INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **14** (2002) 2637–2645 PII: S0953-8984(02)30267-4

Charged point defects in semiconductors and the supercell approximation

J Lento, J-L Mozos and R M Nieminen

Laboratory of Physics, Helsinki University of Technology, PO Box 1100, FIN-02015 HUT, Finland

E-mail: Juha.Lento@hut.fi

Received 25 October 2001
Published 18 March 2002
Online at stacks.iop.org/JPhysCM/14/2637

Abstract

The effects of the supercell approximation in first-principles calculations for isolated, charged point defects in semiconductors and insulators are studied. The convergence of the Coulomb energy with respect to the supercell size is investigated. Quantitative numerical results for the standard uniform compensating charge and the newly proposed localized compensating charge scheme are presented for a prototypical defect, the doubly positive silicon self-interstitial.

1. Introduction

Many of the calculations of point defects in solids are performed using powerful computational methods originally developed for the calculation of perfect crystal structures. These methods rely on the use of the periodic boundary conditions (PBCs) and the plane-wave expansion of the Kohn–Sham orbitals and the electron density. The defect properties are calculated using periodically repeated supercells, which contain up to a few hundred atoms using first-principles density-functional methods and present-day computational resources. Although the supercell approximation describes accurately the crucial local rearrangements of bonding between atoms and the underlying crystal structure, it also introduces artificial long-range interactions between the periodic defect images. The most dramatic artefact of the approximation is the divergence of the Coulomb energy for charged defects.

The Coulomb divergence of the charged defects in the supercell approximation is traditionally circumvented with the use of a uniform, neutralizing 'jellium' background charge. In the large-supercell limit the interaction of the defect with the spurious periodic images and with the jellium background becomes negligible, in principle. However, there is no guarantee that the convergence of the Coulomb energy as a function of the linear dimension of the supercell, defined as the cube root of the supercell volume $L = \sqrt[3]{V}$, is particularly fast. In fact, classically one would expect an asymptotic L^{-1} dependence.

Leslie and Gillan [1] used a macroscopic approximation to correct for the Coulomb energy error in the case of jellium compensation. They considered an array of point charges with a neutralizing background, immersed in a structureless dielectric, and found significant contributions in their calculations for ionic crystals. Makov and Payne [2] developed the idea further, and derived a total-energy correction formula for charge distributions in cubic supercells (see equation (3) below).

A different route was recently taken by Schultz, which also utilizes the linearity of the Poisson equation [3, 4]. An *aperiodic* model defect charge distribution $n_{\mathrm{LM}}(\boldsymbol{r})$, matching the electrostatic moments of the system up to a given order, is separated from the supercell charge. The Coulomb energy of the remaining periodic charge is neutral and momentless, and can be calculated normally using PBC. The Coulomb energy of the model defect charge is calculated using 'cluster boundary conditions', with the surrounding polarizable defect-free crystal replaced by perfect non-polarizable bulk crystal. Unfortunately, no analytical correction formula similar to the Makov–Payne one (equation (3)) is available for this local-moment countercharge (LMCC) method.

A direct approach for the embedding of the charged defect cell into a perfect non-polarizable crystal was given by Nozaki and Itoh [5]. Unfortunately the application of their formalism is limited by the missing treatment of screening and polarization effects. They report test calculations of impurity energies independent of supercell sizes. The acquired size independence does not stem from the derived formalism, but instead from the used numerical method, which does not allow any ionic or electronic screening effects (rigid-ion model and identical atomic positions for different supercells).

To summarize the current status, one can say that for systems isolated by large regions of vacuum, i.e. charged ions or molecules, all the above corrections to the Coulomb energy give excellent, consistent results. For these systems, there is no difficulty in defining the localized, aperiodic defect charge, and the surrounding vacuum does not require any extra considerations about polarizability or potential discontinuities. For metals the physical picture is that the charged defects are completely screened already at short distances from the defect, and the Coulomb defect–defect interaction between supercells is negligible. The usual jellium-compensation scheme is consistent with metallic screening without additional corrections.

Most of the work in Coulomb corrections for supercell approximation use ionic solids as test cases. The effects of the corrections for the jellium compensation are more dramatic in absolute magnitude for ionic materials than for semiconductors due to the stronger screening in semiconductors. In this work we show that the supercell approximation can introduce significant errors in the calculated total energies for charged defects in semiconductors. The errors in the calculated total energies transfer directly to the important formation and ionization energy estimates.

The silicon self-interstitial in the doubly positive charge state is chosen as a prototypical charged defect in covalently bonded semiconductor material. The stable tetrahedral position of the interstitial in the region of low electronic bulk density results in a relatively weak interaction with the host lattice [6]. The defect states are well localized and the symmetry conserving lattice relaxation around the tetrahedral interstitial position is relatively small. Energetically the occupied defect state is slightly below the bottom of the valence band.

2. Computational details

The calculations are performed using a standard plane-wave pseudopotential program within the local-density approximation (LDA) for the exchange–correlation potential [7]. The electron charge density is explicitly symmetrized using cubic point-group symmetries.

The plane-wave basis-set cutoff is 17 Ryd, and with the separable first-principles norm-conserving pseudopotentials we find the equilibrium lattice constant of 5.39 Å. Cubic supercells allow us to simplify the discussion, as dipole interactions are cancelled.

The effects of the first Brillouin-zone (BZ) sampling and the elastic interactions are thoroughly discussed by Puska *et al* [7] in the case of the silicon vacancy, and also give a reference to the magnitude of these effects in the case of the interstitial. In the present calculations for the interstitial, the ions are kept fixed on their ideal positions. The order of magnitude of the ionic relaxation energy associated with the breathing mode of the ions around the doubly positive interstitial remains relatively constant, 0.5 eV, for the studied supercell sizes.

The first BZ is sampled using the Monkhorst–Pack (MP) sampling scheme [8] and the zero-nearest-neighbour interaction scheme by Makov, Shah and Payne (MSP) [9]. The MSP scheme is found to perform consistently and accurately from supercell sizes of 64 atoms onwards compared with the $4 \times 4 \times 4$ MP sampling for 32-, 54- and 64-atom supercells, and the $2 \times 2 \times 2$ MP sampling for 128- and 216-atom supercells (figure 1). The 'Γ-point only' sampling was found to be totally unacceptable for the accuracy we require in the extrapolation as a function of the supercell size. The errors in the defect energy compared with the MP sampling are of the order of 1 eV for small supercells, and 0.5 and 0.1 eV for 128- and 216-atom supercells, respectively. We also stress that the errors caused by inaccurate BZ sampling do not cancel when taking the total-energy differences between defect and perfect-crystal supercells [9].

In addition to the standard uniform jellium charge compensation, we have calculated the Coulomb energies using the newly proposed LMCC scheme [4]. The model density n_{LM} consists of a single Gaussian charge distribution with a width parameter $\alpha = 0.44$ bohr^{-2} centred at the interstitial site.

3. Formation energy

The formation energy E_{f}^{q} of the self-interstitial in charge state q as a function of the electron chemical potential μ_{e} is calculated as

$$E_{\mathrm{f}}^{q}(\mu_{\mathrm{e}}) = E_{\mathrm{d}}^{q} - E_{\mathrm{b}} - \mu_{\mathrm{Si}} + q(\mu_{\mathrm{e}} + E_{\mathrm{v}}), \tag{1}$$

where E_{d}^{q} is the total energy of the defect-containing supercell and E_{b} is the corresponding defect-free bulk supercell energy. The silicon chemical potential is calculated from $\mu_{\mathrm{Si}} = E_{\mathrm{b}}/N_{\mathrm{at}}$, where N_{at} is the number of atoms in a bulk supercell. The electron chemical potential is referenced to to the valence-band maximum E_{v} in the perfect crystal.

In the following we study the simple difference of the supercell energies as a function of L

$$E_{\mathrm{def}}^{q}(L) = E_{\mathrm{d}}^{q}(L) - E_{\mathrm{b}}(L), \tag{2}$$

which contains the artificial, size-dependent defect–defect Coulomb-interaction energy. The errors made in the determination of $E_{\mathrm{def}}^{q}(L)$ transfer directly to the formation energies and depend on the charge state as $\propto q^2$. The errors in ionization levels, where the ionization levels are defined as the values of the electron chemical potential at which the most stable configuration changes charge state, depend on the charge as $\propto q$.

The calculated energy differences $E_{\mathrm{def}}^{2+}(L)$ of the doubly positive silicon self-interstitial are presented in figure 1. We have presented the energies versus $1/L$, anticipating the classical $1/r$ asymptotic convergence. The dashed lines are drawn so that they pass through the points corresponding to the two largest supercells. It is gratifying to notice that the lines nearly meet at the infinite-supercell limit, the limiting values being -121.45 and -121.61 eV for the uniform compensation and LMCC methods, respectively.

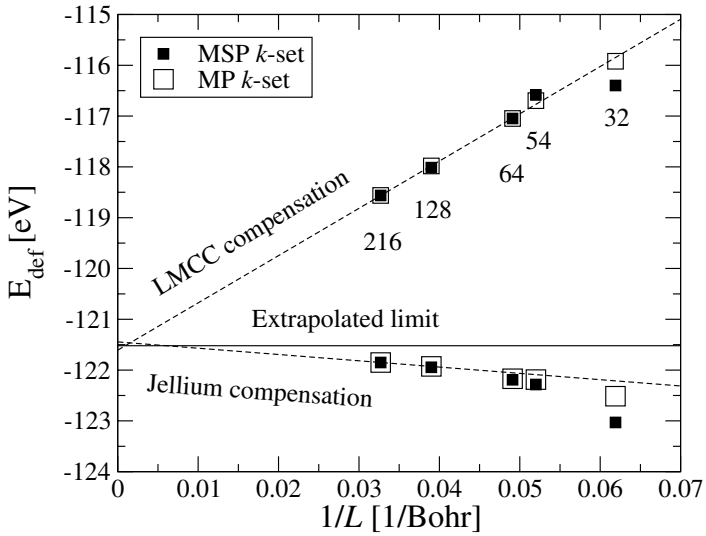

Figure 1. The energy differences E_{def}^{2+} for different supercell sizes calculated using jellium and LMCC charge compensation schemes. A horizontal line is drawn at the average value of the jellium and LMCC extrapolations at 121.5 eV.

The energy difference E_{def}^{2+} converges asymptotically as L^{-1} with the both compensation schemes. The slope of the line corresponding to the jellium compensation in figure 1 is smaller than the slope of the line corresponding to the LMCC compensation. The slopes are determined by the prefactors of the L^{-1} error components, and depend on the strength of the dielectric screening of the medium. The prefactor for materials with strong screening is smaller (larger) for jellium (LMCC) compensation, and vice versa for more ionic materials—the two extreme cases being metallic media and vacuum.

Makov and Payne [2] have derived a (post-) correction formula for the electrostatic energy of an isolated charge distribution within PBC and uniform compensation in cubic lattices [2],

$$E(L) = E_0 - \frac{q^2\alpha}{2\epsilon L} - \frac{2\pi q Q}{3\epsilon L^3} + \mathrm{O}[L^{-5}], \tag{3}$$

which contains an L^{-3} term in addition to the L^{-1} term identical to that derived by Leslie and Gillan [1]. The first correction term with the charge q of the defect and the appropriate superlattice Madelung constant α describes the electrostatic energy of the point charge array in a uniform background in the presence of screening medium described by the static dielectric constant ϵ. The numerical value of the Madelung constant α depends on the definition of the size parameter L and whether the factor $1/2$ is included or not. The numerical values corresponding to the definition of L as the cube root of the supercell volume and equation (3) are 2.8373, 2.8883 and 2.885 for SC, BCC and FCC supercells, respectively.

The second correction term describes the interaction of the defect charge distribution with the neutralizing jellium. The parameter Q is defined as the second radial moment of the aperiodic defect charge density, $\rho_{\mathrm{ap}} = \rho_{\mathrm{d}} - \rho_{\mathrm{b}}$, which is defined as the difference between the valence electron densities of the defect containing supercell and the corresponding perfect crystal supercell. The definitions of the parameters ϵ and Q in the L^{-3}-order term contain some ambiguity (see the discussion below). The value $Q = 14.89$ bohr^2 used in table 1 is taken from the calculation of an isolated atom in vacuum, and the value $\epsilon = 12$ roughly corresponds to the experimental value of the static dielectric constant.

Table 1. Makov–Payne correction applied to E_{def}^{2+} energy differences calculated with jellium compensation and MP k-sets. Columns Δ_1 and Δ_2 give the contributions from L^{1-} and L^{3-} order terms. Energies are given in eV.

N_{at}	E_{def}^{2+}	Δ_1	Δ_2	$E_{\mathrm{def}}^{2+} + \Delta_1$	$E_{\mathrm{def}}^{2+} + \Delta_1 + \Delta_2$
32	−122.51	0.81	0.03	−121.70	−121.67
54	−122.19	0.68	0.02	−121.51	−121.49
64	−122.17	0.63	0.02	−121.54	−121.52
128	−121.93	0.51	0.01	−121.42	−121.41
216	−121.85	0.42	0.00	−121.43	−121.42

The values for the energy differences corrected using the analytic expression equation (3) in table 1 agree with the extrapolated limit of the numerical data in figure 1, and differ significantly from the uncorrected values. We can provide the result with an absolute error estimate $E_{\mathrm{def}}^{2+} = 121.5 \pm 0.1$ eV, the error bounds including the BZ sampling and defect–defect Coulomb interaction error contributions. We notice however that there still remains a small systematic error in the corrected results in table 1, which is related to the under-estimation of the L^{-3}-order correction term (see discussion).

4. Discussion

A quantitative example of the effects of the Makov–Payne correction on the formation energies and ionization levels can be found in the recent calculation of the monovacancies in silicon–germanium [10]. The calculated energies include ion-relaxation contributions. The effect of the correction on the model system of a 64-atom zincblende supercell and a Si vacancy is illustrated in figure 2. Without any correction for the spurious defect–defect and defect–jellium interactions, the results imply that the stable configuration of the Si vacancy in SiGe changes first from doubly positive to negative (2+/−) and then from negative to quadruply negative (−/4−) as the electron chemical potential is increased. This would be interpreted as a strong negative-effective-U phenomenon ($U_{\mathrm{eff}} < 0$), that is, the energy gain from the lattice–lattice relaxations exceeds the Coulomb repulsion energy U of electrons on (nearly) degenerate orbitals. The calculated lattice relaxation patterns do not support this interpretation, not to mention the unlikelihood of finding a 4− charge state within the gap region. The lattice relaxations for charge states from 1 to 4 are relatively small and the NN symmetry does not change. The Makov–Payne Coulomb correction restores the repulsion between the electrons on degenerate orbitals, resulting in an almost constant U_{eff} for the higher-charge states, consistent with the static relaxation pattern. It also removes the negative-effective-U effect from the positive-charge states. We note that the negative-effective-U character of the (2+/0) transition of the Ge vacancy is not destroyed by the correction.

Another illustrative example of the practical effects of the correction to the defect properties is the comprehensive study of monovacancies and antisites in 4H–SiC by Torpo *et al* [11]. They present numerous formation energy plots with and without the first-order correction.

All the above-mentioned schemes, the correction for the jellium compensation and the mixed boundary condition methods, are derived with the assumption that the defect charge is localized within the supercell. Strictly speaking, in the classical theory, this assumption is not valid, as the polarization and thus the aperiodic screening charge distribution in principle extend to the whole macroscopic crystal.

If we compare the two charge-compensation schemes on the qualitative, classical electrostatic level, we can see that in many respects the ideal uniform compensating charge scheme corresponds to metallic screening, and the localized compensating scheme to the limit

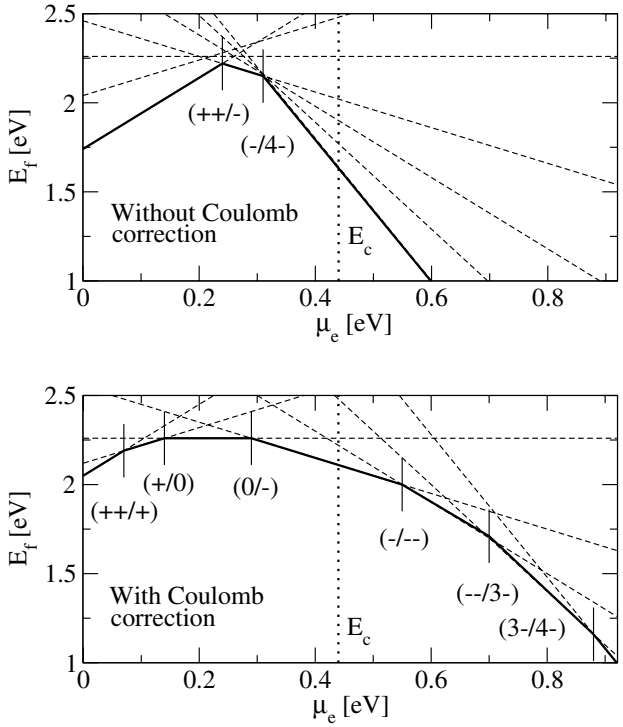

Figure 2. Formation energy of the Si vacancy in SiGe as a function of the electron chemical potential, with (above) and without (below) the Makov–Payne correction [10]. The bottom of the conduction band (LDA) is shown as a vertical dotted line.

where screening is absent outside the supercell, i.e. non-polarizable media or vacuum. This is illustrated in figure 3, which shows the integrated aperiodic charge $Z(r) = \int_0^r \mathrm{d}r\, 4\pi r^2 \langle \rho(r) \rangle$ used in the calculation of the Coulomb energy.

Figure 3 shows that at the border of the supercell the integrated charge (including the jellium contribution) approaches the number of the valence electrons of the neutral interstitial in the case of the jellium compensation. Counting the ionic charge of the interstitial, the defect appears completely screened, whereas in the case of the LMCC method the charge inside the supercell is the same as the nominal charge state 2+ of the defect, and at the border of the supercell the defect appears completely unscreened. Classically, the amount of charge, including the screening charge, should be a fraction of the nominal charge q^2/ϵ inside the supercell.

Looking at the integrated aperiodic valence electron distributions (without the jellium contribution) in figure 4, we notice that the electron densities are nearly identical inside the first NN shell. The valence electron distributions calculated using the LMCC charge compensation extend correctly further away from the defect than in the jellium compensation scheme. The positive tails (negative electron density) of the aperiodic densities are in the border regions of the corresponding supercells.

The aperiodic charge ρ_{ap} of the supercell does not fully correspond to the classical picture of an isolated localized positive charge at the defect site. At close distances, the defect is well screened and the charge distribution ρ_{ap} is similar to the charge distribution of a *neutral* pseudo-atom in vacuum. Since the screening electrons must come from somewhere,

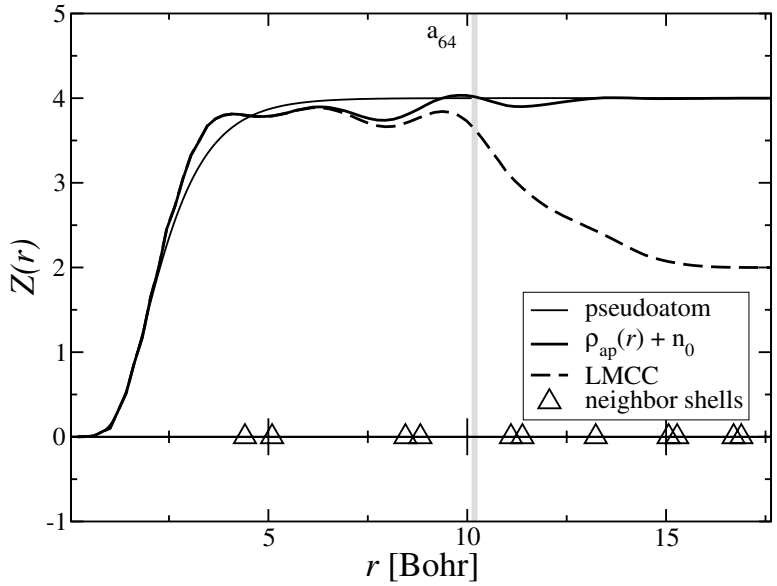

Figure 3. Integrated aperiodic charge used in the calculation of the Coulomb energy at the distance r from the 2+ interstitial in the 64-atom supercell. The thin solid curve shows the valence electron density of an isolated atom for reference, the thicker solid curve is the aperiodic density from the jellium-compensated calculation plus the jellium contribution and the dashed curve is the aperiodic charge density from the LMCC-compensated calculation. Triangles show the distances to the neighbouring atom shells. The shortest distance from the defect to the border of the supercell is indicated with a vertical line with a label a_{64}. Beyond a_{64} the contribution to the integral $Z(r)$ is from the 'corner regions' of the supercell.

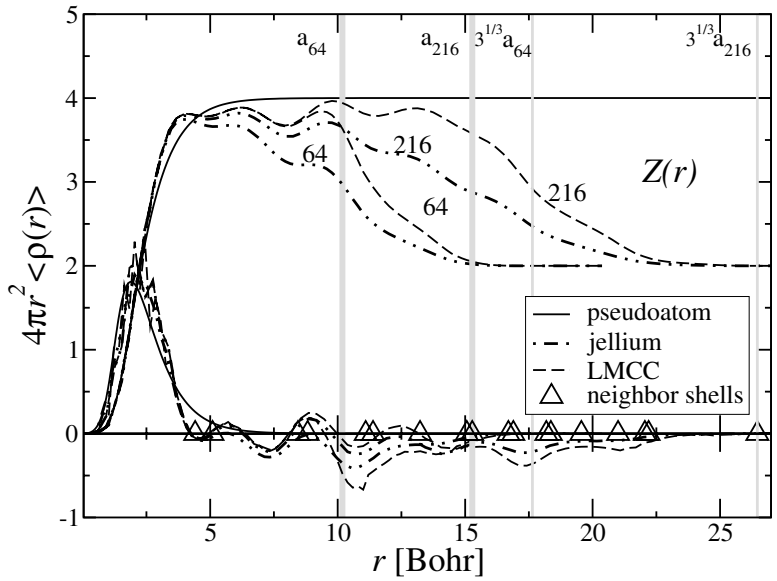

Figure 4. The aperiodic valence-electron density $4\pi r^2 \langle \rho_{\rm ap}(\boldsymbol{r}) \rangle$ and its cumulative sum $Z(r)$ at the distance r from the interstitial.

and the charge of the supercell is fixed, the screening charge comes from where it is energetically most favourable—dilutely, from the borders of the supercell. The density ρ_{ap} has a substantial positive region near the borders of the supercell. This region gives a major contribution to the integral defining the second radial moment $Q = \int \mathrm{d}^3 r\, r^2 \rho_{\mathrm{ap}}(r)$. Thus, the above definition of Q depends on the size of the supercell and does not lead to a well defined value.

The classical treatment of the Coulomb interaction in dielectric materials separates the external charges and the induced, screening charges. Since the effects of the screening in the Makov–Payne correction formula (equation (3)) are taken into account using a dielectric constant, the aperiodic charge ρ_{ap} should correspond to the external charge. Exactly the same holds for the determination of the model charge n_{LM} in the LMCC method. We have used the valence electron distribution of an isolated silicon atom to estimate the second radial moment Q in the Makov–Payne correction formula, and in the determination of the width parameter of n_{LM} in the LMCC method.

As mentioned, there is still a small systematic error in the corrected energies in table 1. We attribute this error to the L^{-3}-order term, which seems to be underestimated. The L^{-3}-order term describes the interaction between the aperiodic charge distribution and the jellium. This interaction is not as long ranged as the defect–defect interaction. Instead, it is short ranged and the screening is reduced from the long-range limit of ϵ. We propose that this is the reason for the apparent under-estimation of the L^{-3}-order correction.

The final estimate of the defect formation energies in semiconductors also contains a number of factors besides the Coulomb energy discussed here. For example, heavy doping can be argued to turn a semiconductor material metallic, influencing the screening [12].

5. Conclusions

The magnitude of the Coulomb correction for a given supercell size can and should be estimated at least by calculating the first-order correction, corresponding to a point charge array and uniform compensating background in structureless dielectric. The correction systematically moves the formation energies up in energy, and pushes the positive ionization levels in the gap towards the valence band maximum and the negative levels towards the conduction band minimum (figure 2). The general effect is to reduce the (over-) binding of the charged defects.

In practice, both LMCC and jellium charge-neutralization schemes have to rely on the use of a screening parameter or the direct extrapolation of the results for supercells of different sizes. To obtain confidence in the absolute convergence with respect to supercell size, we propose the use of a plot similar to figure 1.

Acknowledgments

This research has been supported by the Academy of Finland through its Centre of Excellence Programme (2000–2005). The computations presented in this document have been performed with Cray T3E and Compaq Alpha Cluster equipment provided by CSC. JL acknowledges the financial support by Wihuri, Väisälä and Nokia foundations.

References

[1] Leslie M and Gillan M J 1985 *J. Phys. C: Solid State Phys.* **18** 973
[2] Makov G and Payne M C 1995 *Phys. Rev.* B **51** 4014
[3] Schultz P A 1999 *Phys. Rev.* B **60** 1551

[4] Schultz P A 2000 *Phys. Rev. Lett.* **85** 1942
[5] Nozaki H and Itoh S 2000 *Phys. Rev.* E **62** 1390
[6] Mozos J-L and Nieminen R M 1999 *Properties of Crystalline Silicon* ed R Hull (London: INSPEC, The Institution of Electrical Engineers) p 319
[7] Puska M J, Pöykkö S, Pesola M and Nieminen R M 1998 *Phys. Rev.* B **58** 1318
[8] Monkhorst H J and Pack J D 1976 *Phys. Rev.* B **13** 5188
[9] Makov G, Shah R and Payne M C 1996 *Phys. Rev.* B **53** 15 513
[10] Lento J, Mozos J-L and Nieminen R M 2000 *Appl. Phys. Lett.* **77** 232
[11] Torpo L, Marlo M, Staab T E M and Nieminen R M 2001 *J. Phys.: Condens. Matter* **13** 1
[12] Mäkinen J, Corbel C, Hautojärvi P, Moser P and Pierre F 1989 *Phys. Rev.* B **39** 10 162