding the characterisation of negative-U defects – probing the intermediate metastable state.

In the particular case of V_C the metastable state is the singly negative defect. Access to this state was achieved by optical excitation [54] and keeping the sample at low temperatures to avoid carrier re-emission, or providing a small amount of electrons to a reverse-biased n-type diode, via short injection [59, 53] or optical pulses [24]. The injection level must be well below saturation limit so that the fraction of defects in the $V_C^=$ state is much smaller than that in the V_C^- state.

Although we can only observe a single emission peak by means of junction spectroscopy, the underlying double emission sequence leads to a variation of the capacitance twice as large to that observed for a transition involving a single emission. As we will see in the next Section, this stands as a rather characteristic feature which is often helpful in the identification of negative-U defects.

4. Methods of characterisation

4.1. Probing negative-U defects

For the sake of convenience, below we refer to the effective correlation energy simply as $U \equiv U_{eff}$. It has already been mentioned above that under equilibrium conditions, defects with negative-U properties emit or capture charge carriers by pairs. So, a confirmation of the negative-U property of a defect requires evidence of such paired emission or capture of charge carriers.

The first definitive experimental evidence of negative-U properties of a defect in a crystalline semiconductor, the lattice vacancy in silicon (V_{Si}), was obtained from DLTS measurements [6]. This technique is frequently used for the determination of the concentration of defects with deep levels. Accordingly, the magnitude of a measured signal (change in sample capacitance due to carrier emission from a defect), is usually directly proportional to the concentration the defect trap it refers to, $\Delta C \sim [D]$ [60]. However, it was found in Ref. [6] that the magnitude of the DLTS signal due to hole emission from V_{Si}^{++} was proportional to $2[V_{Si}]$, so confirming that each emission event was in fact a two-hole emission sequence, so the vacancy in Si is a defect with $U < 0$.

It should be mentioned that for the conclusion about $\Delta C \sim 2[D]$ in the recorded DLTS spectra, an independent method of determination of the vacancy concentration was used. Reliable independent methods for the determination of the concentration of a defect in semiconductor samples for DLTS measurements are rarely possible, so from a conventional analysis of emission signals in the DLTS spectra, it is usually not possible to judge if a carrier emission signal is related to a defect with positive or negative effective correlation energy. Simple analysis of temperature dependencies of carrier emission rates measured with DLTS for a negative-U defect gives only information about parameters for the emission (activation energy for emission and apparent capture cross section) of the first, more strongly bound, charge carrier.

4.1.1. Elucidation of the nature of a defect by means of analysis of its occupancy with charge carriers at equilibrium conditions Solid evidence of negative-U properties for a number of defects in semiconductors has been obtained from studies of their occupancy with charge carriers at equilibrium conditions. The

Negative-U defects in semiconductors

14

general statistics of charge distribution at equilibrium conditions for defects with several trapping levels in semiconductors has been presented in 1957 by Shockley and Last [46]. It has been shown that the concentration ratio of defects in two charge states differing by one electron is

$$\frac{[D^{(N)}]}{[D^{(N-1)}]} = \exp\left(-\frac{E(N/N-1) - E_F}{k_B T}\right), \quad (37)$$

where N is the number of electrons in the most negative state and $E(N/N-1)$ is the energy level of the defect (*c.f.* Section 3.3).

For a defect with two energy levels, the relationship between them can by represented by

$$E(N+1/N) = E(N/N-1) + U. \quad (38)$$

If $U < 0$ and $|U| \gg k_B T$, it follows from Eq. 37 that for any position of E_F, $[D^{(N)}] \ll [D^{(N+1)}] + [D^{(N-1)}]$, *i.e.* $[D^{(N+1)}] + [D^{(N-1)}] \cong [D]$, where $[D]$ is the total concentration of the defect. Furthermore, the electron occupancy function of the $|N+1\rangle$ state is

$$f_{U<0}^{(N+1)} = \frac{[D^{(N+1)}]}{[D]} =$$

$$= \left\{1 + \exp\left[2\frac{E(N+1/N-1) - E_F}{k_B T}\right]\right\}^{-1}, \quad (39)$$

where the quantity $E(N+1/N-1) = [E(N+1/N) + E(N/N-1)]/2$ represents the occupancy level of the defect with $U < 0$. When $E_F > E(N+1/N-1)$, the defect has $N+1$ electrons, when $E_F < E(N+1/N-1)$, the defect has $N-1$ electrons. It is the difference between Equation 39 and the Fermi function, the later describing the occupancy of single-electron defect levels, that allows us to determine the nature of a defect (the sign of the effective electron correlation energy) from a study of its occupancy with charge carriers.

The free energy of ionisation with respect to a reference level (for instance the conduction band bottom), can be presented as

$$\Delta E(N+1/N-1) = E_c - E(N+1/N-1) \quad (40)$$

$$= \Delta H - T\Delta S, \quad (41)$$

where ΔH and ΔS are the changes in enthalpy and entropy due to the $(N+1/N-1)$ ionisation event, and it is temperature dependent if $\Delta S(N+1/N-1) \neq 0$. As an illustration, Figure 6 compares dependencies of charge occupancy versus Fermi level position in the gap (i) for a defect with $U < 0$ and $E_c - E(N+1/N-1) = 0.3$ eV and (ii) for a defect with $U > 0$ and $E_c - E(N+1/N) = 0.3$ eV in n-type silicon, with fully ionised shallow donors providing a density $n = 1 \times 10^{15}$ cm^{-3} of free electrons. The dependencies have been calculated assuming that $n \gg [D]$, and
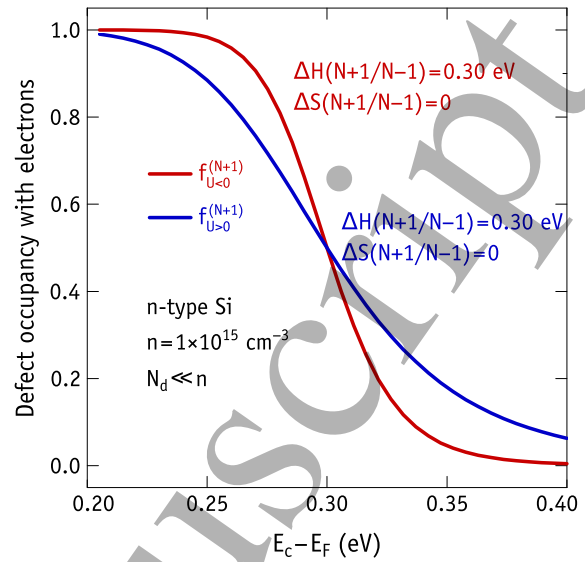
Figure 6. Dependencies of occupancy with electrons ($f^{(N+1)}$) versus the Fermi level position relative to the conduction band edge for a defect with $U < 0$ and $E_c - E(N+1/N-1) = 0.3$ eV and for a defect with $U > 0$ and $E_c - E(N+1/N) = 0.3$ eV. The dependencies have been calculated with the use of Eq. 39 (for $U < 0$) and the Fermi function (for $U > 0$) upon the assumption $n = 1 \times 10^{15}$ cm$^{-3} \gg [D]$.

changes in the Fermi level position have been induced by temperature variations.

The simplest way to probe the occupancy of a defect with charge carriers is to change the temperature of a semiconductor sample in a certain range and monitor associated changes either in free carrier concentration measured by means of Hall effect, or in magnitude of a signal related to emission or capture of charge carriers measured by means of DLTS. Negative-U properties for a number of defects in semiconductors have been elucidated from analyses of temperature dependencies of free carrier concentrations, $n(p) \sim f(T)$. Particularly, evidence of negative-U properties of the Si vacancy [61, 62], oxygen-related thermal double donors (TDDs) in silicon and germanium [63, 20, 21], and complexes consisting of a Si interstitial atom with oxygen dimer (IO_2) [64] and interstitial carbon, oxygen and hydrogen atoms (C_iO_iH) [65] in Si has been obtained from analyses of temperature dependencies of electron (hole) concentrations.

The experimental $n(p)$ *v.s.* T dependencies can be analysed either by methods based on solving charge carrier neutrality equations [66] or by the differential method proposed by Hoffmann [67, 68]. According to the differential method, the concentration of a defect and position of its energy level in the gap can be determined from a dependency of $X = k_B T(dn/dE_F)$ versus E_F in semiconductors of n-type [$X = k_B T(dp/dE_F)$ in semiconductors of p-type]. The $X(E_F)$ dependencies look like spectra with spectral
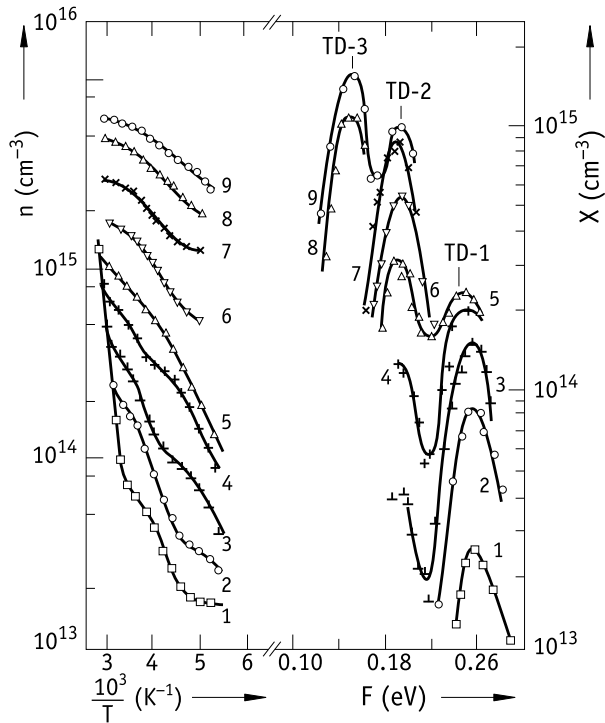
Figure 7. Temperature dependencies of free electron concentrations, n, (left part) and its derivative $X = k_B T(dn/dE_F)$ (right part) versus the Fermi level with respect to the conduction band edge $F = E_c - E_F$, for an oxygen-rich Ge crystal subjected to heat-treatments at 623 K for (1) 5 min, (2) 10 min, (3) 25 min, (4) 60 min, (5) 120 min, (6) 240 min, (7) 480 min, (8) 1020 min, and (9) 1860 min. (Reproduced with permission from Litvinov *et al.* [21], © 1988, WILEY-VCH Verlag GmbH & Co.).

bands induced by ionisation of defects. The magnitude of a peak (X_m) is proportional to the concentration of a defect, while the peak position (E_m) corresponds to the position of an energy level in the gap. It has been shown in Refs. [67] and [68] that for defects with $U > 0$, $X_m = 0.25[D]$ and the half-width of a band, δE_F, is about $3.5k_B T_m$, where T_m is the temperature which corresponds to E_m. For defects with $U < 0$ on the other hand, $X_m = [D]$ and $\delta E_F \approx 1.8k_B T_m$. So, an analysis of the half-width of bands in the $X(E_F)$ dependencies can be considered as a quick test on the sign of effective correlation energy.

As an example, Fig. 7 shows $n(10^3/T)$ and $X(F)$ dependences with $F = E_c - E_F$, for an oxygen-rich Ge crystal, which was subjected to heat-treatments of different durations at 623 K [21]. Such heat-treatments are known to result in the introduction of oxygen-related thermal double donors (TDDs), which consist of a family of subsequently formed (at least nine) defect species in Ge:O crystals [69]. An analysis of the $X(F)$ spectra in Fig. 7 showed that the half-width of the three appearing bands is about $1.8k_B T_m$, so giving solid evidence that the first three species of the TDD family

of defects in Ge are centres with negative-U [21].

From investigations of the occupancy of negative-U defects with charge carriers at equilibrium conditions, only a value of free energy for the two-electron charge state change, $\Delta E(N+1 / N-1)$, can be determined. Further information about the $E(N+1 / N)$ and $E(N / N-1)$ values, charges states, structural configurations and energy barriers between them can be obtained from investigations of non-equilibrium processes of carrier emission and capture.

4.1.2. Observations of negative-U defects in metastable configurations For the majority of defects with $U < 0$, rather large energy barriers exist between their atomic configurations. Because of these barriers, under certain conditions, it is possible to 'freeze' a negative-U defect in a metastable atomic configuration with a shallower level, and to obtain information about the electronic properties of the defect in this configuration. Such 'freezing' or, in other words, deep-shallow excitation can be induced by different methods, *e.g.*, by (i) injection of minority carriers in certain temperature ranges either by above-bandgap-energy light pulses or forward bias pulses in n-p diodes, (ii) a quick change in temperature (quenching), usually from higher to lower T, (iii) cooling down of a semiconductor sample under external illumination with photons having above-bandgap energies, (iv) cooling down an n-p diode with an applied reverse bias voltage. The significant changes in $n(p)$ *v.s.* T dependencies, in DLTS, EPR and various optical (infrared absorption, photoluminescence, *etc.*) spectra induced by the 'freezing'/excitation methods listed above, usually indicate the presence of metastable (negative-U) defect states in a semiconductor material. From analysis of the appearing signals in the excited spectra, electronic and structural characteristics of a negative-U defect in the shallow level configuration can be determined.

In EPR studies, a number of defects, *e.g.*, positively charged vacancy and interstitial boron in silicon [6, 9, 10], Se-related DX centre in AlSb [70] or carbon vacancy in 4H-SiC [54], could only be observed in the spectra after external illumination with above-band-gap light. It has even been argued in Ref. [70] that the absence of an EPR signal in the samples cooled in the dark and its appearance after illumination can be considered as a direct evidence of the negative-U properties. It should be noted, however, that such effects can be related to a defect with a single deep level in the gap and large energy barrier for capture of a charge carrier.

In Si and Ge crystals containing oxygen-related thermal double donors, significant changes in $n(T)$ dependencies and IR absorption spectra have been

Negative-U defects in semiconductors 16

observed after cooling the investigated samples down under external illumination with energy $h\nu \geq E_g$ [63, 20, 71, 21, 72, 73, 69]. In the 'illuminated' IR absorption spectra of Si crystals, three sets of absorption bands, which have not been seen in the spectra recorded after cooling down in the dark, were detected [20, 71, 72, 73]. From the analysis of positions of the electronic-transition related absorption lines in the 'illuminated' spectra, the exact locations of energy levels of three bistable TDD species (negative-U defects) in the shallow donor configuration could be determined in both Si and Ge [20, 71, 72, 73, 69].

Changes in DLTS spectra that resulted from different cooling down conditions of investigated diodes have been reported for semiconductor samples containing both defects with negative-U properties and metastable defects with $U > 0$ [74, 75, 76, 73, 64]. The spectra are usually recorded upon heating the diodes up after their cooling down either with or without a reverse bias applied. Cooling down with the applied reverse bias, when there are no free electrons/holes in the probed 'depletion' region, results in 'freezing' of a negative-U defect in a configuration with a shallower level, which is the minimum energy configuration at high temperatures. So, typically, in the DLTS spectra recorded after cooling down the diodes with the applied reverse bias, a signal due to charge emission from the shallower level appears, and a signal due to emission from the deeper level either disappears or decreases. As an example, Figure 8 compares the DLTS spectra for a Czochralski silicon (Cz-Si) sample containing the IO_2 complex (with I standing for a Si self-interstitial). This complex is a negative-U defect with its first and second donor levels being at $E(0/+) = E_v + 0.12$ eV and $E(+/++) = E_v + 0.36$ eV [64, 77]. The spectra were recorded after cooling down with either bias-on or bias-off. Cooling down with the applied reverse bias resulted in the disappearance of the DLTS signal due to hole emission from the double positively charged state (the minimum energy configuration at low temperatures) of the IO_2 defect and appearance of the signal due to hole emission from the metastable single positively charged state.

Further, back transformations from a metastable to stable configuration can be induced by *e.g.* an increase in temperature if majority charge carriers are available. From studies of kinetics of these backward processes at different temperatures in semiconductor crystals with different free carrier concentrations, a comprehensive set of information about the electronic structure and concentration of defects with negative-U can be obtained.

4.1.3. Non-equilibrium occupancy statistics for defects with $U < 0$
The non-equilibrium occupancy statistics
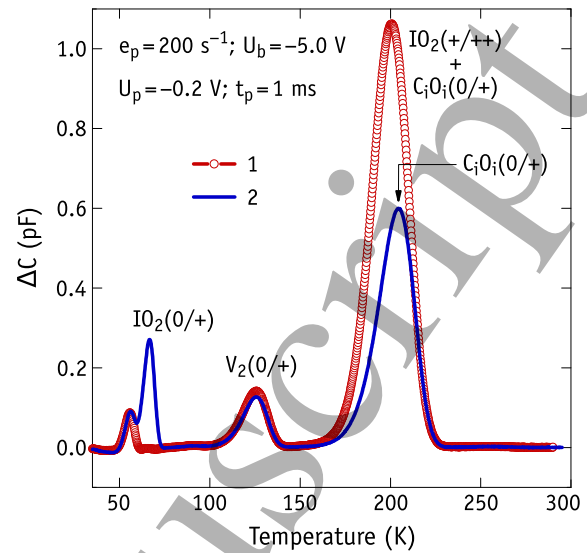
Figure 8. DLTS spectra for a boron-doped Cz-Si sample with low carbon content ($[C] \leq 10^{15}$ cm^{-3}) after its irradiation with 4 MeV electrons. The dose of irradiation was 4×10^{14} cm^{-2}. The spectra were recorded upon heating the sample from 35 K to 300 K with the measurement settings shown in the graph. Spectrum 1 was recorded after cooling the sample down without reverse bias applied to the Schottky diode, spectrum 2 was recorded after cooling down with the reverse bias on.

for defects with $U < 0$ have been developed in Refs. [78, 79, 80]. For an elucidation of the details of these statistics let us consider a total concentration N_d of an amphoteric defect (having acceptor and donor levels) with negative-U level ordering, which exchange charge carriers (electrons) with the conduction band. Figure 9 shows a general configuration-coordinate diagram for such a defect. In the absence of minority carriers, transitions between the stable acceptor (A^-) and donor (D^+) states occur through the metastable X^0 and D^0 states according to the following sequence of reactions:

$$A^- \rightleftarrows A^0 + e^- \rightleftarrows X^0 + e^- \rightleftarrows D^+ + 2e^- \qquad (42)$$
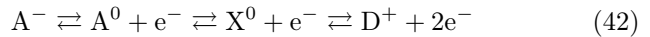
The changes in the density of defect states and in the free electron concentration can be described by the following set of differential equations,

$$\frac{d[A^-]}{dt} = + c_{X^0}[X^0] - e_{A^-}[A^-],$$

$$\frac{d[X^0]}{dt} = - c_{X^0}[X^0] + e_{A^-}[A^-] - \omega_{XD}[X^0] + \omega_{DX}[D^0],$$

$$\frac{d[D^0]}{dt} = + c_{D^+}[D^+] + e_{D^0}[D^0] + \omega_{XD}[X^0] - \omega_{DX}[D^0],$$

$$\frac{d[D^+]}{dt} = - c_{D^+}[D^+] + e_{D^0}[D^0],$$

$$\frac{dn}{dt} = - c_{X^0}[X^0] + e_{A^-}[A^-] - c_{D^+}[D^+] + e_{D^0}[D^0],$$

where emission and capture rates, e_Q and c_Q, are
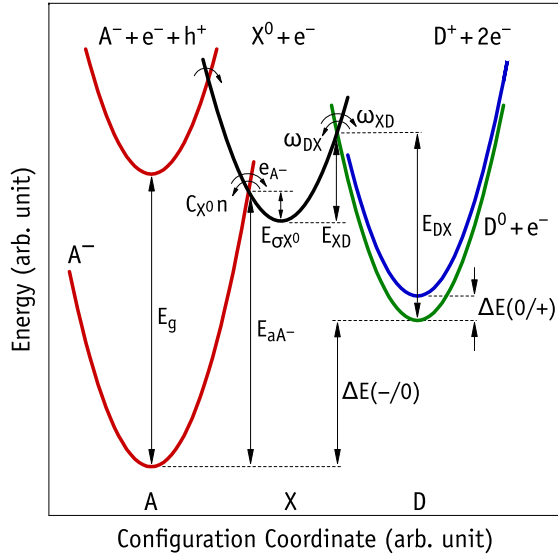
Figure 9. Configuration coordinate diagram for an amphoteric defect (having an acceptor and a donor levels) with $U < 0$.

defined as

$$e_Q = c_{Q'}(N_c/n)\exp[-\Delta E(Q/Q')/k_B T] \qquad (43)$$

and

$$c_Q = \sigma_{\infty Q} n\langle v_c\rangle_{th}\exp(-E_{\sigma Q}/k_B T), \qquad (44)$$

respectively, with Q and Q' referring to neighbouring configurations with different charge. Structural transformations $X \rightleftarrows D$ without a charge state change, occur at a rate $\omega_{QQ'} = \omega_{\infty,QQ'}\exp(-E_{QQ'}/k_B T)$, where now Q and Q' refer to neutral X or D states. Defect concentrations $[Q^q]$ of configuration Q in charge state q, are subject to $\sum_q [Q^q] = N_d$. For the transformation process, $\omega_{\infty,QQ'}$ and $E_{QQ'}$ are the high-temperature attempt frequency and transformation barrier, respectively.

The full solution of the above set of the differential equations is not an easy task. It should be noted, however, that some of the reaction rates are usually much faster than others, so some of the states are in quasi-equilibrium. Furthermore, upon the assumption that $N_d \ll n \approx$ 'const', the above set of equations can be transformed to a set of first order differential equations, which can be solved analytically. Below we will consider in more detail the changes in the density of the A^- state, which is usually monitored in DLTS measurements. The changes in the density of the A^- state at a constant temperature after an initial deviation from equilibrium, $\Delta[A^-]_0$, can be expressed by the following equation,

$$\Delta[A^-](t) = \Delta[A^-]_0\exp(-t/\tau), \qquad (45)$$

where τ is the characteristic time of the decay process. The characteristic time depends on state transition rates as [80]:

$$\tau^{-1} = \frac{\omega_{XD}e_{A^-} + c_{X^0}\omega_{DX}f_D}{\omega_{DX} + c_{X^0}}, \qquad (46)$$

where

$$f_D = \frac{[D^0]}{[D^0] + [D^+]} = \left[1 + \exp\left(-\frac{E_F - E(0/+)}{k_B T}\right)\right]^{-1} \quad (47)$$

$$= \left[1 + \frac{N_c}{n}\exp\left(-\frac{\Delta E(0/+)}{k_B T}\right)\right]^{-1}, \qquad (48)$$

is the occupancy function for the defect in the donor configuration. Equation 46 can be expressed in a common way as

$$\tau^{-1} = e_{eff} + c_{eff}, \qquad (49)$$

with

$$e_{eff} = \frac{e_{A^-}}{1 + c_{X^0}/\omega_{XD}}, \qquad (50)$$

and

$$c_{eff} = \frac{c_{X^0}\omega_{DX}\omega_{XD}^{-1}f_D}{1 + c_{X^0}/\omega_{XD}}. \qquad (51)$$

An analysis of Eqs. 46, 49, 50 and 51 indicates that several terms, with their specific activation energies and power dependence on the free carrier concentration, can dominate the temperature dependence of τ^{-1}. Particularly influential factors are the Fermi level position with respect to $E(-/0)$, $E(0/+)$ and $E(-/+)$, and the c_{X^0}/ω_{XD} ratio. The effective emission rate, e_{eff}, is the dominant term in Eq. 49 when the Fermi level is below the $E(-/+) = E_c - [\Delta E(-/0) + \Delta E(0/+)]/2$ occupancy level of a negative-U defect. At these conditions, $\tau^{-1} = e_{eff}$. When $c_{X^0} \ll \omega_{XD}$, $\tau^{-1} = e_{A^-}$, so the temperature dependence of τ^{-1} can be described by a commonly used equation for single electron emission. For the cases of $c_{X^0} \gg \omega_{XD}$, we have $\tau^{-1} = \omega_{XD}e_{A^-}/c_{X^0}$, so the transformation rate (actually the rate of occupancy of the A^- state) is inversely proportional to the free electron concentration. Such unusual dependence of τ^{-1} versus n has been clearly observed for a complex consisting of a substitutional boron atom and oxygen dimer in silicon [26, 81].

The capture term in Eq. 49 is dominant when $E_F > E(-/+)$. In this case, $\tau^{-1} = c_{eff}$ and up to four specific terms with different values of activation energy and power dependence on n can occur, depending on the c_{X^0}/ω_{XD} ratio and the Fermi level position with respect to $E(0/+)$. For a number of defects, *e.g.* bistable TDDs in Si and Ge crystals, under certain conditions, $\tau^{-1} \sim n^2$ dependencies have been observed

Negative-U defects in semiconductors   18

[78, 21, 79, 82]. Such dependencies clearly show a two-fold change in the charge state of a defect upon capture of carriers, so indicating its negative-U properties.

From the analysis of temperature dependencies of transition rates between configurations of a defect with $U < 0$ monitored by measurements of changes in either free carrier concentration or density of a specific defect state in semiconductor crystals with different n (p), it is possible to elucidate the electronic structure of the defect and determine practically all its energy differences and transformation barriers. This has been done for a number of defects in Ge and Si crystals [78, 21, 79, 80, 64, 82, 26, 81].

4.2. Modelling negative-U defects

4.2.1. Calculation of electronic transition levels The calculation of electronic transition levels is essentially a problem of finding the energy of an electronic reservoir (Fermi level), for which the exchange of electrons with a defective sample becomes energetically favourable (see Section 3.3). This is a topic which has been widely revised in the past (see for instance Ref. [49] and [83]), so we leave here the essential features. We start by recalling Eq. 29, which determines the energy needed to create a defect in a crystal (defect formation energy)

$$E_{\mathrm{f}}^{(q)}(\mathbf{R}, \mu_i, E_{\mathrm{F}}) = E^{(q)}(\mathbf{R}) - \sum_i n_i \mu_i + q E_{\mathrm{F}}. \quad (52)$$

Here, $E^{(q)}(\mathbf{R})$ is the energy of a large portion of a sample enclosing a single defect, including its electronic wave functions. The defective sample volume is made of n_i atoms of species i with chemical potential μ_i and collective atomic coordinate $\mathbf{R} = \{\mathbf{R}_\alpha\}$ (α being an atomic index). It can be easily shown that a transition energy level between charges states q and q' (with q being more negative) is located at

$$E(q/q') = -\frac{E^{(q)}(\mathbf{R}) - E^{(q')}(\mathbf{R}')}{q - q'}. \quad (53)$$

Equation 53 accounts for the fact that charge states q and q' may correspond to radically different atomistic geometries $\mathbf{R}$ and $\mathbf{R}'$.

The large chunk of material with a defect referred above is normally approximated to a defective supercell. The inherent periodic boundary conditions imply the cell to be neutral, even when we add (remove) electrons to (from) a defect state located in the gap. Technically, a uniform background counter-charge of density $-q/\Omega$ (with Ω being the supercell volume) is superimposed to the electronic density. This implies that the aperiodic energy $E^{(q)}$ that enters in Equations 52 and 53, is related to the periodic energy $\widetilde{E}^{(q)}$ obtained from the supercell calculation as $E^{(q)} = \widetilde{E}^{(q)} + E_{\mathrm{pcc}}^{(q)}$, where $E_{\mathrm{pcc}}^{(q)}$ is a correction

often referred to as *periodic charge correction*. Several schemes have been proposed for the calculation of $E_{\mathrm{pcc}}^{(q)}$ (see for instance Ref. [84] and references therein).

Of course, in a semiconductor, transitions are only observable if they are located in the range $E_{\mathrm{v}} < E(q/q') < E_{\mathrm{c}}$. We have therefore to calculate E_{v} or E_{c} in order to cast the levels as measurable quantities, *i.e.*, $E(q/q') - E_{\mathrm{v}}$ or $E_{\mathrm{c}} - E(q/q')$. One possibility is to assume that $E_{\{\mathrm{v,c}\}} = \epsilon_{\{\mathrm{v,c}\},\mathrm{bulk}}$, which are the highest occupied ($\epsilon_{\mathrm{v,bulk}}$) and lowest unoccupied ($\epsilon_{\mathrm{c,bulk}}$) states of a single-particle Hamiltonian (like the Kohn-Sham equations). Another option is to follow the ΔSCF (delta self-consistent field) method to obtain $E_{\mathrm{v}} = \widetilde{E}_{\mathrm{bulk}}^{(0)} - \widetilde{E}_{\mathrm{bulk}}^{(+)}$ or $E_{\mathrm{c}} = \widetilde{E}_{\mathrm{bulk}}^{(-)} - \widetilde{E}_{\mathrm{bulk}}^{(0)}$ from total energies of bulk supercells. These must be identical in size and shape to those used to obtain $\widetilde{E}^{(q)}(\mathbf{R})$ values. In any case, we should be aware that the accuracy of the calculated defect levels, including the energies of the band gap edges, strongly depends on the quality of the Hamiltonian. Particularly important is the level of detail put in the description of the exchange and correlation interactions between electrons (see for instance Ref. [85]).

4.2.2. Calculation of carrier capture cross-sections in multi-phonon assisted transitions Inelastic scattering of free carriers by defects essentially consists of the transfer of energy and momentum of a propagating electron or hole, to another carrier bound to a defect (trap-Auger process), its conversion into radiation (luminescence process), into nuclear motion (*e.g.* capture enhanced defect migration) or heat dissipation (multi-phonon emission process) [58]. Non-radiative capture of carriers via multi-phonon emission (MPE) is therefore a special case of inelastic scattering – the energy drop of the traveling carrier is fully dissipated into atomic vibrations. This is the most common capture mechanism at deep traps involving negative-U defects. Other important processes, like recombination, generation or emission, can be described either as a sequence of consecutive capture events, or by reversing the operation using detailed balance.

The relevant quantity to be evaluated is the carrier capture rate, $c_{n/p}$, which relates to the apparent capture cross-section of Eq. 33 as $c_n = \sigma_n \langle v_{\mathrm{c}} \rangle_{\mathrm{th}} n$ for the case of electrons (with an obvious analogous expression for holes). While first-principles calculations of electronic transition levels have been routinely reported in the literature, the calculation of capture rates is clearly lagging behind. For a historical account regarding the development MPE theory, we divert the reader to Refs. [58, 86, 87]. In general, the calculation of the non-radiative transition rate between free and bound states starts with the description of

a many-body Hamiltonian $\hat{H}$ and the choice of a practical basis to describe the total initial and final wave functions involved, Ψ_i and Ψ_f, respectively. The initial state represents the defect plus uncorrelated free carrier, whereas the final state stands for the trapped state. The *static approximation* is a rather convenient starting point in the context of popular electronic structure methods like Hartree-Fock and density functional theory. Here, the electronic wave functions $\psi_{s;\mathbf{R}_0}(\mathbf{r})$ (s being an electronic state index and $\mathbf{r}$ all electronic degrees of freedom) are found from a static Hamiltonian $\hat{H}_{\mathbf{R}_0}$, constructed for a fixed atomic geometry $\mathbf{R}_0$. The nuclear wave functions are treated separately. Assuming the harmonic approximation, we write the initial state as $\Psi_{im} = \psi_{i;\mathbf{R}_0}(\mathbf{r}, \{Q_{ik}\})\chi_{im}(\{Q_{ik}\})$ and likewise, the final state as $\Psi_{fm} = \psi_{f;\mathbf{R}_0}(\mathbf{r}, \{Q_{fk}\})\chi_{fn}(\{Q_{fk}\})$. Here, m and n are quantum numbers for the vibrational states χ_i and χ_f, respectively, while k indexes are used to identify the $3N$ individual vibrational modes $Q_{\{i,f\}k}$, N being the number of atoms in the sample.

The transition probability between states Ψ_{im} and Ψ_{fn} relates to the off-diagonal matrix elements of the Hamiltonian when perturbed by a change in geometry (away from $\mathbf{R}_0$) induced by the *promoting modes* $\{Q_{ik}\}$ of the initial electronic state. Such quantities may be obtained by first-order expansion of $\hat{H}$ in a Taylor series of Q_{ik} (*linear coupling approximation*),

$$\Delta H_{im,fn} = \sum_k \langle \psi_{i,\mathbf{R}_0}| \frac{\partial \hat{H}}{\partial Q_{ik}} |\psi_{f,\mathbf{R}_0}\rangle \langle \chi_{im}| \Delta Q_{ik} |\chi_{fm}\rangle, \quad (54)$$

where $\Delta Q_{ik} = Q_{ik} - Q_{0k}$ with Q_{0k} representing a generalised coordinate, equivalent to the collective nuclear coordinate $\mathbf{R}_0$ in the $\{Q_{ik}\}$ basis. The transition (capture) rate is then cast in the form of Fermi's golden rule,

$$c = \frac{2\pi}{\hbar} \sum_{m,n} w_m(T)|\Delta H_{im,fn}|^2 \delta(\Delta E_{im,fn}), \quad (55)$$

which sums up all possible vibronic transitions, weighted by the occupancy fraction w_m of the promoting vibronic states. Under thermal equilibrium w_m is simply a normalised Boltzmann distribution. The allowed transitions in Equation 55 are also restricted to those which conserve the energy, hence the use of a Dirac delta $\delta(\Delta E_{im,fn}) = \delta(E_i - E_f + m\hbar\omega_m - n\hbar\omega_n)$. In practice, the δ function is replaced by a normalised Gaussian function with full width $\sim k_BT$ to account for thermal broadening.

Up to this point, several bold approximations were already made. Still, the formulation of $\Delta H_{im,fn}$ as its stands in Eq. 54, makes its calculation a rather cumbersome exercise. The problem lies essentially on the identification of the vibrational modes which lead to non-vanishing $\langle \chi_i|\Delta Q_{ik}|\chi_f\rangle$ terms. To give an idea,

using a triply hydrogenated vacancy (VH_3) defect in silicon as a benchmark, Barmparis and co-workers [87] used a Monte Carlo sampling scheme, being able to find convergence for $\Delta H_{im,fn}$ when about 12 distinct phonon modes were considered.

Motivated by the fact that a judicious choice of a single effective phonon mode resulted in a promising agreement between the observed and calculated luminescence intensity of radiative transitions (see for instance Refs. [88] and [89]), several authors evaluated the transition matrix elements within the single-phonon approximation. Accordingly, the effective mode is the one that maximises the coupling to the distortion during the capture process,

$$Q^2 = \sum_\alpha m_\alpha \lambda^2 |\Delta \mathbf{R}_\alpha|^2. \quad (56)$$

Equation 56 is a linear interpolation between initial and final atomic coordinates $\Delta \mathbf{R}_\alpha = \mathbf{R}_{f,\alpha} - \mathbf{R}_{i,\alpha}$ with α running over all N atoms with mass m_α. The displacement amplitude along the effective mode is governed by a unitless factor $0 \leq \lambda \leq 1$. The units of Q are amu$^{1/2}$ Å (amu being the atomic mass unit). From the effective mode we arrive at the corresponding effective vibrational frequency $\omega_{\{i,f\}}$ of the initial or final states by extracting $\omega_{\{i,f\}} = \partial E_{\{i,f\}}/\partial Q$ from total energy calculations. With this in mind, Equation 55 simplifies into a factorisation of purely electronic and vibrational terms,

$$c = \frac{2\pi}{\hbar} W_{if}^2 G_{if}(T), \quad (57)$$

where the transition matrix element is now $W_{if} = \langle \psi_i|\partial\hat{H}/\partial Q|\psi_f\rangle$ and G_{if} is the temperature-dependent line shape factor

$$G_{if}(T) = \sum_{m,n} w_m(T)|\langle \chi_{im}|\chi_{fn}\rangle|^2 \delta(\Delta E_{im,fn}). \quad (58)$$

Equations 57 and 58 translate the *Franck-Condon approximation*, which separates an instantaneous electronic transition from the *sluggish* phonon dissipation process that follows. The calculation of G_{if} boils down to the Frank-Condon integrals $\langle \chi_{im}|\chi_{fn}\rangle$ for all vibrational promoting and accepting states. For that, recursive techniques have been proposed [90, 91, 92, 93] and applied on different contexts [94, 88, 95]. Sampling methods have also been applied [87], although inclusion of a few phonons implied the evaluation of millions of $\langle \chi_{im}|\chi_{fn}\rangle$ configurations that matched the electronic energy off-set.

Regarding the evaluation of W_{if}, different approaches have also been used. Alkauskas *et al.* [86] replaced the many-body Hamiltonian and wave functions by single-particle counterparts $\hat{h}$ and $\phi_{\{i,f\}}$, which were taken from the Kohn-Sham equations of density-functional theory. Thus, W_{if} becomes

$$W_{if} = \langle \phi_i|\partial\hat{h}/\partial Q|\phi_f\rangle, \quad (59)$$

which can be solved by first-order perturbation theory [86]. Alternatively, a phenomenological approach was pursued by Krasikov *et al.* [95] who used Eq. 88 in the paper of Henry and Lang [16],

$$|W_{\text{if}}|^2 = \frac{2\pi}{\Omega} \left(\frac{\hbar^2}{2m^*} \right)^{3/2} \epsilon^{1/2}, \tag{60}$$

where Ω is the crystal volume, m^* is the effective mass of the free carrier and ϵ is the energy separation between the initial and final states near the intersection point (see discussion in Section 3.5).

4.2.3. Calculation of transformation barriers Defect reorientation and transformation are important processes that occur during the experimental characterisation of negative-U defects. Understanding the underlying physics of such processes is fundamental in order to explain and ultimately to control defect behaviour. The problem is usually addressed in the spirit of transition state theory [96], and usually involves finding the lowest free energy barrier that separates specific reactants from products, or in the case of defects, initial from final structures.

Although being a rather challenging problem, particularly when we have no clue about the mechanisms involved, the search for a saddle-point along a potential energy landscape can be investigated in a number of ways. In general, defect transformation mechanisms are investigated via constrained-relaxation of the atomistic structure, ideally using a highly accurate quantum-mechanical method to evaluate the forces. Among the existing algorithms we highlight:

The *dimer method* [97], which requires knowledge of a single minimum energy configuration, performs a search for a nearby saddle point. This method is particularly useful to look for saddle points along paths with unknown final configurations. The dimer method works on two structures (the *dimer*) separated by a small distance in configurational space. The direction along the lowest curvature of the potential energy is obtained by rotation of the dimer, and by finding the forces for each structure. This result is used to move the dimer uphill, from a stable configuration to a near saddle point.

The *Lanczos method* was originally referred to as ART algorithm by Malek and Mousseau [98] and is a close relative to the dimer method. It can be used to search for a saddle point close to an arbitrary initial configuration. Unlike the dimer method, the Lanczos scheme estimates the Hessian eigenvector that corresponds to the lowest eigenvalue by expansion of the potential energy surface in a Krylov subspace. After this stage, the method proceeds in the same way as the dimer.
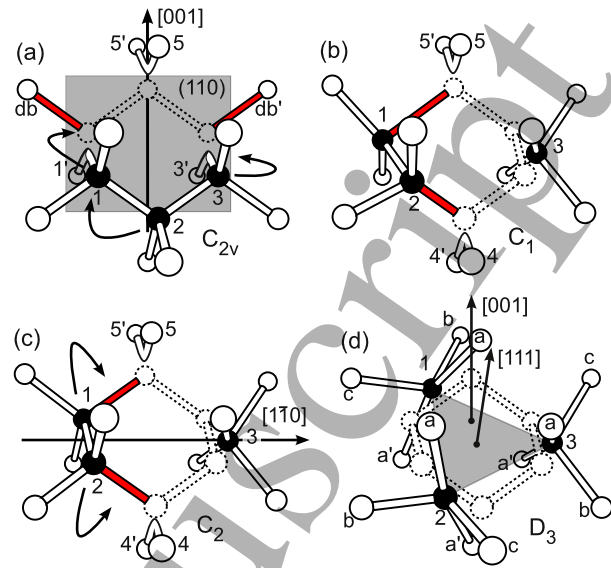
Figure 10. Atomic structure of stable and metastable trivacancy defects in silicon. The transformation mechanism is depicted in four stages (a), (b), (c) and (d) by the arrows indicating the motion of three Si atoms at the core (shown in black for a better perception). Perfect lattice sites are drawn with a dashed line. Broken bonds are shown as solid red sticks. (Reprinted with permission from Ref. [101]. © 2012 by the American Physical Society).

The *nudged elastic band* (NEB) *method* [99, 100] is a robust algorithm that works when the initial and final states of the mechanism are known *a priori*. It basically involves the relaxation of a structure sequence (referred to as intermediate images) along the multi-dimensional path that separates the end-structures. Usually a starting guess for the intermediate images is obtained by linear interpolation of the end-structures. The latter are kept fixed during the relaxation process. During the relaxation of each image, equal spacing between neighbouring images is maintained by inclusion of artificial spring forces.

An example of a defect-related mechanism studied using the NEB method is depicted in Figure 10, showing the transformation of the bistable trivacancy (V_3) complex in Si [101]. This is one of the most important radiation products in heavy particle irradiated Si [102]. Right after irradiation the observed defect has the structure depicted in Figure 10(a). The Si broken bonds on this structure are responsible for several deep donor and acceptor levels observed by DLTS [102]. However, storage for a few weeks of the irradiated Si samples at room temperature result in the transformation to the structure of Figure 10(d). This is accompanied by the disappearance of the deep levels and the introduction of shallow electron traps at 75 meV bellow the conduction band bottom.

The energy barrier for the above transformation was estimated using the NEB method as 1.15 eV

Negative-U defects in semiconductors   21

and 1.14 eV for neutral and negatively charged V_3, respectively [101]. These figures match very well their respective experimental counterparts of 1.16 eV and 1.15 eV [102].

5. Showcase of negative-U defects and their characterisation

In this section we provide some examples of negative-U centres in widely used semiconductors. In the selection we try to illustrate a range of behaviour, the convergence of theory and experiment, and cases which have provided particular challenges in understanding. Among this selection are examples of negative-U centres which are of importance in semiconductor devices. The section is divided in two parts, firstly mini reviews of published work on four defects, 1) B_sO_2 (LID) complex in silicon, 2) atomic hydrogen in silicon, 3) DX centres in III-V alloys and 4) The Ga-O_{As}-Ga defect in GaAs. This is followed by a table which covers a wider range of important negative-U centres providing basic parameters and key references. The table includes the four defects listed above and negative-U defects referred to previously in the text to illustrate particular aspects of negative-U behaviour, and theoretical and experimental challenges in their study.

5.1. B_sO_2 (LID) complex in silicon

A complex consisting of a substitutional boron atom and an oxygen dimer (B_sO_2) in Si was suggested to be responsible for light induced degradation (LID) of solar cells produced from silicon crystals, which contain boron and oxygen impurity atoms [103, 104, 105]. The suggestion was based on the experimental observations of linear and quadratic dependencies of concentration of recombination active defects responsible for LID on substitutional boron and interstitial oxygen concentrations, respectively [103, 104]. However, for almost two decades, there was no clear understanding of the electronic structure of the complex or the mechanisms of its formation and transformations [105]. Answers to the above uncertainties have been found recently in a complex study consisting of a range of experimental techniques (various junction spectroscopy methods, photoluminescence, microwave detected photo-conductance decays) and *ab-initio* modelling [26, 81], where it was concluded that B_sO_2 has negative-U properties.

Figure 11 shows the configuration-coordinate diagram and atomic configurations of the B_sO_2 complex in different charge states. The negative-U properties of the complex are related to the transformations of the oxygen dimer in the vicinity of a substitutional boron atom between the so called 'squared' configuration (D and X in Figure 11,

Figure 11. Configuration coordinate diagram and atomic configurations of the B_sO_2 complex in Si on different charge states. (Reprinted from Ref. [26], with the permission of AIP Publishing).

with two three-fold coordinated oxygen atoms) and 'staggered' configurations (configurations A, A$'$, and A$''$ in Figure 11, with two-fold coordinated oxygen atoms). In p-type Si crystals when $E_F < E_v + 0.3$ eV, the B_sO_2 centre is in the positive charge state (state D$^+$ in Figure 11). In this state, the defect can either capture an electron (created by either light or forward bias pulses) or emit a hole (thermally when the temperature is high enough) and transform to the shallow acceptor A$^-$ state according to the following sequence of transitions: D$^+$ + h$^+$ + e$^-$ $\rightleftarrows$ X^0 + h$^+$ $\rightleftarrows$ A^0 + h$^+$ $\rightleftarrows$ A$^-$ + 2h$^+$. Kinetics of the forward and back transitions between the D$^+$ and A$^-$ have been monitored at different temperatures in the p-type Si crystals with different hole concentrations and analysed with the use of Equations 49, 50 and 51 [26, 81]. From this analysis, the values of energy differences and barriers between different configurations of the complex have been derived.

Further, it has been found that upon relatively long (tens of hours at room temperature) minority-
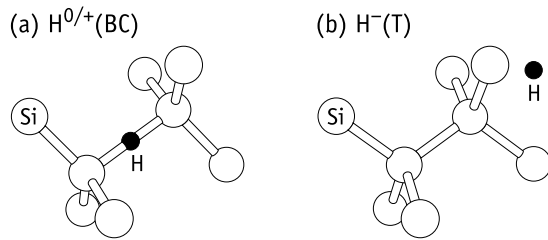
Figure 12. Illustration of the minimum-energy atomic configurations for hydrogen interstitials in silicon, as derived from first-principles calculations (a) the positive charge and neutral charge states, and (b) the negative charge state.

carrier-injection treatments the B_sO_2 defect transforms to a metastable configuration with a shallow acceptor level and stronger recombination activity (than that for B_s and the A state of the B_sO_2 complex). A possible mechanism of this transformation has been considered in Ref. [26]. *Ab-initio* modelling results have shown the existence of two metastable shallow acceptor configurations (A$'$ and A$''$ configurations in Fig. 11) of the B_sO_2 complex [26]. So, it appears that the B_sO_2 complex is a multi-stable defect with at least three experimentally confirmed atomic configurations for the neutral charge state (X^0, A^0, and A$'^0$/A$''^0$). Because of the existence of a rather high energy barrier between the A$'^0$/A$''^0$ and A^0 configurations the metastable A$'^0$/A$''^0$ configurations are long-living at room temperature. Only at temperatures above 150 °C the A$'$/A$'' \to$ A/(D) reactions occur efficiently.

5.2. Atomic hydrogen in silicon

Hydrogen is a very common impurity in silicon. It is quite mobile at room temperature and extremely reactive in its atomic form. As a result, it readily combines with other impurities and itself, leaving only a small fraction of isolated hydrogen atoms. These exist in three charge states and exhibit negative-U behaviour with the H-related $(0/+)$ donor level lying above the $(-/0)$ acceptor level, where neutral H^0 is a metastable species [106]. *Ab-initio* calculations predict that both H^0 and H$^+$ have minimum energy positions at the Si-Si bond centre (BC), while H$^-$ is located at an interstitial tetrahedral (T) lattice site [107]. Representative structures are shown in Figure 12. The resultant negative-U characteristic of isolated hydrogen is technologically very significant, impacting on hydrogen's ability to passivate defects and its diffusivity. Several comprehensive reviews have been published which include the negative-U aspects of hydrogen in silicon [22, 108, 109, 110].

The experimental study of isolated hydrogen in silicon has proved to be very difficult and it is only by the application of multiple techniques and calculations that a good understanding has been achieved. The fundamental problem is that hydrogen is both highly reactive and very mobile. The end result has been that some experiments claiming to characterise isolated hydrogen actually derive from hydrogen complexes, commonly H-O and H-C. A methodology to avoid this, which has met with success, is to implant protons at low temperature and undertake measurements at low temperature. The technique was first used by Stein [111] who showed unambiguously that the line at 1998 cm^{-1} seen in LVM absorption studies was related to hydrogen at the bond centre position. Low temperature implantation at $T = 45$ K was used by Holm [112] to study n-type Si using DLTS, and an $E_c - E(0/+) = 0.16$ eV donor level (E3$'$) was observed, being later ascribed to bond-centred H(BC) . This was followed by detailed studies by Bonde-Nielsen *et al.* [113] on hydrogen near the T site, which quantified the impact of oxygen on nearby sites and determined the energy of the hydrogen acceptor to be $E_c - E(-/0) = 0.68$ eV.

5.3. DX centres in III-V alloys

DX centres are a special case of negative-U systems in which the addition of dopants, expected to form shallow donors in III-V alloys, gives rise to persistent photoconductivity and deep acceptors. The phenomenon was first reported in detail by Craford *et al.* in 1968 [114] describing Hall effect measurements of GaAs$_{1-x}$P$_x$ with $x \approx 0.3$ doped with Te or S. The phenomenon was erroneously believed to be the result of a donor 'D' reacting with an unknown defect 'X' to produce a deep state (hence 'DX'). This was a result of studies of Al$_x$Ga$_{1-x}$As with $x > 0.22$ doped with Si, Ge, Sn, S, Se or Te using Hall, photo-capacitance and DLTS (see for example Refs. [16, 17]). The story of the DX centres is a classic research history spanning well over 25 years starting with a technological driver of achieving adequately high doping in III-V heterojunction lasers and transistors and involving some of the world's leading semiconductor groups combining theory and experiment. After many plausible explanations were explored it is now evident that the DX centre is a classic example of negative-U behaviour. The saga up to 1990 has been reviewed by Mooney [115] and some key issues are covered in the introductory section of the current paper.

The most extensively studied DX centre is that of Si in Al$_x$Ga$_{1-x}$As. DX behaviour is observed for the case where $x > 0.22$ or at lower aluminium concentrations under hydrostatic pressure, for example DX behaviour of Si in GaAs is observed at pressures $P > 2$ GPa. In the DX configuration the silicon accepts an electron and is negatively charged. The charge state has been determined experimentally by
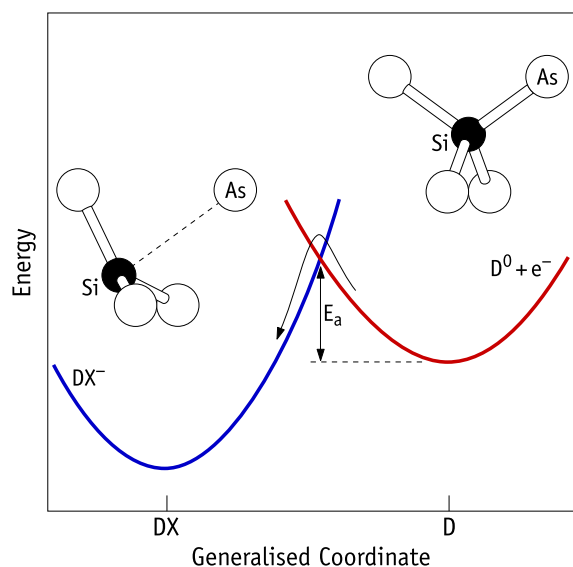
Figure 13. Configuration coordinate diagram of the Si *DX* centre in Al$_x$Ga$_{1-x}$As. In the *DX* configuration (left), the Si atom is displaced from its substitutional site. In the shallow-donor configuration (right), the Si atom occupies the substitutional site.

Dobaczewski et al. [19] from direct measurement of the capture cross section. As shown in Figure 13, this change of charge state results in the relaxation of the Si from its substitutional position on the group III site, to an interstitial site neighbouring three As atoms. In this siting, the Si does not act as a conventional thermally ionised donor and does not contribute to the conductivity.

If the system is exposed to light of a sufficiently high energy, electrons are excited from the Si to the conduction band and the *DX* centre reverts to the substitutional site where it is a metastable neutral donor. At temperatures $T < 180$ K the electron does not have enough energy to surmount the barrier ($E_a \sim 0.2$ eV) shown by the arrow in Fig. 13, and so the additional conductivity resulting from the shallow donor state induced by the light remains for hours or at lower temperatures for days or even weeks. This bizarre behaviour is referred to as persistent photoconductivity and is a signature of *DX* behaviour.

The identification of the substitutional-interstitial motion evolved from a sequence of calculations and experiments discussed in Mooney's review [115]. However, a series of revealing experiments published subsequently provided unambiguous detail of this behaviour [116, 117]. In this work, high resolution Laplace DLTS was analysed to provide detail of the influence of the local environment on the electron thermal emission process from *DX* centres in a range of compositions of Al$_x$Ga$_{1-x}$As. The direct comparison of this process for centres related to Si which can

replace gallium or aluminium, with that observed for a group-VI donor (tellurium), which resides in the arsenic sublattice, enabled the configuration of atoms to be determined when the centre is in the ground state. This shows that the *DX*(Si) defect ionisation process is associated with interstitial-substitutional motion of the silicon atom. However, in the case of the *DX*(Te) centre, the spectra shows peaks from two groups which are attributed to interstitial substitutional motion of the neighbouring aluminium or gallium atoms, respectively.

DX behaviour has been observed in many III-V and II-VI semiconductors, and recently, persistent photoconductivity has been reported in oxides (*e.g.* ZnO and SrTiO$_3$), which has been interpreted as *DX* behaviour associated with negative-*U* characteristics [118, 119].

5.4. The Ga-O$_{As}$-Ga defect in GaAs

The Ga-O$_{As}$-Ga defect in GaAs is often referred to as substitutional oxygen (on the arsenic site) or arsenic-vacancy-oxygen (V_{As}-O) complex [120]. It shows up as a group of three local vibrational mode (LVM) absorption bands with identical fine structure. The bands, labelled A, B$'$ and B, appear at 731 cm^{-1}, 714 cm^{-1} and 715 cm^{-1}. Band B$'$ emerges during the intermediate stages of optical-induced conversion from A to B using light with $h\nu > 0.8$ eV (below band-gap light). This was interpreted as an electron-transfer process from EL2 to the conduction band, and subsequent capture by the Ga-O$_{As}$-Ga complex. EL2 is an As antisite-related defect, responsible for a pinning $E(0/+)$ level, effectively working as a reservoir of electrons at $E_c - 0.75$ eV [121]. Hence, A, B$'$ and B bands were ascribed to LVMs of the same complex in charge states q, $q - 1$ and $q - 2$, respectively [122].

The A and B bands were first reported in semi-insulating material and assigned to EL2 by Song *et al.* [123]. However, this connection could not be correct and was quickly supplanted by a model involving one O atom with two equivalent Ga-O bonds [124]. That way, the triple-peak fine structure with intensity ratio of about 2:3:1 on each band (from high to low frequency peaks), could be explained by a combination of natural occurring Ga isotopes. Considering that ^{69}Ga and ^{71}Ga are found in a proportion of about 3:2, a natural population of ^{69}Ga-X-^{69}Ga, ^{69}Ga-X-^{71}Ga, and ^{71}Ga-X-^{71}Ga units corresponds to an intensity ratio of 9 : (2×6) : 4 = 2.25 : 3 : 1, respectively [124, 125]. Based on sample history arguments and semi-empirical calculations, the X element was suggested to be oxygen [124]. This was confirmed shortly after with the observation of a ^{18}O-isotope shift of the bands by Schneider and co-workers [125].

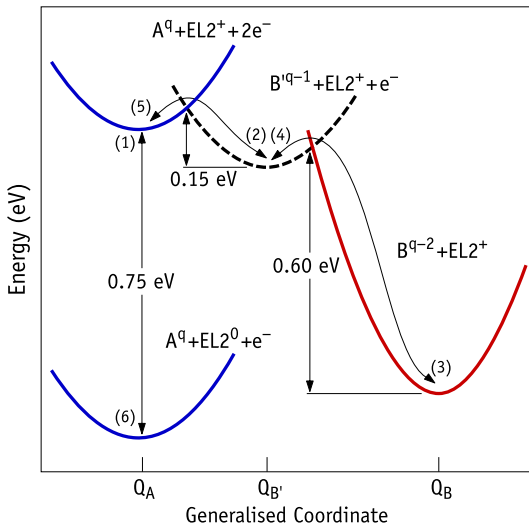Besides possessing a Ga-O-Ga unit with two

Figure 14. Configuration coordinate diagram for the Ga-O-Ga defect in GaAs based on experimental observations. Arrow heads indicate electron capture and emission processes described in the text.

equivalent Ga-O bonds, we know from stress-splitting FTIR absorption that the A band involves a ⟨110⟩-oriented vibration dipole [126] and the defect should have C_{2v} symmetry. These experiments also concluded that Ga-O_{As}-Ga does not reorient like VO in Si.

The negative-U nature of the Ga-O_{As}-Ga defect was found by Alt, who performed a series of enlightening experiments, combining Fourier-transform infrared (FTIR) absorption, photo-excitation using monochromatic light with energy spanning the band gap and annealing [122, 127]. The configuration coordinate diagram presented in Figure 14 was constructed from the experimental data. Upon cooling the sample from room temperature to 10 K in dark, only band A was observed, indicating that this should be the ground state in semi-insulating GaAs (for $E_F = E_c - 0.75$ eV). State A is here assumed to have a reference charge state q, thus being referred to as A^q.

After applying a short pulse of light with 0.8 eV $< h\nu < 1.5$ eV, band A was partially quenched and that corresponded to a proportional growth of band B′. As referred to above, this was assigned to the photoionisation of EL2 followed by electron capture by Ga-O_{As}-Ga. Importantly, band B was not visible at this point. This process is indicated in Figure 14 by the arrow heads with labels (1) and (2). The B′ band was therefore assigned to the $q-1$ charge state (B^{q-1}).

Applying a 95 K dark annealing treatment to samples where only A and B′ bands were present, led to the quenching of B′ and to the concurrent growth of bands A and B. The argument was that by warming the sample at 95 K, electrons bound to B'^{q-1} were thermally promoted into the conduction band, thus forming A^q again (process with arrow head (5) in Figure 14). The relatively small capture cross section of ionized EL2, also allowed free electrons to be captured by other B'^{q-1} defects, thus leading to the growth of band B. This was ascribed to the formation of a $q-2$ charge state, namely B^{q-2} (process number (3) in Figure 14). The activation energy for this conversion was measured as $E_a = 0.15$ eV, and was assigned to the thermal emission barrier of $B'^{q-1} \rightarrow A^q + e^-$ [122, 128]. As pointed out by Alt [122], process (3) and (5) consisted on a disproportionation of metastable B'^{q-1} with negative-U ordering of levels. It was also noted that full A → B band conversion could be achieved by subjecting the samples for long illumination times with 0.8 eV $< h\nu < 1.5$ eV at $T < 70$ K. At these temperatures, the thermal emission channel (5) in Figure 14 was suppressed and all defects ended up in state B^{q-2}.

At higher temperatures ($T \geq 180$ K), band B starts to decrease in intensity at a rate identical to the recovery of band A. Full recovery of band A is achieved when B vanishes. This was suggested to result from a two-electron emission sequence $B^{q-2} \rightarrow B'^{q-1} + e^- \rightarrow A^q + 2e^-$, where the limiting step is the first one (with larger barrier). The above emission sequence is represented in Figure 14 as steps with arrow heads (4) and (5), respectively. At these temperatures, the thermally emitted electrons were shown to be captured and trapped at the deeeper EL2 state (step (6) of Figure 14). The activation energy for the B → A conversion was measured as $E_a = 0.60$ eV, thus suggesting that the binding energy of the second electron captured by the Ga-O_{As}-Ga defect, exceeded the binding energy first one (0.15 eV).

We note that a more cautious analysis of the above picture would include the capture barriers for steps (2) and (3). DLTS measurements give activation energies of 0.58 eV and 0.14 eV for electron emission from B^{q-2} B'^{q-1}, respectively [129]. As far as we are aware, capture barriers were not measured. However, judging from Hall effect measurements, which find a disproportionating $(q-2/q)$ transition at $E_c - 0.43$ eV, and considering the above emission barriers, the sum of first and second capture barriers should be in the meV range [130, 127]. Also interestingly, based on its DLTS fingerprint, electron emission from B^{q-2} was assigned to EL3, a previously known deep trap with unknown composition [129].

The first atomistic model claiming to account for the above data was a ⟨001⟩-displaced substitutional oxygen atom at the arsenic site (O_{As}) [125], resembling the vacancy-oxygen complex in silicon. O_{As} possesses a Ga-O-Ga unit aligned along the ⟨110⟩ direction, thus being compliant with the conclusions from the polarised absorption experiments [126]. Jones and

Negative-U defects in semiconductors   25

Öberg [131] inspected the O_{As} defect in H-terminated GaAs spherical clusters using local density functional theory. The paramagnetic $O_{As}^=$ state was found to be metastable and was ascribed to band B′. Bands A and B were therefore connected to diamagnetic states with O_{As}^- and O_{As}^{-3}, respectively. While the calculations of Ref. [131] accounted for the experimentally observed red-shift of the oxygen vibrational frequency of B with respect to A, this result could not be reproduced by subsequent calculations [132, 133, 134]. Additionally, and perhaps more problematic for the O_{As} model, is that it cannot explain the lack of reorientation of the defect under the effect of uniaxial stress (although this was speculated to be due to a nearly isotropic stress tensor [126]).

Alternatively, Pesola *et al.* [134] proposed a defect model in which the two Ga radicals (not connected to oxygen) in the off-site O_{As} defect are replaced by As. This complex was referred to as $(As_{Ga})_2$-O_{As} and incorporates a Ga-O-Ga unit like in O_{As}. The first-principles calculations of Ref. [134] indicated that $(As_{Ga})_2$-O_{As} shows a negative-U ordering of $E(0/+)$ and $E(-/0)$ transitions, and the calculated O-LVM frequencies of charge states $+$, 0 and $-$ closely follow the observed frequency variations of bands A, B′ and B, respectively. The model naturally accounts for the lack of reorientation under uniaxial stress. However, it clashes with optically detected electron-nuclear double resonance (OD-ENDOR), which reported that as much as 60% of spin-density of the $B′^0$ state (with spin-1/2) is localized on the first shell of Ga atoms [135] – Figure 2 of Ref. [134] clearly shows that the $(As_{Ga})_2$-O_{As} model has most of the density on As shells and virtually none on Ga shells.

The origin for the negative-U ordering of levels in the Ga-O_{As}-Ga complex is not clear. The simpler O_{As} substitution, although incompatible with the vibrational absorption data, could explain the effect based on strong charge-dependent Ga-Ga reconstructions. This feature was found theoretically for the As vacancy (V_{As}) in GaAs [136], and used as a justification for V_{As}^0 and $V_{As}^=$ showing a negative-U. On the other hand, the $(As_{Ga})_2$-O_{As} model has a mid-gap state localised on two As radicals, which were suggested to induce strong forces on neighbouring atoms as a function of occupancy [134]. This effect does not find a parallel in the (positive-U) Ga vacancy (V_{As}), whose As radicals induce weak deformations [136].

Recently, the O_{As} defect was investigated using modern hybrid density functional theory [137]. O_{As}^+ and O_{As}^0 were found to be most stable at the perfect tetrahedral site. The highest occupied state of O_{As}^0, which is paramagnetic, is an s-like orbital on the oxygen atom, and therefore not compliant with the OD-ENDOR data. O_{As}^- had an off-site Ga-O_{As}-Ga

configuration with a localised acceptor state on a reconstructed Ga-Ga pair. The $E(0/+)$ and $E(-/0)$ levels were calculated at $E_c - 0.71$ eV and $E_c - 0.54$ eV, thus showing a positive-U ordering. Clearly the Ga-O_{As}-Ga defect should be revisited by both theory and experiments.

5.5. <mark>Electrical levels</mark> of selected negative-U defects in semiconductors

Table 1 lists the electronic characteristics of a selection of negative-U defects in semiconductors. It should be noted that the table is not comprehensive. We only considered defects for which the negative-U behaviour is confirmed by solid experimental results, where their electronic properties are relatively well understood and the atomic structures in different charge states have been predicted by experiments and *ab-initio* modelling. Negative-U properties have been suggested for a number of other defects in semiconductor materials, *e.g.*, for interstitial boron-interstitial oxygen complex and hydrogen-related shallow donors in silicon [138, 76]. Besides silicon, atomic hydrogen has also been claimed to show negative-U behaviour in many other semiconductors [139, 140]. Further examples could be enumerated. However, no solid experimental evidence confirming these suggestions has been presented yet. Perhaps it is worth mentioning the interesting case of the Si trivacancy, which has been briefly discussed in Section 4.2.3. It has been argued in Refs. [102, 101] that V_3 in Si can be considered as a defect with negative-U properties as it has the $E(=/-)$ acceptor level at $E_c - 0.21$ eV below the $E(-/0)$ level at $E_c - 0.07$ eV. However, the levels mentioned above are energy levels of V_3 in different configurations (part of the hexagonal ring and four-fold coordinated shown in Figures 10(a) and 10(d), respectively), which are separated by rather high energy barriers. Because of these barriers the paired emission or capture of electrons (a signature of negative-U defects) has not been observed for the V_3 defect, and therefore, it has not been included into the table.

The transition levels of Table 1 are cast as $E(q/q+1)$ and $E(q-1/q)$, where q is a reference charge state. For negative-U defects $E(q/q+1) > E(q-1/q)$, and the q charge state corresponds to the metastable state which, under equilibrium conditions, disproportionates into stable $q+1$ and $q-1$ charge states.

Negative-U defects in semiconductors   26

Table 1. Electrical levels of selected negative-U defects in semiconductors and respective key references. All energies in eV. Starred (*) defects are discussed in detail either below or elsewhere in this paper.

Host	Defect	$q/q+1$	$E(q/q+1)$	$q-1/q$	$E(q-1/q)$	Comment	Key references
Si	Isolated vacancy (*)	+/++	$E_v + 0.13$	0/+	$E_v + 0.05$	DLTS, EPR, p-type, low temperature irradiation	[5, 6]
Si	Interstitial B (*)	0/+	$E_c - 0.13$	−/0	$E_c + 0.37$	DLTS, ODLTS, EPR, p- and n-type, n+p & p+n diodes, low temperature irradiation	[9, 10, 11]
Si	Hydrogen-Oxygen	0/+	$E_c - 0.16$	−/0	$E_c - 0.79$	DLTS, n-type Cz, low temperature proton implantation	[113]
Si	Hydrogen (*)	0/+	$E_c - 0.16$	−/0	$E_c - 0.68$	C transients, PH dissociation, EPR, n-type, low temperature proton implantation	[106]
Si	B_sO_2 (*)	0/+	$E_v + 0.56$	−/0	$E_v + 0.04$	DLTS, LDLTS, admittance spectroscopy, p type CZ, heat-treatments at 400-500 °C	[26, 81]
Si	C_iO_iH	0/+	$E_c - 0.043$	−/0	$E_c - 0.11$	Hall effect, DLTS, FT-IR, n-type Cz, heat-treatments 350 °C electron irradiated hydrogenated samples	[141, 80, 142]
Si	Self-interstitial (I)	+/++	$E_c - 0.39$	0/+	?	DLTS, EPR, p-type, low temperature implantation with alpha particles	[143]
Si	IO_2 (*)	+/++	$E_v + 0.36$	0/+	$E_v + 0.12$	Hall effect, DLTS, FT-IR, carbon lean p-type Cz, room temperature electron irradiation	[144, 64, 77]
Si	O-related TDDs	+/++	$E_c - 0.16$	0/+ 0/+ 0/+	TDD-0, $E_c - 0.75$ / TDD-1, $E_c - 0.48$ / TDD-2, $E_c - 0.29$	Hall effect, photoconductivity, FT-IR, n-type Cz, heat treatments 400-500 °C	[63, 79, 20]
Ge	O-related TDDs	+/++	$E_c - 0.04$	0/+ 0/+ 0/+	TDD-0, $E_c - 0.48$ / TDD-1, $E_c - 0.34$ / TDD-2, $E_c - 0.26$	Hall effect, photoconductivity, FT-IR, n type Cz, heat treatments 300-400 °C	[21, 145]
4H-SiC	C vacancy (*)	−/0	$E_c - 0.42$	=/−	$E_c - 0.64$	DLTS, Laplace-DLTS, EPR, n-type, traces in all as-grown samples	[24, 54, 53]
6H-SiC	C vacancy (*)	−/0	$E_c - 0.20$	=/−	$E_c - 0.43$	DLTS, Laplace-DLTS, n-type, traces in all as-grown samples	[25, 59]
InP	M-center	+/++	$E_c - 0.18$	0/+	$E_c - 0.42$	TSCAP, DLTS, optical ionisation, irradiated samples	[12, 13, 14, 15]
GaAs	Ga-O$_{As}$-Ga (*)	0/+	$E_c - 0.14$	−/0	$E_c - 0.58$	DLTS, optical ionisation	[122, 129, 134]
GaAsP	DX centre	See text, properties depend on alloy composition				Hall effect, optical absorption	[114]
AlGaAs	DX centre	See text, properties depend on alloy composition				DLTS, photo-capacitance, Hall effect	[146, 17, 18, 19, 147]
ZnO	Oxygen vacancy	+/++	$E_c - 0.4$	0/+	$E_c - 1.6$	DLTS, theory	[148, 149]
GaSb	DX(S)	0/+	$E_c - 0.22$	−/0	$E_c - 0.25$	DLTS	[19]

Negative-U defects in semiconductors　　　　　　　　　　　　　　　　　　　　　　27

6. Conclusions

The present review begins with an historical perspective of the topic of 'negative-U defects in semiconductors'. We start from the early 1980's, where the first reports of this class of defects emerged in the literature. The account includes several difficulties identified at the time, mostly due to meta-stability, or in some cases multi-stability, of one of the three charge states involved. We provide the reader with an introduction to the essential physics of the workings of negative-U defects (electronic correlation, electron-phonon coupling, disproportionation and defect transition levels). We present the state-of-the-art regarding the experimental and theoretical methods that are currently used for the characterisation of negative-U defects. These subjects are presented in detail and they are accompanied with several successful studies from the literature. This includes experimental methods based either on defect thermodynamics or defect kinetics (monitored by several techniques, like Hall effect, DLTS or optical absorption), as well as numerical methods that can help us in the construction of detailed configuration coordinate diagrams (with the calculation of transition levels, transition rates and transformation barriers). To wrap up, we showcase a few experimental and theoretical analyses of some of the most impacting negative-U defects in group-IV and compound semiconductors, leaving a summary of the characteristics of many others, along with the most relevant references in a tabular format.

A negative-U defect has a metastable state with N electrons that disproportionate into $N-1$ and $N+1$ species under equilibrium conditions. This is only possible in many-atom systems, and thanks to the conversion of electron correlation energy into bond formation energy, either during ionisation or capture of carriers. In these defects, the stable $N-1$ and $N+1$ states are normally diamagnetic close-shell systems, making them unnoticeable by paramagnetic resonance techniques, unless some sort of excitation is provided to the sample. The metastability of the N-electron state works like a barrier separating the $N-1$ and $N+1$ species. When one of the latter is a shallow donor or a shallow acceptor, their occupation via deep-shallow excitation induced by illumination, usually results in a concentration increase of free carriers. This effect is often measurable as a persistent photoconductivity, which depending on the temperature, may last for several hours, until the defects surmounts a barrier associated with a charge capture event followed by transformation and return to the ground state configuration.

The above features make junction spectroscopy methods a primary choice for the characterisation of negative-U defects in semiconductors. However, the complexity of their behaviour often requires the usage of complementary techniques, as well as special tricks, like illumination or thermal quenching, which allow researchers to play with thermodynamics and kinetics, and by doing so, to isolate the defect in one of its three states, or to investigate the conversions between them. Defect modelling has also been pivotal. Since the pioneering work on the silicon vacancy, theories, numerical algorithms and hardware have evolved dramatically. Density functional related methods are perhaps the most successful cases. These have been able to provide us with a detailed quantum mechanical description of the interactions between trapped carriers and the many-body cloud of electrons comprising the defective semiconductor.

Finally, we would like to emphasise the following concluding remark – we have achieved a rather deep understanding of the electronic and atomistic structure of many negative-U defects in semiconductors. For sure, that was achieved thanks to concurrent and combined experimental and theoretical work.

ORCID iDs

J Coutinho https://orcid.org/0000-0003-0280-366X
V P Markevich https://orcid.org/0000-0002-2503-6144
A R Peaker https://orcid.org/0000-0001-7667-4624

All authors contributed equally to this work.

References

[1] Anderson P W 1975 *Physical Review Letters* **34** 953–955
[2] Hubbard J 1963 *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **276** 238–257
[3] Street R A and Mott N F 1975 *Physical Review Letters* **35** 1293–1296
[4] Kastner M, Adler D and Fritzsche H 1976 *Physical Review Letters* **37** 1504–1507
[5] Baraff G A, Kane E O and Schlüter M 1979 *Physical Review Letters* **43** 956–959
[6] Watkins G D and Troxell J R 1980 *Physical Review Letters* **44** 593–596
[7] Benton J L 1989 *Journal of Electronic Materials* **18** 199–206
[8] Langer J M 1983 *Radiation Effects* **72** 55–72
[9] Troxell J R and Watkins G D 1980 *Physical Review B* **22** 921–931
[10] Harris R D, Newton J L and Watkins G D 1982 *Physical Review Letters* **48** 1271–1274

Negative-U defects in semiconductors 28

[11] Harris R D, Newton J L and Watkins G D 1987 *Physical Review B* **36** 1094–1104

[12] Levinson M, Benton J L and Kimerling L C 1983 *Physical Review B* **27** 6216–6221

[13] Levinson M, Stavola M, Benton J L and Kimerling L C 1983 *Physical Review B* **28** 5848–5855

[14] Stavola M, Levinson M, Benton J L and Kimerling L C 1984 *Physical Review B* **30** 832–839

[15] Wager J F and Vechten J A V 1985 *Physical Review B* **32** 5251–5258

[16] Henry C H and Lang D V 1977 *Physical Review B* **15** 989–1016

[17] Lang D V, Logan R A and Jaros M 1979 *Physical Review B* **19** 1015–1030

[18] Chadi D J and Chang K J 1988 *Physical Review Letters* **61** 873–876

[19] Dobaczewski L and Kaczor P 1991 *Physical Review B* **44** 8621–8632

[20] Latushko Y A I, Makarenko L F, Markevich V P and Murin L I 1986 *physica status solidi (a)* **93** K181–K184

[21] Litvinov V V, Palchik G V and Urenev V I 1988 *physica status solidi (a)* **108** 311–321

[22] 1991 *Hydrogen in semiconductors* (*Semiconductors and Semimetals* vol 34) ed Pankove J I and Johnson N M (Boston: Academic Press) ISBN 978-0-12-752134-3

[23] Pearton S J, Corbett J W and Stavola M 1992 Hydrogen in crystalline semiconductors Springer Series in Materials Science (Berlin, Heidelberg: Springer) ISBN 978-3-540-55491-2

[24] Hemmingsson C G, Son N T, Ellison A, Zhang J and Janzén E 1998 *Physical Review B* **58** R10119–R10122

[25] Hemmingsson C G, Son N T and Janzén E 1999 *Applied Physics Letters* **74** 839–841

[26] Vaqueiro-Contreras M, Markevich V P, Coutinho J, Santos P, Crowe I F, Halsall M P, Hawkins I, Lastovskii S B, Murin L I and Peaker A R 2019 *Journal of Applied Physics* **125** 185704

[27] Levine I N 2014 *Quantum chemistry* Pearson Advanced Chemistry Series (Boston: Pearson Education) ISBN 0321803450

[28] Blinder S M 2019 *Mathematical Physics in Theoretical Chemistry* Developments in Physical & Theoretical Chemistry ed Blinder S M and House J E (Amsterdam: Elsevier) chap Chapter 1 - Introduction to the Hartree-Fock method, pp 1–30 ISBN 9780128136515

[29] Löwdin P O 1955 *Physical Review* **97** 1474–1489

[30] Szabo A and Ostlund N S 1996 *Modern quantum chemistry: introduction to advanced electronic structure theory* (Dover Publications Inc.) ISBN 0486691861

[31] Hohenberg P and Kohn W 1964 *Physical Review* **136** B864–B871

[32] Parr R G and Yang W 1989 *Density-functional theory of atoms and molecules* International Series of Monographs ion Chemistry (New York: Oxford University Press) ISBN 0195042794

[33] Burke K and Wagner L O 2012 *International Journal of Quantum Chemistry* **113** 96–101

[34] Kohn W and Sham L J 1965 *Physical Review* **140** A1133–A1138

[35] Dirac P A M 1930 *Mathematical Proceedings of the Cambridge Philosophical Society* **26** 376–385

[36] Chachiyo T 2016 *The Journal of Chemical Physics* **145** 021101

[37] Lieb E H and Wu F Y 2003 *Physica A: Statistical Mechanics and its Applications* **321** 1–27

[38] Perdew J P, Parr R G, Levy M and Balduz J L 1982 *Physical Review Letters* **49** 1691–1694

[39] Chaibi W, Peláez R J, Blondel C, Drag C and Delsart C 2010 *The European Physical Journal D* **58** 29–37

[40] Shaffer D W, Xie Y and Concepcion J J 2017 *Chemical Society Reviews* **46** 6170–6193

[41] Franck J and Dymond E G 1926 *Transactions of the Faraday Society* **21** 536–542

[42] Wilson H N and Bremner J G M 1948 *Quarterly Reviews, Chemical Society* **2** 1–24

[43] Kremer M L and Stein G 1959 *Transactions of the Faraday Society* **55** 959–973

[44] Giefers H, Porsch F and Wortmann G 2005 *Solid State Ionics* **176** 199–207

[45] Dalpian G M, Liu Q, Varignon J, Bibes M and Zunger A 2018 *Physical Review B* **98** 075135

[46] Shockley W and Last J T 1957 *Physical Review* **107** 392–396

[47] Look D C 1981 *Physical Review B* **24** 5852–5862

[48] Sah C T 1996 *Fundamentals of Solid State Electronics* (Singapore: World Scientific Publishing Company) ISBN 978-981-4366-25-0

[49] Todorova M and Neugebauer J 2014 *Physical Review Applied* **1** 014001

[50] Grabowski B, Ismer L, Hickel T and Neugebauer J 2009 *Physical Review B* **79** 134106

[51] Estreicher S K, Sanati M, West D and Ruymgaart F 2004 *Physical Review B* **70** 125209

[52] Coutinho J, Torres V J B, Demmouche K and Öberg S 2017 *Physical Review B* **96** 174105

[53] Capan I, Brodar T, Coutinho J, Ohshima T, Markevich V P and Peaker A R 2018 *Journal of Applied Physics* **124** 245701

[54] Son N T, Trinh X T, Løvlie L S, Svensson B G, Kawahara K, Suda J, Kimoto T, Umeda T, Isoya J, Makino T, Ohshima T and Janzén E 2012 *Physical Review Letters* **109** 187603

[55] Ayedh H M, Nipoti R, Hallén A and Svensson B G 2015 *Applied Physics Letters* **107** 252102

[56] Stoneham A M 1981 *Reports on Progress in Physics* **44** 1251–1295

[57] Peaker A R, Markevich V P and Coutinho J 2018 *Journal of Applied Physics* **123** 161559

[58] Stoneham A M 2001 *Theory of defects in solids: electronic structure of defects in insulators and semiconductors* (New York: Oxford University Press) ISBN 9780198507802

[59] Koizumi A, Markevich V P, Iwamoto N, Sasaki S, Ohshima T, Kojima K, Kimoto T, Uchida K, Nozaki S, Hamilton B and Peaker A R 2013 *Applied Physics Letters* **102** 032104

[60] Lang D V 1974 *Journal of Applied Physics* **45** 3023–3032

[61] Mukashev B N, Frolov V V and Kolodin L G 1982 *Physics Letters A* **91** 358–360

[62] Emtsev V V, Margaryan M A and Mashovets T V 1983 *Soviet Physics - Semiconductors* **19** 1192

[63] Makarenko L F, Markevich V P and Murin L I 1985 *Soviet Physics - Semiconductors* **19** 1192

[64] Markevich V P, Murin L I, Lastovskii S B, Medvedeva I F, Komarov B A, Lindström J L and Peaker A R 2005 *Journal of Physics: Condensed Matter* **17** S2331–S2340

[65] Markevich V P and Murin L I 1996 *Semiconductors* **30** 626

[66] Blakemore J S 1987 *Semiconductor statistics* (Mineola, New York: Dover Publications) ISBN 0486653625

[67] Hoffmann H J 1979 *Applied Physics* **19** 307–312

[68] Hoffmann H J 1982 *Applied Physics A Solids and Surfaces* **27** 39–47

[69] Clauws P 2007 Oxygen in germanium *Germanium-based technologies* ed Clayes C and Simoen E (Amsterdam: Elsevier) chap 4, pp 97–130 ISBN 9780080449531

[70] Stallinga P, Walukiewicz W, Weber E R, Becla P and Lagowski J 1995 *Physical Review B* **52** R8609–R8612

[71] Wruck D and Spiegelberg F 1986 *physica status solidi (b)* **133** K39–K43

Negative-U defects in semiconductors   29

[72] Wagner P and Hage J 1989 *Applied Physics A* **49** 123–138
[73] Murin L I, Markevich V P, Lindström J L and Kleverman M 2003 *Physica B: Condensed Matter* **340-342** 1046–1050
[74] Chantre A 1987 *Applied Physics Letters* **50** 1500–1502
[75] Chantre A 1989 *Applied Physics A* **48** 3–9
[76] Mukashev B N, Abdullin K A and Gorelkinskii Y V 2000 *Physics-Uspekhi* **43** 139–150
[77] Markevich V P, Murin L I, Lastovskii S B, Medvedeva I F, Lindström J L, Peaker A R, Coutinho J, Jones R, Torres V J B, Öberg S and Briddon P R 2005 *Solid State Phenomena* **108-109** 273–278
[78] Tkachev V D, Makarenko L F, Markevich V P and Murin L I 1984 *Soviet Physics - Semiconductors* **18** 324–328
[79] Makarenko L F and Murin L I 1988 *physica status solidi (b)* **145** 241–253
[80] Markevich V P, Murin L I, Sekiguchi T and Suezawa M 1997 *Materials Science Forum* **258-263** 217–222
[81] Markevich V P, Vaqueiro-Contreras M, Guzman J T D, Coutinho J, Santos P, Crowe I F, Halsall M P, Hawkins I, Lastovskii S B, Murin L I and Peaker A R 2019 *physica status solidi (a)* **216** 1900315
[82] Markevich V P, Vaqueiro-Contreras M, Lastovskii S B, Murin L I, Halsall M P and Peaker A R 2018 *Journal of Applied Physics* **124** 225703
[83] Drabold D A and Estreicher S K 2007 *Theory of defects in semiconductors* Topics in Applied Physics ed Drabold D A and Estreicher S K (Berlin, Heidelberg: Springer-Verlag)
[84] Komsa H P, Rantala T T and Pasquarello A 2012 *Physical Review B* **86** 045112
[85] Alkauskas A, Broqvist P and Pasquarello A 2008 *Physical Review Letters* **101** 046405
[86] Alkauskas A, Yan Q and de Walle C G V 2014 *Physical Review B* **90** 075202
[87] Barmparis G D, Puzyrev Y S, Zhang X G and Pantelides S T 2015 *Physical Review B* **92** 214111
[88] Schanovsky F, Gös W and Grasser T 2011 *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena* **29** 01A201
[89] Alkauskas A, Lyons J L, Steiauf D and de Walle C G V 2012 *Physical Review Letters* **109** 267401
[90] Manneback C 1951 *Physica* **17** 1001–1010
[91] Doktorov E V, Malkin I A and Man'ko V I 1977 *Journal of Molecular Spectroscopy* **64** 302–326
[92] Zapol B P 1982 *Chemical Physics Letters* **93** 549–552
[93] Schmidt P P 2010 *Molecular Physics* **108** 1513–1529
[94] Borrelli R and Peluso A 2003 *The Journal of Chemical Physics* **119** 8437–8448
[95] Krasikov D N, Scherbinin A V, Knizhnik A A, Vasiliev A N, Potapkin B V and Sommerer T J 2016 *Journal of Applied Physics* **119** 085706
[96] Flynn C P 1972 *Point defects and diffusion* International Series of Monographs on Physics (New York: Oxford University Press) ISBN 0198512600
[97] Henkelman G and Jónsson H 1999 *The Journal of Chemical Physics* **111** 7010–7022
[98] Malek R and Mousseau N 2000 *Physical Review E* **62** 7723–7728
[99] Henkelman G, Uberuaga B P and Jónsson H 2000 *The Journal of Chemical Physics* **113** 9901–9904
[100] Henkelman G and Jónsson H 2000 *The Journal of Chemical Physics* **113** 9978–9985
[101] Coutinho J, Markevich V P, Peaker A R, Hamilton B, Lastovskii S B, Murin L I, Svensson B J, Rayson M J and Briddon P R 2012 *Physical Review B* **86** 174101
[102] Markevich V P, Peaker A R, Lastovskii S B, Murin L I, Coutinho J, Torres V J B, Briddon P R, Dobaczewski L, Monakhov E V and Svensson B G 2009 *Physical Review B* **80** 235207
[103] Schmidt J and Bothe K 2004 *Physical Review B* **69** 024107
[104] Bothe K and Schmidt J 2006 *Journal of Applied Physics* **99** 013701
[105] Niewelt T, Schon J, Warta W, Glunz S W and Schubert M C 2017 *IEEE Journal of Photovoltaics* **7** 383–398
[106] Herring C, Johnson N M and de Walle C G V 2001 *Physical Review B* **64** 125209
[107] de Walle C G V, Denteneer P J H, Bar-Yam Y and Pantelides S T 1989 *Physical Review B* **39** 10791–10808
[108] de Walle C G V and Neugebauer J 2006 *Annual Review of Materials Research* **36** 179–198
[109] Peaker A R, Markevich V P and Dobaczewski L 2008 Hydrogen-related defects in silicon, germanium, and silicon-germanium alloys *Defects in microelectronic materials and devices* ed Fleetwood D M and Schrimpf R D (Boca Raton: CRC Press) chap 5, pp 27–50 ISBN 1420043765
[110] Estreicher S K, Stavola M and Weber J 2014 Hydrogen in Si and Ge *Silicon, germanium, and their alloys* ed Kissinger G and Pizzini S (Boca Raton: Taylor & Francis Ltd.) chap 7, pp 217–254 ISBN 9781466586659
[111] Stein H J 1975 *Journal of Electronic Materials* **4** 159–174
[112] Holm B, Nielsen K B and Nielsen B B 1991 *Physical Review Letters* **66** 2360–2363
[113] Nielsen K B, Dobaczewski L, Søgård S and Nielsen B B 2002 *Physical Review B* **65** 075205
[114] Craford M G, Stillman G E, Rossi J A and Holonyak N 1968 *Physical Review* **168** 867–882
[115] Mooney P M 1990 *Journal of Applied Physics* **67** R1–R26
[116] Dobaczewski L, Kaczor P, Missous M, Peaker A R and Żytkiewicz Z 1992 *Physical Review Letters* **68** 2508–2511
[117] Dobaczewski L, Kaczor P, Missous M, Peaker A R and Zytkiewicz Z R 1995 *Journal of Applied Physics* **78** 2468–2477
[118] Bhatt V, Kumar M, Kim J, Chung H J and Yun J H 2019 *Ceramics International* **45** 8561–8570
[119] Tarun M C, Selim F A and McCluskey M D 2013 *Physical Review Letters* **111** 187403
[120] Stavola M and Fowler W B 2018 *Journal of Applied Physics* **123** 161561
[121] Look D C and Fang Z Q 1999 *Solid-State Electronics* **43** 1317–1319
[122] Alt H C 1990 *Physical Review Letters* **65** 3421–3424
[123] Song C, Ge W, Jiang D and Hsu C 1987 *Applied Physics Letters* **50** 1666–1668
[124] Zhong X, Jiang D, Ge W and Song C 1988 *Applied Physics Letters* **52** 628–630
[125] Schneider J, Dischler B, Seelewind H, Mooney P M, Lagowski J, Matsui M, Beard D R and Newman R C 1989 *Applied Physics Letters* **54** 1442–1444
[126] Song C, Pajot B and Porte C 1990 *Physical Review B* **41** 12330–12333
[127] Alt H C 1991 *Semiconductor Science and Technology* **6** B121–B129
[128] Skowronski M, Neild S T and Kremer R E 1990 *Applied Physics Letters* **57** 902–904
[129] Kaufmann U, Klausmann E, Schneider J and Alt H C 1991 *Physical Review B* **43** 12106–12109
[130] Alt H C 1989 *Applied Physics Letters* **54** 1445–1447
[131] Jones R and Öberg S 1992 *Physical Review Letters* **69** 136–139
[132] Mattila T and Nieminen R M 1996 *Physical Review B* **54** 16676–16682
[133] Taguchi A and Kageshima H 1998 *Physical Review B* **57** R6779–R6782
[134] Pesola M, von Boehm J, Sammalkorpi V, Mattila T and Nieminen R M 1999 *Physical Review B* **60** R16267–R16270

Negative-U defects in semiconductors 30

[135] Koschnick F K, Linde M, Pinheiro M V B and Spaeth J M 1997 *Physical Review B* **56** 10221–10227

[136] El-Mellouhi F and Mousseau N 2005 *Physical Review B* **71** 125207

[137] Colleoni D and Pasquarello A 2014 *Applied Surface Science* **291** 6–10

[138] Makarenko L F, Lastovskii S B, Yakushevich H S, Moll M and Pintilie I 2014 *physica status solidi (a)* **211** 2558–2562

[139] de Walle C G V and Neugebauer J 2003 *Nature* **423** 626–628

[140] de Walle C G V and Neugebauer J 2006 *Annual Review of Materials Research* **36** 179–198

[141] Markevich V P, Suezawa M, Sumino K and Murin L I 1994 *Journal of Applied Physics* **76** 7347–7350

[142] Coutinho J, Jones R, Briddon P R, Öberg S, Murin L I, Markevich V P and Lindström J L 2001 *Physical Review B* **65** 014109

[143] Gorelkinskii Y V, Abdullin K A, Mukashev B N and Turmagambetov T S 2009 *Physica B: Condensed Matter* **404** 4579–4582

[144] Lindström J L, Hallberg T, Hermansson J, Murin L I, Komarov B A, Markevich V P, Kleverman M and Svensson B G 2001 *Physica B: Condensed Matter* **308-310** 284–289

[145] Clauws P, Callens F, Maes F, Vennik J and Boesman E 1991 *Physical Review B* **44** 3665–3672

[146] Lang D V and Logan R A 1977 *Physical Review Letters* **39** 635–639

[147] Baraff G A 1992 *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* **341** 195–202

[148] Alkauskas A and Pasquarello A 2011 *Physical Review B* **84** 125206

[149] Hupfer A, Bhoodoo C, Vines L and Svensson B G 2016 *Journal of Applied Physics* **119** 181506