PHYSICAL REVIEW B VOLUME 12, NUMBER 6 15 SEPTEMBER 1975

Lattice-relaxation effects at point imperfections in semiconductors*

Sokrates T. Pantelides†
Department of Applied Physics, Stanford University, Stanford, California 94305
(Received 3 March 1975)

We discuss some of the fundamentally inherent limitations involved in calculations of lattice-relaxation effects at point imperfections. We show that differences of tetrahedral covalent radii between impurity and host atoms cannot reliably determine whether the nearest-neighbor atoms move inward or outward. A massive minimization-of-total-energy calculation may yield reliable information for deep levels but for shallow levels the problem is inherently intractable; even if one knew which way the neighboring atoms move, neither the sign nor the size of the change in binding energy could be determined reliably, either by a full calculation or, even more so, by qualitative arguments.

Point imperfections in semiconductors, such as impurities, vacancies or interstitials, are known to produce local distortions in the lattice of the otherwise perfect crystal. The importance of these distortions and their contribution to the electron or hole-binding energies has been the subject of much study in the literature. Rigorous calculations are of course very difficult, and as a result the size of the effect has, in general, been estimated in terms of simple models. Unfortunately, however, these models are flexible enough to produce the "right" answer in various circumstances with the net result of confusing rather than clarifying the issue. A case at hand is that of the chemical shifts of shallow donors in Si and Ge. Until very recently, it was thought that a one-valley effective-mass calculation was adequate for the ground state of a screened-Coulomb potential in these materials. The resulting discrepancies between theory and experiments, the well-known chemical shifts, were the subject of much controversy for a long time. Several authors[1-4] showed that the right order of magnitude for the chemical shifts was obtainable in terms of simple models of the strains produced by local lattice distortions. When it recently became clear that most of the so-called chemical shift was due to intervalley mixing rather than chemical or strain effects,[5-8] the models could easily be toned down to yield smaller corrections. This, for example, was done recently by Schechter,[9] who calculated binding energies for the shallow donors in silicon by using pseudopotentials and including intervalley mixing. He then proceeded to include contributions from strain fields using Weinreich's[1] model. Despite the fact that Weinreich's model gives a vanishing correction for isotropic wave functions (and Schechter's wave-functions are isotropic), Schechter used the following formula obtained by Weinreich from perturbation theory

$$\Delta E = -4\Xi_u r_0 \Delta r / a_0^2 \tag{1}$$

where Ξ_u is the deformation potential constant, r_0 is the tetrahedral covalent radius of the donor, Δr is the difference in the tetrahedral radii of donor and host, and a_0 is the transverse Bohr radius of the unperturbed wave function. In the case of P in Si, Schechter calculated a binding energy of -53 meV. He then estimated the contribution from strain fields from Eq. (1) to be $+7$ meV for a final binding energy of -46 meV, in perfect agreement with the experimental value of -45.5 meV. In using (1), however, he used the *hydrogenic* value for a_0 (20 Å, which corresponds to a binding energy of ~30 meV), had he used the average Bohr radius of *his own unperturbed wave function* instead, namely ~9 Å, he would obtain a much larger correction, a whopping $+34.6$ meV, for a disastrously low net binding energy of -18.4 meV.

A very analogous case is that of shallow acceptors in silicon and germanium. The chemical shifts, particularly in silicon, have long been known to be both large and sensitively dependent on the chemical species. Morgan[10] used the differences of covalent radii between impurity and host atoms to calculate strain and Stark fields and roughly accounted for the observed shifts. More recently, however, Lipari and Baldereschi[11] showed that the bulk of the "chemical" shift for Al in Si (for which the Coulomb potential is expected to work best) is actually due to anisotropic terms in the effective-mass kinetic-energy part of the Hamiltonian. In order to calculate this correction, one has to include spherical harmonics of $l=4$ in the trial function. More recent work by the present author[12] in terms of the pseudoimpurity theory[6,8,13] (PIT), previously developed and applied to donors in silicon, shows that the shifts of the binding energies of B, Ga, and In in Si, as measured from the binding energy of Al in Si, are truly chemical in nature and arise mostly from the *reorthogonalization-potential* terms.[6,8,13]

In other cases, strain fields, lattice distortions, and the like have been used to explain noted discrepancies at least qualitatively: The present author[6,13] attributed the discrepancy between theoretical and experimental binding energies of the heavy Sb and Te impurities in Si to distortions of the impurity-atom core orbitals caused by the nearest-neighbor atoms which squeeze on them; Allen[14] pointed out the importance of strain fields in the case of isoelectronic traps and suggested the use of covalent radii in determining them; Phillips[15] suggested that lattice distortions actually dominate the determination of their binding energies; Baldereschi and Hopfield[16] studied the same problem in more detail and used the differences in bond lengths to obtain quantitative results; Kukimoto *et al.*[17] interpreted their experimental data on the binding energy of a second electron at oxygen sites in GaP by invoking a very large local lattice distortion; Jaros and Ross[18] included lattice distortions in their calculation of the binding energy and found that "reasonable" distortions could produce sizable additional binding, but nowhere close to the observed value; Grimmeiss *et al.*[19] remeasured and reinterpreted the data on oxygen in GaP. In their model the second electron occupies a p-like state with fairly large binding energy which Morgan[20] suggested can be understood by invoking strain fields. By similar arguments in terms of strain fields, Morgan[21] also suggested an explanation for the unusually large binding energy of the Ge donor in GaP (~200 meV compared with 65.5 meV for Sn and 82.5 meV for Si in GaP, all of them substitutional at a Ga site).

All the arguments about lattice distortions and strain fields, such as those just mentioned, are in fact plausible, and numerical estimates do or are likely to support the conclusions. The purpose of this note is not to offer unfair criticism of good-faith attempts to offer reasonable explanations for complicated situations. However, in view of a growing tendency to rely on strain fields to explain important effects, we wish to explore here some of the fundamentally inherent limitations one is faced with in attempting to calculate lattice-relaxation effects.[22]

First we wish to illustrate the fact that it is not a trivial or perhaps even feasible task to determine whether the nearest-neighbor atoms move outward or inward, which is the one factor that determines the nature of the strain fields. As we saw already, a procedure that has been used extensively is to compare the covalent tetrahedral radii of host and impurity atoms and conclude that if the impurity radius is larger, the neighbors move out, in order to allow for the embedding of a larger "sphere" in the crystal, and vice versa.[23] Such arguments, if valid, would suggest that the atoms around a vacancy move inward since the "impurity radius" is zero.[24] Indeed, many authors studied this question in terms of the "defect-molecule" model and concluded that the atoms do move inward.[25] However, recently, Messmer and Watkins[26] and Yip[27] carried out *ab initio* calculations, minimized the total energy, and showed that the atoms around a neutral vacancy *in fact move outward*. For other charge states, Yip found the movement to be inward or outward depending on whether the crystal is diamond, silicon, or germanium (Table I). These calculations of course involve many drastic approximations and one should be cautious in view of the difficulty in predicting crystal structures and lattice constants of perfect crystals via minimization-of-total-energy *ab initio* calculations, but they do serve to illustrate our first point: Unless one resorts to a massive total-energy calculation, it is impossible to tell reliably which way the nearest-neighbor atoms move by using intuitively appealing arguments such as differences in covalent radii of atoms.[28] Particulary, in the case of shallow impurities and isoelectronic traps where the energy changes are very tiny, it appears that no argument can ever be conclusive as to which way the atoms go.

Let us now move to our second point. Suppose one first calculates the position of an impurity level with the lattice frozen (by effective-mass theory or in some other way), as was done in Refs. 1–3, 9, and 10. Suppose further that one could estimate or guess which way the atoms move. Could then one tell reliably whether the impurity level moves up or down in energy? The answer is by no means trivial, since the lowering of the *total* energy of the whole system does not necessarily imply lowering of the single bound level of interest. Let us first see what happens if we attempt to obtain an answer by perturbation theory which has often been the case. This means that we first solve the one-electron problem,

$$[-(\hbar^2/2m)\nabla^2 + V]\Psi = E\Psi , \qquad (2)$$

where V is the total potential of the crystal containing the defect, but with all the nuclei frozen at their perfect-crystal postions. If we then let the

TABLE I. The movement of the four atoms neighboring a vacancy in diamond, Si, and Ge after Yip (Ref. 24); "in" indicates inward movement, "out" indicates outward movement.

Crystal \ Charge state	V^+	V^0	V^-	V^{--}
Diamond	out	out	out	out
Silicon	out	out	out	in
Germanium	out	out	in	in

nuclei move to their new positions, an additional perturbation potential ΔV is introduced, which may be written as

$$\Delta V = \sum_j [v(\vec{r} - \vec{R}_j) - v^0(\vec{r} - \vec{R}_j^0)] , \tag{3}$$

where $\vec{R}_j^0$ are the atomic positions in the perfect crystal, and $\vec{R}_j$ are the atomic positions in the perturbed crystal. In the notation used in (3), the perfect-crystal potential V^0 would be

$$V^0 = \sum_j v^0(\vec{r} - \vec{R}_j^0) . \tag{4}$$

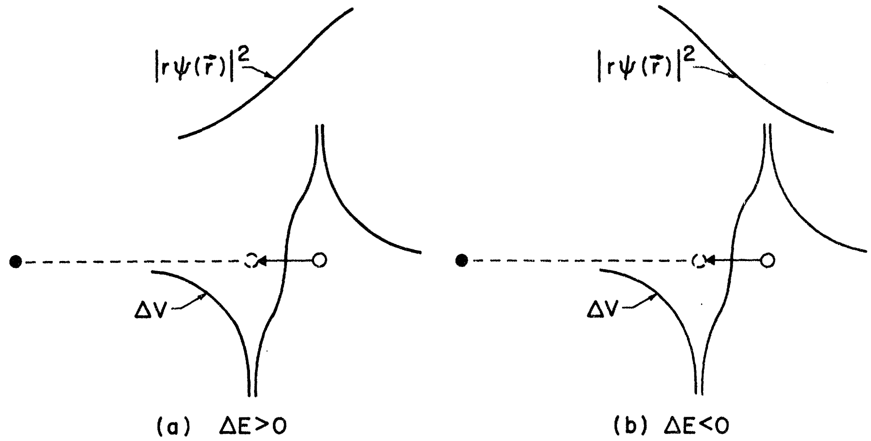
Clearly, our task would be to evaluate the change in energy ΔE, given to first order by

$$\Delta E = \langle \Psi | \Delta V | \Psi \rangle . \tag{5}$$

The form of the perturbation ΔV resulting from an atom moving *inward* is illustrated in Fig. 1. Since ΔV is essentially symmetric about the midpoint of the displacement vectors, it is clear from Fig. 1 that the sign of ΔE will be different depending on whether the slope of $|r\Psi(\vec{r})|^2$ is positive or negative in the region (See Fig. 1). Since the atomic site is the region where Ψ has nodal oscillations, the sign (and magnitude) of ΔE becomes a very sensitive quantity and practically impossible to determine. Even if the problem is formulated in terms of pseudopotentials, in which case Ψ is a pseudo-wave function, some wiggles may still be left in the region of the atomic sites to make the determination of ΔE a very tricky task. In practice, the situation is even worse, since (2) is generally solved by making serious approximations, which result in wave functions that are not very reliable. Most of the time , (2) is solved variationally and the flexibility allowed in the trial form in general is not adequate to produce wiggles of the kind that can determine the sign or size of ΔE accurately. A case at hand is the effective-mass theory and its variants in which the trial funciton is generally taken to be an exponential-like envelope function times a Bloch function. In most cases the envelope function has an upward slope at the nearest neighbors (which move the most), as in Fig. 1(a), whereby it is guaranteed to give a positive ΔE when the atoms move in and a negative ΔE when the atoms move out.[29] When the envelope function is multiplied by the Bloch function (or pseudo-Bloch function), as it should, no guarantee remains that the slope of the total wave function at the atomic sites remains the same, leaving even the sign of the effect an unresolved matter.[30]

Finally, what about including ΔV in V of (2) and solving directly, rather than by perturbation theory? The essence of the problem here is solving (2) in approximations better than those of effective-mass theories. This is necessary because we are really trying to probe the detailed oscillatory nature of Ψ at lattice sites, whereas in effective-mass theories the discreteness of the lattice is washed out. Going beyond effective-mass theories, however, is a task involving massive computations. The approaches of Messmer and Watkins,[26] Yip,[27] and that of Jaros and Ross,[18,31] if carried to convergence, can be suitable for deep levels. With caution, in view of the variational nature of the wave function, they may supply us many answers. On the other hand, these methods would be far less reliable for fairly shallow levels where tiny changes in large energies are the object. In these cases, it appears that the determination of which way the atoms move and of the sign and size of the change in binding energy will remain a very uncertain matter.

In conclusion, the main messages we wish to convey in this note are: (i) Comparison of the covalent radii of impurity and host atoms cannot tell



(a) ΔE>0 (b) ΔE<0

FIG. 1. Illustration of how the change in binding energy ΔE may be either positive or negative depending on the slope of the wave functions at the neighboring atoms. We show the case when the neighboring atoms (open circles) move *toward* the impurity atom (dark circle). A similar illustration can be made when the neighboring atoms move out, away from the impurity atom.

us whether the atoms neighboring an impurity atom move inward or outward. (ii) Even if we knew where the neighboring atoms move, perturbation theory cannot determine reliably even the sign, nonetheless the size, of the correction to the binding energy. (iii) Only massive minimization-of-total-energy calculations have any hope of determining both the positions of the neighboring atoms and the net binding energy. However, this is only workable for fairly deep levels. For shallow levels, the changes in the total energy due to the motion of the neighbors are too tiny to be computable.

---

[*] Work supported by the National Science Foundation Grant No. GH-39811.

[†] Present address: IBM T. J. Watson Research Center, Yorktown Heights, N. Y. 10598.

[1] G. Weinreich, J. Phys. Chem. Solids 8, 216 (1959).

[2] S. Shinohara, Nuovo Cimento 22, 18 (1961).

[3] A. Morita and H. Nara, J. Phys. Soc. Jpn. 21, S234 (1966).

[4] J. C. Phillips, Phys. Rev. B 1, 1540 (1970).

[5] The importance of valley-orbit mixing was first pointed out by Narita and Shimizu [J. Phys. Soc. Jpn. 16, 2588 (1961)] but their short note apparently went largely unnoticed. Baldereschi [Phys. Rev. B 1, 4673 (1970)] calculated the valley-orbit splitting in Si and Ge and showed that it is due to the reduced screening of intervalley potential terms. A direct calculation of the ground-state energy of a screened-Coulomb potential was subsequently given in Refs. 6 and 7.

[6] S. T. Pantelides, Doctoral thesis (University of Illinois, 1973) (unpublished).

[7] S. T. Pantelides and C. T. Sah, Solid State Commun. 11, 1713 (1972); Phys. Rev. B 10, 621 (1974).

[8] S. T. Pantelides, in *Proceedings of the Twelfth International Conference on the Physics of Semiconductors*, edited by M. H. Pilkuhn (Teubner, Stuttgart, 1974), p. 396.

[9] D. Schechter, Phys. Rev. B 5, 652 (1973).

[10] T. N. Morgan, in *Proceedings of the Tenth International Conference on the Physics of Semiconductors*, edited by S. P. Keller, J. C. Hensel and F. Stern (USAEC, Cambridge, Mass., 1970), p. 266.

[11] N. O. Lipari and A. Baldereschi, Bull. Am. Phys. Soc. 19, 198 (1974); and (unpublished).

[12] S. T. Pantelides (unpublished).

[13] S. T. Pantelides and C. T. Sah, Phys. Rev. B 10, 638 (1974).

[14] J. W. Allen, J. Phys. C 1, 1136 (1968).

[15] J. C. Phillips, Phys. Rev. Lett. 22, 285 (1969).

[16] A. Baldereschi and J. J. Hopfield, Phys. Rev. Lett. 28, 171 (1972).

[17] H. Kukimoto, C. H. Henry, and F. R. Merritt, Phys. Rev. B 7, 2486 (1973).

[18] M. Jaros and S. F. Ross, Ref. 8, p. 401.

[19] H. G. Grimmeiss, L. A. Lebedo, C. Ovren, and T. N. Morgan, Ref. 8, p. 386.

[20] T. N. Morgan, Ref. 8, p. 391 and private communication.

[21] T. N. Morgan (unpublished), see also F. Mehran, T. N. Morgan, R. S. Title, and S. E. Blum, Phys. Rev. B 11, 661 (1972).

[22] We will not discuss at all Jahn-Teller-type effects in this paper. Many excellent accounts of them exist.

[23] This approach is usually coupled to elastic-medium theory and deformation potentials; see e.g., Refs. 1–3, 10, and T. N. Morgan and H. Maier, Phys. Rev. Lett. 27, 1200 (1971).

[24] Viewing the vacancy as an impurity with zero radius is the natural, though perhaps unreasonable limit of the elastic medium theories (Refs. 1–3, 10, and 23). In such theories, one first carves out a cavity (vacancy) in the continuous medium and then inserts in it a sphere of different radius (the impurity radius) but with the same elastic properties as the rest of the medium. On the other hand, one might argue [J.A. Van Vechten (private communication)], using Pauling's bond-order–bond-length relations, that the electrons of the dangling bonds redistribute themselves in the backbonds, which become stronger and thus shorter. The atoms nearest the vacancy would then move out. According to this argument, an extra electron at the vacancy (V^-) should have the same effect and make the outward movement even more pronounced. Yip's calculations (Table I), however, indicate that it is for some V^- and V^{--} states that the atoms in fact move in.

[25] J. Friedel, M. Lannoo, and G. Leman, Phys. Rev. 164, 1056 (1967); M. Lannoo and A. M. Stoneham, J. Phys. Chem. Solids 29, 1987 (1968); C. A. Coulson and F. P. Larkins, J. Phys. Chem. Solids 32, 965, 2245 (1971); F. P. Larkins, J. Phys. Chem. Solids 32, 2123 (1971); Radiat. Eff. 9, 5 (1971); F. P. Larkins and A. M. Stoneham, J. Phys. C 4, 143 (1971).

[26] R. P. Messmer and G. D. Watkins, Phys. Rev. B 7, 2568 (1973).

[27] K. L. Yip, Phys. Status Solidi 66, 619 (1974).

[28] The failure of these simple intuitive arguments is evident in the results of Morgan and Maier, Ref. 23. These authors begin by proposing a theory for the observed splittings in pair spectra in terms of a model based on strain fields produced by the misfit of the impurity-atom "sphere" in the "spherical" cavity left by the host atom plus electrostatic forces. The experimental data in fact fit to values of K_T and K_ϵ which are *independent* of the donor impurity, whereas the differences in atomic radii predict a strong dependence on the donor impurity (see Fig. 3 of Ref. 23).

[29] The results of Refs. 1, 9, and 10 obey this pattern.

[30] We may note in conjunction with Fig. 1 that one could determine the slope of the wave function ψ at the nearest neighbors if one could independently determine *both* the sign of the change in binding energy and the movement of the nearest neighbors.

[31] It should be noted that the procedure of Jaros and Ross (Ref. 18) can be applied only when the positions of the neighboring atoms are known from an alternative calculation or experiment.