Ann. Rev. Mater. Sci. 1985. 15:411–453

THEORY OF 3d TRANSITION ATOM IMPURITIES IN SEMICONDUCTORS[1]

Alex Zunger

Solar Energy Research Institute, Golden, Colorado 80401

INTRODUCTION

The utility of semiconductors in many device applications, such as high-speed logic circuits (1–2), optoelectronic devices (3–4), microwave devices (5), and solar cells (6), rests on the characteristics of the intentionally introduced impurities (dopants) as much as on the properties of trace amounts of unintentional contaminations. Transition atom (TA) impurities in semiconductors form a special class of such contaminants. They were studied experimentally in great detail (e.g. review articles in References 7–10) using a broad range of techniques, including optical absorption, luminescence, photocapacitance, photoconductivity, electron paramagnetic resonance (EPR), electron nuclear double resonance (ENDOR), deep level transient spectroscopy (DLTS), and Hall effect. This review article is concerned with the theoretical understanding of the electronic properties of TA impurities in Si, III–V, and II–VI semiconductors.

First, we establish the nomenclature. When a transition atom takes up a substitutional site, say on a cation, its formal oxidation state when neutral (labeled A^0) becomes that of the site it replaces, e.g. TA^{3+} if it replaces a column III element. When the impurity captures an electron its charge state becomes negative (denoted A^-), and the oxidation state is TA^{2+}. Conversely, when the impurity loses an electron its charge state becomes positive (labeled A^+), and the oxidation state becomes TA^{4+}. The tenfold degenerate atomic d orbitals can split in the cubic environment into a

[1] The US Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

sixfold degenerate t_2 representation and a fourfold degenerate e representation. This nomenclature is illustrated in Figure 1. Few of the elements of the structural and electronic properties of transition atom impurities can be appreciated from a cursory look at the experimental data and from simple atomic considerations. Assume first that the crystalline environment is a weak perturbation on the oxidation states of the transition atom impurity, and examine the consequences and contradictions of this assumption vis-à-

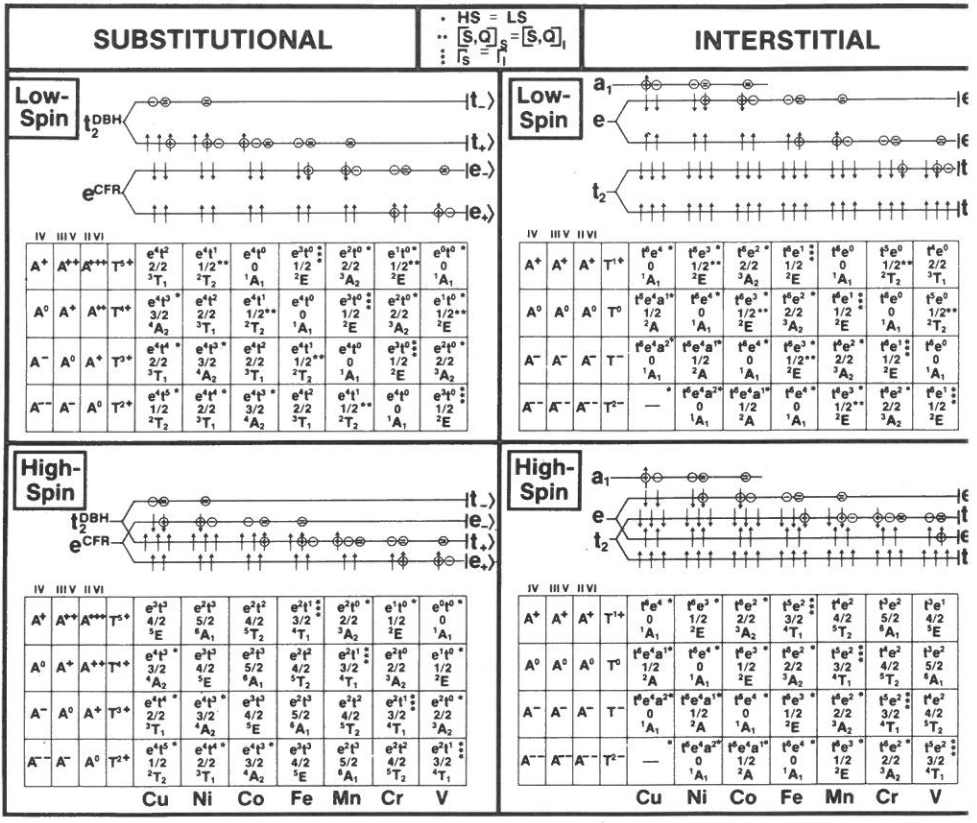

Figure 1 Summary of one-electron levels (t_+, t_-, e_+, e_-) and multiplets for substitutional (*left*) and interstitial (*right*) 3d impurities in semiconductors. Different charge states are denoted A^+, A^{++}, A^0, A^-, and $A^=$. Different oxidation states are denoted as T^{2+}, T^{3+}, T^{4+}, and T^{5+}. We give results for impurities in Si (denoted IV), III–V, and II–VI semiconductors. A single asterisk indicates that the high-spin and low-spin configurations give the same spin (S). A double asterisk indicates that the spin (S) and charge (Q) of the substitutional impurity are identical to those of the interstitial. A triple asterisk indicates that the multiplets for substitutional and interstitial configurations are identical.

vis experimental evidence. Figure 2 displays the calculated (11) 3d orbital energies of the free ions Ni through V in these three oxidation states. The figure shows both the spin-up (d_+) and the spin-down (d_-) energies; their separation (*shaded area* in Figure 2) is the exchange splitting. First, observe that the orbital energies become less negative with decreasing impurity atomic number Z, going from Ni to V (a Coulomb attraction effect). The slope of this reduction in binding energies is about 1–3 eV per increment of Z. Since this is also the order of magnitude of the semiconductor band gap (shown schematically at the top of Figure 2), at this rate at most one impurity will have an energy level inside the semiconductor band gap. In contrast, all impurities shown in Figure 2 were observed experimentally to have band gap levels in GaP, InP and GaAs (7–8).

Second, note in Figure 2 that the binding energy increases rapidly as the oxidation state changes from 2+ to 3+ to 4+, at a rate of about 20 eV per increment in the oxidation state. This is related to the (Mott-Hubbard)

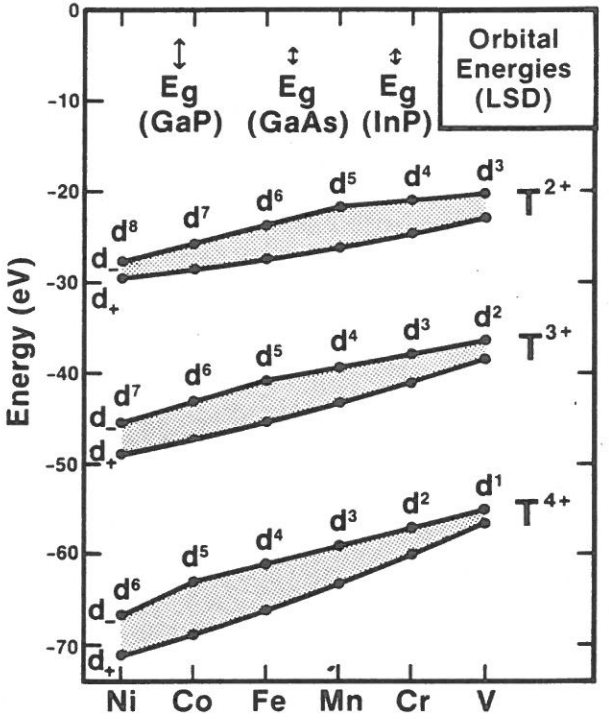

Figure 2 Calculated (local spin density, Reference 11) orbital energies for free ions of transition (T) atom elements.

single-site Coulomb repulsion effect; interelectronic repulsions of the order $U \approx 20$ eV are characteristic of 3d atoms (12). At this rate, if a level of a particular transition atom exists inside the semiconducting gap it will have at most a single stable oxidation state (say TA^{4+}), and will be pushed into the conduction band when a second electron is captured (producing TA^{3+}). Similarly, if a level exists in GaAs or GaP (i.e. TA^{3+} when neutral) it will not exist in ZnS (TA^{2+} when neutral). However, the experimental data (7–8) show many charged states: For GaAs and GaP the impurities Mn and Co exist in the gap in two oxidation states, Fe and Ni exist in three oxidation states, and Cr was observed to exist in four oxidation states. Furthermore, these impurities give rise to gap levels both in trivalent host cations (e.g. GaP) and in divalent host cations (e.g. ZnS) (7, 10). For this to occur the Coulomb repulsion energies of the free ions would have to be reduced in the solid by 1–2 orders of magnitude, i.e. $U \sim 0.2$–2.0 eV. This is surprising because EPR and ENDOR measurements suggest that these impurities have localized orbitals resembling atomic 3d orbitals. Hence, much as in biological electron-transporting molecules (13), e.g. cytochrome C, where the different oxidation states of the transition atom (i.e. Fe) are separated by less than 1 eV, so the different ionization states of a TA are compressed into a narrow band gap range when the atom exists as an impurity in a semiconductor.

Third, the absorption spectra of TA-containing semiconductors at sub–band gap photon energies show sharp lines (7–10) that resemble the familiar multiplet transitions in free atoms. This suggests that the atomic nature of the atoms is preserved in the solid, in sharp contrast with the previous point. Fourth, whereas the free-ion orbital ionization energies of Figure 2 show a change in slope of 2.5 eV between high-spin ionizations and low-spin ionizations (Mn^{2+} d^5 and Fe^{2+} d^6, respectively), the change is five times smaller for the corresponding impurities (0.45 eV between the Mn and Fe first acceptors in GaP). Finally, EPR (7–8) shows these systems occur at their maximum spin. Hence, despite the fact that interelectronic repulsions are reduced substantially in the solid, the exchange splittings (2–5 eV in ions, cf Figure 2) must necessarily remain reasonably high to prefer the high-spin ground state (Hund's rule).

These observations lead one to conclude that transition atom impurities in semiconductors show dual behavior with respect to localization: They are sufficiently localized to maintain a large exchange splitting and a multiplet structure, but are also sufficiently delocalized to exhibit small Coulomb repulsion with the attendant plurality of charge states, a compressed excitation spectrum, and a slow reduction in binding energies with decreasing Z.

From these simple observations it appears that a theory is needed to

explain the chemical shift in binding energies (i.e. the reduction of their slope with Z), the Coulomb reduction (i.e. the attenuation of repulsion energies by 1–2 orders of magnitude), and the survival of the many-electron multiplet effects in the solid with their attendant high-spin ground states.

TWO TYPES OF ELECTRONIC TRANSITIONS

Absorption, luminescence, Hall effect, and photoconductivity experiments (surveyed in References 7–10) indicate two types of electronic transitions in this system: those that occur within a fixed formal charge state, and those that alter the formal charge state of the center.

Intra-Center $d \rightarrow d^*$ Excitations

The intra-center transitions are excitations within a fixed formal charge where the one-electron configuration of the impurity $e^m t^n$ changes to $e^{m-1}t^{n+1}$ or $e^{m+1}t^{n-1}$, etc. They are often referred to as $d^N \rightarrow (d^N)^*$ internal transitions, or crystal-field excitons. The change in the total many-electron energy (E_T) that accompanies such transitions is given by

$$\Delta^{(ij)} = E_T^{(j)}[e^{m'}t^{n'}] - E_T^{(i)}[e^m t^n], \qquad 1.$$

where $E_T^{(j)}[e^{m'}t^{n'}]$ and $E_T^{(i)}[e^m t^n]$ are, respectively, the total energies of multiplets $|j\rangle$ and $|i\rangle$. The number of intra-center transitions is limited by the size of the band gap E_g; hence only a rather small number of transitions are observed in GaP ($E_g = 2.35$ eV) compared with wider gap materials such as ZnS or NiO (none observed so far in silicon).

Ionizations

Whereas intra-center excitations are charge conserving, the second type of electronic transitions, "charge-transfer" thermal ionizations or photo-ionizations, changes the formal charge state of the center. We refer to these as ionizations, although the removed electron or added hole remains in the crystal and can have a finite amplitude on the excited atom. Two types of such ionization processes are pertinent here: acceptors and donors.

The single acceptor ionizations correspond to, for example, a valence band (VB) electron being transferred to an impurity level of type e or t, thereby changing the charge state from A^0 to A^-. If the fully occupied valence band has M electrons such a single acceptor ionization (a reduction reaction) can be described for a t level as

$$A^0, (VB)^M t^n \rightarrow A^-, (VB)^{M-1} t^{n+1}. \qquad 2.$$

This type of single acceptor process is often referred to as hole emission to

the valence band. The associated change in the total free energy E_T of the system is denoted as $H(-/0)$, or

$$H^{(i)}(-/0) = E_T^{(j)}[A^-,(VB)^{M-1}t^{n+1}] - E_T^{(i)}[A^0,(VB)^M t^n], \quad 3.$$

where each total energy term corresponds to the lowest multiplet (i,j) of the corresponding charge state. A complementary acceptor ionization process can occur when a conduction band (CB) electron is captured by an impurity center. This can be described as an electron capture process:

$$A^0,(CB)^1t^n \rightarrow A^-,(CB)^0t^{n+1}. \quad 4.$$

Double acceptor ionizations correspond similarly to a second valence band electron being transferred to the impurity levels, thereby changing the charge state from A^- to A^{2-}. This can be described as a second hole emission process:

$$A^-,(VB)^M t^{n+1} \rightarrow A^{2-},(VB)^{M-1}t^{n+2}. \quad 5.$$

The change in total free energy attendant upon this process can be denoted as $E(-/2-)$, or

$$E^{(i)}(-/2-) = E_T^{(j)}[A^{2-},(VB)^{M-1}t^{n+2}] - E_T^{(i)}[A^-,(VB)^M t^{n+1}]. \quad 6.$$

Again, the complementary process of capture of a second electron from the conduction band is possible, but in this case less likely.

In single donor ionization (an oxidation reaction) the impurity level emits an electron to the conduction band, changing the formal charge state from A^0 to A^+. This can be described as

$$A^0,(CB)^0t^n \rightarrow A^+,(CB)^1t^{n-1}. \quad 7.$$

The change in total free energy attendant upon this electron emission is denoted as $E(0/+)$ and given by

$$E^{(i)}(0/+) = E_T^{(j)}[A^+,(CB)^1t^{n-1}] - E_T^{(i)}[A^0,(CB)^0t^n]. \quad 8.$$

Double donor transitions can be constructed in a manner analogous to the single donor electron emission process.

Mott-Hubbard Coulomb Energies

Using the above definition of donor and acceptor energies, one can define the apparent Mott-Hubbard Coulomb repulsion energy $[U^{(\alpha\beta)}]$ as the energy required to transfer an impurity electron from orbital α on one site to orbital β on a distant impurity site. If we carry out this process for the neutral center A^0, we first remove an electron from orbital α to form A^+, investing an energy $E^{(\alpha)}(0/+)$ (Equation 8), and then add an electron to orbital β on a distant neutral center A^0, transforming it to an A^- center by

investing the energy $H^{(\beta)}(-/0)$ (Equation 3). If all three species, A^0, A^+, and A^-, involve the same space orbital ($\alpha = \beta$) in the electron transfer processes then the diagonal Mott-Hubbard energy for A^0 is given by the difference in single donor and single acceptor energies (referred to the same band edge) as

$$U^{\alpha\alpha}(A^0) = E^{(\alpha)}(0/+) + (H^{(\alpha)}(-/0) - E_g]. \qquad 9.$$

For most impurities $U^{\alpha\alpha}(A^0)$ is positive, i.e. the donor level is lower in the gap than the acceptor level if the two transitions are associated with the same center.

A different apparent Mott-Hubbard energy can be obtained for the A^- center by first removing an electron from orbital α on A^-, transforming the center to A^0 through an energy change $H^{(\alpha)}(-/0)$ (Equation 3), and then adding an electron to orbital β of an A^- center, transforming it to an A^{2-} center through an energy change $E^{(\alpha)}(-/2-)$ (Equation 6). The diagonal Mott-Hubbard energy for A^- is given by the difference between double and single acceptor energies as

$$U^{(\alpha\alpha)}(A^-) = E^{(\alpha)}(-/2-) - E^{(\alpha)}(0/-), \qquad 10.$$

where we have referred both acceptor energies to the valence band edge. For most impurities $U^{(\alpha\alpha)}(A^-)$ is positive, i.e. the single acceptor level is lower in the gap than the double acceptor level if both belong to the same center. Experimental values for these Mott-Hubbard energies are (1–10) $U^{(tt)}(A^-) = 0.73$ and 1.05 eV for GaP:Cr and GaP:Ni, respectively; $U^{(ee)}(A^0) = 1.39$ eV for GaP:Fe; and $U^{(tt)}(A^0) = 0.62 \pm 0.1$ eV for GaP:Cr.

SEPARATION OF TRANSITION ENERGIES INTO MEAN-FIELD AND MULTIPLET CONTRIBUTIONS

It is clear from the foregoing discussion that many of the properties of transition atom impurities in semiconductors involve a convolution of many-electron (multiplet) effects and one-electron (mean-field) effects. Separating those two contributions must be our next task. This will establish if the understanding of the properties of transition atom impurities requires the intellectual tools pertinent to band theory of solids and effective-mass impurities (i.e. if mean-field effects are dominant), or those pertinent to atomic physics and Mott insulators (if many-electron effects prevail).

In a recent work, Fazzio et al (FCZ) constructed a method for deconvoluting the observed excitation and ionization energies of Equations 1–10 into a component derived from mean-field (MF) effects and one due to multiplet corrections (MC). Specifically, the intra-center excitation energy

$\Delta^{(ij)}[e^m t^n; e^{m'} t^{n'}]$ of Equation 1 was formulated as

$$\Delta^{(ij)}[e^m t^n; e^{m'} t^{n'}] = \Delta_{eff}[e^m t^n; e^{m'} t^{n'}] + [E^{(j)}_{MC} - E^{(i)}_{MC}]$$
$$\equiv \Delta_{eff}[e^m t^n; e^{m'} t^{n'}] + \Delta E^{(ij)}_{MC}. \quad 11.$$

Here, the first term $\Delta_{eff}[e^m t^n; e^{m'} t^{n'}]$ (the "effective crystal-field splitting") is the mean-field total energy difference that separates the two one-electron configurations:

$$\Delta_{eff}[e^m t^n; e^{m'} t^{n'}] = E_{T,MF}[e^{m'} t^{n'}] - E_{T,MF}[e^m t^n]. \quad 12.$$

The effective crystal field energy Δ_{eff} differs from the classic crystal-field splitting used in the Tanabe-Sugano (TS) model (14–15). The Δ_{eff} includes (*a*) bare-ion contributions, (*b*) covalency effects, i.e. screening by the impurity's electrons, and (*c*) average multiplet effects. All are incorporated in mean-field calculations described in this article; the last two are absent (at least formally) in the classical TS approach. The second term in Equation 12 includes multiplet corrections, where $E^{(j)}_{MC}[e^{m'} t^{n'}]$ and $E^{(i)}_{MC}[e^m t^n]$ are, respectively, the energy of multiplet $|j\rangle$ derived the predominant one-electron configuration $e^{m'} t^{n'}$ and the energy of multiplet $|i\rangle$ derived from $e^m t^n$, both measured relative to the respective multiplet average energies.

Similarly, for the ionization process of Equations 3, 8, and 6, the deconvolution is formulated, respectively, as

$$H^{(\alpha)}_{ij}(-/0) = H^{(\alpha)}_{MF}(-/0) + [E^{(j)}_{MC}(A^-) - E^{(i)}_{MC}(A^0)]$$
$$\equiv H^{(\alpha)}_{MF}(-/0) + \Delta H^{(ij)}_{MC}(-/0), \quad 13.$$

$$E^{(\alpha)}_{ik}(0/+) = E^{(\alpha)}_{MF}(0/+) + [E^{(k)}_{MC}(A^+) - E^{(i)}_{MC}(A^0)]$$
$$\equiv E^{(\alpha)}_{MF}(0/+) + \Delta E^{(ik)}_{MC}(0/+), \quad 14.$$

and

$$E^{(\alpha)}_{j\ell}(-/2-) = E^{(\alpha)}_{MF}(-/2-) + [E^{(\ell)}_{MC}(A^{2-}) - E^{(j)}_{MC}(A^-)]$$
$$\equiv E^{(\alpha)}_{MF}(-/2-) + \Delta E^{(j\ell)}_{MC}(-/2-). \quad 15.$$

Here $H^{(\alpha)}_{MF}(-/0)$, $E^{(\alpha)}_{MF}(0/+)$, and $E^{(\alpha)}_{MF}(-/2-)$ are the differences in mean-field total energies pertinent to a single acceptor, single donor, and double acceptor, respectively. They have expressions analogous to those of Equations 3, 8, and 6, respectively, except that the many-electron total energy E_T is replaced by the mean-field total energy $E_{T,MF}$. The correction terms for many-electron multiplet effects are given in square brackets in Equations 13–15. Here, $E^{(i)}_{MC}(A^0)$, $E^{(j)}_{MC}(A^-)$, $E^{(k)}_{MC}(A^+)$, and $E^{(\ell)}_{MC}(A^{2-})$ are the many-electron shifts of the ground state multiplets $|i\rangle$, $|j\rangle$, $|k\rangle$, and $|\ell\rangle$ of the species A^0, A^-, A^+, and A^{2-}, respectively.

The apparent Mott-Hubbard (16) Coulomb repulsion energies of Equations 9 and 10 can be deconvoluted similarly, yielding

$$U_{ijk}^{(\alpha\alpha)}(A^0) = U_{MF}^{(\alpha\alpha)}(A^0) + \Delta U_{ijk}^{(\alpha\alpha)}(A^0), \qquad 16.$$

and

$$U_{ij\ell}^{(\alpha\alpha)}(A^-) = U_{MF}^{(\alpha\alpha)}(A^-) + \Delta U_{ij\ell}^{(\alpha\alpha)}(A^-). \qquad 17.$$

Here, $U_{MF}^{(\alpha\alpha)}(A^0)$ and $U_{MF}^{(\alpha\alpha)}(A^-)$ are the mean-field Coulomb energies and $\Delta U_{ijk}^{(\alpha\alpha)}(A^0)$ and $\Delta U_{ij\ell}^{(\alpha\alpha)}(A^-)$ are the corresponding multiplet corrections (MC), calculated from the corresponding differences of Equations 13–15.

Next, one has to evaluate the magnitude of the many-electron correction terms in Equations 12–17. In the theory of FCZ (10), the multiplet corrections can be obtained in terms of the known (11) free-ion Racah parameters B_0 and C_0 of the free impurity ions (14), and the three physical parameters characterizing each impurity-host system. This is similar to the outline suggested by Allen (17). These are the orbital deformation parameters λ_e and λ_t of the impurity Wannier orbitals of type e and t, respectively, and the effective crystal-field splitting Δ_{eff} of Equation 11. In the limit of atomically localized impurity states the orbital deformation parameters (measuring the strength of interelectronic interactions in the solid relative to the free ion) approach unity, whereas in the opposite limit of completely delocalized impurity states they diminish to zero.

In principle, λ_e and, λ_t could be calculated from mean-field theory by constructing the impurity Wannier orbitals [in the sense of Kohn & Onffroy (18)] using both the gap level and the resonance wave functions. FCZ took the alternative approach; they evaluated the magnitude of the multiplet corrections underlying the experimental data itself; i.e. deconvoluting the observed transition energies into an MF and an MC, using their multiplet theory with its parameters as an intermediate step. To this end, they fitted the observed intra-center $A^- \rightarrow (A^-)^*$ excitation spectra, obtaining λ_e, λ_t, and Δ_{eff} directly from experiment. Due to the scarcity of data, a single Δ_{eff} value was used, and spin-orbit corrections (which would have introduced more parameters) were neglected. When the number of observed transitions was smaller than the number of parameters, a range of these parameters consistent with the data was determined. The interested reader is referred to Reference 10 for a complete description of the details of the method. We next apply this method to transition atom impurities in semiconductors, and discuss the content of this deconvolution.

Figures 3*a–c* reveal chemically transparent trends in the mean-field parameters. The complicated S-shaped curve of Δ_{CF} obtained in the conventional models (19) and unmatched by state-of-the-art mean-field calculations is replaced by a simple one. This curve shows a minimum at

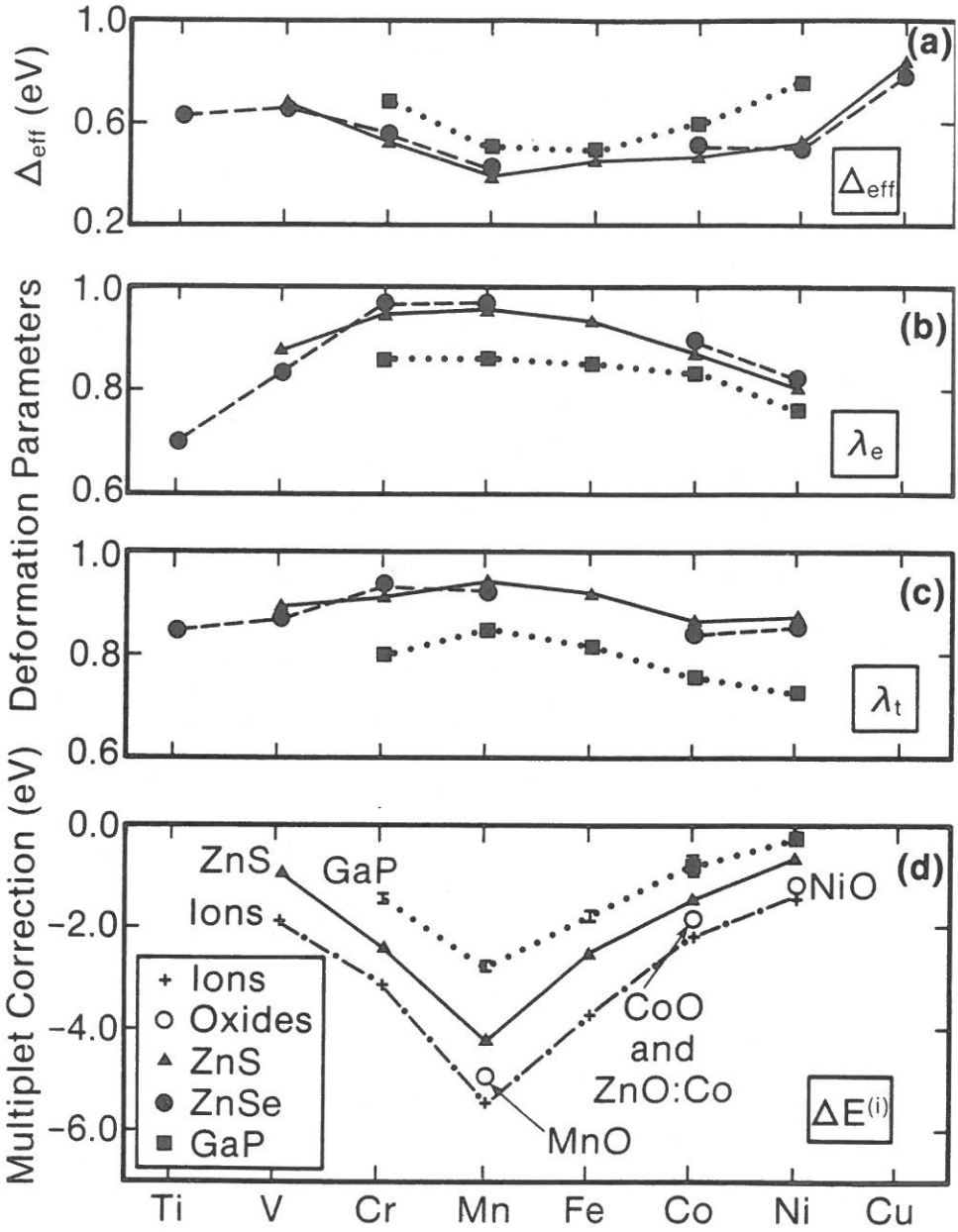

Figure 3 Mean-field parameters obtained by fitting the d → d* absorption spectra (*a*) Δ_{eff}, (*b*) λ_e, (*c*) λ_t, and (*d*) multiplet corrections to the ground-state energy for free ions, bulk oxides, and 3d impurities in semiconductors.

Mn and a maximum for the heavier 3d impurities, as found in recent calculations (9); an increase in Δ_{eff} with covalency (GaP versus II–VI compounds); and a decrease in Δ_{eff} with lattice parameter. The orbital deformation parameters decrease rapidly with covalency (Figures 3*b* and *c*), but are nevertheless substantially higher than those deduced from orbital-integrated charges (20). The MC to the ground-state energies (Figure 3*d*) show a maximum for the high-spin impurity at the center of the 3d series (Mn^{2+}), and a rapid decrease with increasing covalency and decreasing spin at the two ends of the 3d series. Interestingly, we find a similar correction for the impurity system ZnO : Co and the pure bulk CoO (Figure 3*d*), which suggests that most of the ground-state multiplet effects are induced by the nearest ligands. We find that the MC for bulk Mott insulators MnO, CoO (21), and NiO are intermediate between those of the free ions and the impurities in the sulfides (21) (Figure 3d).

Whereas multiplet corrections to the ground-state total energy are relatively small (but sufficient to stabilize the high-spin ground state, cf Figures 3*d* and 4*b*), the corrections for excitations (Figure 4*a*) and ionizations (Figure 4*c*), can be substantial on the physically relevant scale of the optical band gap. Surprisingly, one-electron theory is predicted to describe accurately the lowest excitation energies for all impurities but Mn, even in a relatively ionic host like ZnS (Figure 4*a*). In contrast, MC to the donor and acceptor energies are substantial for all impurities (Figure 4*c*). We next illustrate the variations in the many-electron effects in a more covalent III–V semiconductor (GaP in Figure 5).

Figure 5 displays the mean-field (*shaded*) and MC (*clear*) portions of the first acceptor energies (Equation 13), second acceptor energies (Equation 15), and the lowest $d \rightarrow d^*$ intra-center excitation energies (Equation 11). The solid circles represent the sum $\Delta E_{MF} + \Delta E_{MC}$. The simple chemical trends obtained in the mean-field parameters λ_e, λ_t, and Δ_{eff} [deduced from experiments (10)] result in more complex trends in the many-electron energies. First, on the physically relevant energy scale of the host's crystal band gap ($E_g = 2.35$ eV) the many-electron corrections are seen to constitute a significant portion of the ionization energies, even in the more covalent III–V system. Second, note that the MC can be either negative (e.g. first acceptors in GaP : Cr and GaP : Mn, second acceptor in GaP : Cr) or positive (e.g. first and second acceptors in GaP : Fe and GaP : Ni). When the transition involves an increase in spin (e.g. the $^4T_1 \rightarrow {}^5T_2$ first acceptor in GaP : Cr) the MC is negative, whereas when the transition involves a reduction in spin (e.g. the $^6A_1 \rightarrow {}^5E$ first acceptor in GaP : Mn) the MC is positive. However, even spin-conserving transitions, to which there are no contributions from spin polarization in free ions, are seen to have either small positive (e.g. $^5T_2 \rightarrow {}^5E$ intra-center excitation in GaP : Cr) or small

negative (e.g. $^5E \rightarrow {}^5T_2$ intra-center excitation in GaP:Fe) multiplet corrections.

Finally, note that according to Equations 12–17 the results of Figure 5 also imply that the apparent Mott-Hubbard Coulomb energies and the crystal-field energies have both a mean-field component and a multiplet correction component. For example, we find that for GaP:Cr, GaP:Fe, and GaP:Ni the mean-field contributions $[U_{MF}(A^-)]$ (Equation 17) are 1.04, 1.60, and 1.29 eV, whereas the multiplet corrections $[\Delta U(A^-)]$

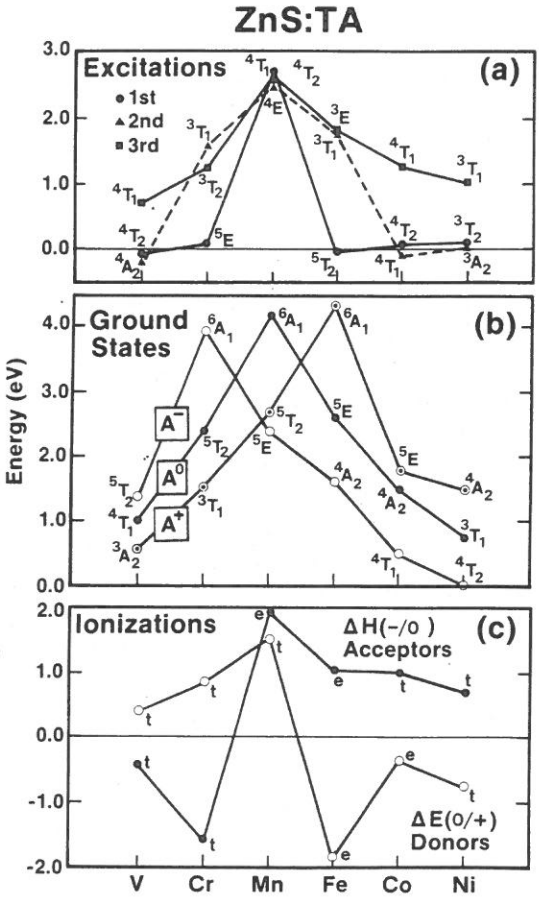

Figure 4 Multiplet corrections for 3d impurities in ZnS. (*a*) Correction to the three lowest high-spin d → d* excitation energies of the neutral (A^0) impurity. (*b*) Negative of the correction to the ground state of the A^0, A^-, and A^+. (*c*) Multiplet corrections for acceptor $[\Delta H(-/0)]$ and donor $[\Delta E(0/+)]$ ionizations. The labels e or t refer to the one-electron orbital being ionized.

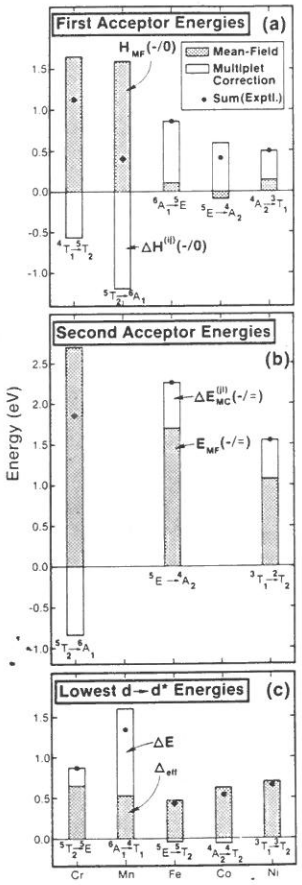

Figure 5 Deconvolution of the experimental transition energies for first acceptors (*a*), second acceptors (*b*), and intra-center d → d* excitation energies (*c*) for TA impurities in GaP into a mean-field contribution (*shaded area*) and many-electron multiplet correction (*clear area*). The *solid circles* denote their sum, i.e. the experimentally observed energies.

(Equation 17) are −0.31, −0.21, and −0.25 eV, respectively. Our analysis of the data is summarized in Figure 1, which displays the different multiplets for both substitutional and interstitial impurities in Si, III–V, and II–VI semiconductors for a series of charge states (A^0, A^-, A^+) and their corresponding oxidation states.

ENERGY LEVEL DIAGRAMS FOR MEAN-FIELD TRANSITIONS

The transition energies, $H_{MF}(-/0)$, $E_{MF}(0/+)$, $E_{MF}(-/2-)$, and Δ_{eff}, and the Coulomb energy U_{MF} correspond to a total energy difference, and hence they could not be readily displayed in terms of one-electron-like energy

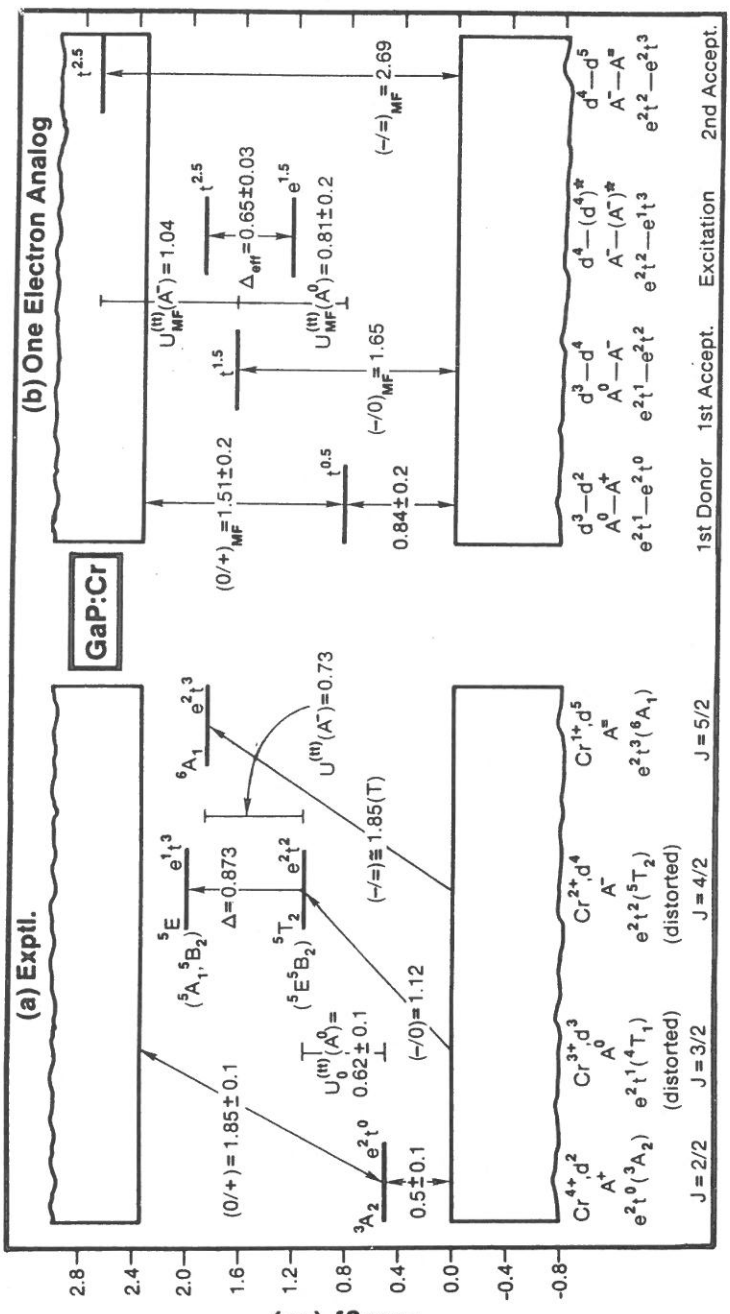

Figure 6 (*a*) Experimentally observed donor, acceptor, and intra-d excitations in GaP : Cr; T denotes tentative values. (*b*) One-electron mean-field analog for donor, acceptor, and intra-d excitations.

level diagrams. Having identified the particular transitions $e^m t^n \rightarrow e^{m'} e^{n'}$ attendant upon the relevant transitions, one could construct approximate mean-field energy level diagrams using Slater's transition state concept (22). The difference in total energies attendant upon a configuration change $V^Q t^m \rightarrow V^{Q'} t^{m'}$, correct to second order, is given by the difference in the corresponding self-consistent density-functional orbital energies, computed at the intermediate occupation numbers $(Q+Q')/2$ and $(m+m')/2$:

$$E_T[V^{Q'} t^{m'}] - E_T[V^Q t^m] \cong [\varepsilon_v - \varepsilon_t] V^{(Q+Q')/2} t^{(m+m')/2}. \qquad 18.$$

The calculation strategy is as follows: We calculate ΔE_{MF} by performing separate self-consistent Green's function calculations for each transition state configuration pertinent to the various donor and acceptor transitions. We assume throughout the calculation the unrelaxed geometry of the ideal crystal. We then compare our results for ΔE_{MF} with the experimentally deduced mean-field energies $\Delta E - \Delta E_{MC}$. The discrepancies with experiments, where they exist, reflect a combination of the amount by which the transition state local density model fails to be a perfect MF theory, and the change in lattice distortions attendant upon the particular transition that was left out.

The mean-field excitation and ionization energies deduced from experiments are illustrated for GaP:Cr in Figure 6, with their corresponding transition state occupation numbers. For example, the total energy difference attendant upon the first acceptor transition in Ga:Cr [(−/0) in Figure 6*a*], i.e. $[A^0, Cr^{3+}, {}^4T_1(e^2t^1)] \rightarrow [A^-, Cr^{2+}, {}^5T_2(e^2t^2)]$, is represented within the mean-field transition state approximation [(−/0) MF in Figure 6*b*] as the energy separation between the t level and the top of the valence band, using the intermediate $(VB)^{5.5} e^2 t^{1.5}$ configuration for the self-consistent calculation.

METHOD OF CALCULATION OF MEAN-FIELD ELECTRONIC STRUCTURE

Having isolated the mean-field part of the observed excitation and ionization energies of deep impurities, our next task is to calculate these energies from self-consistent electronic structure theory. To do this we use the quasi-band crystal field (QBCF) Green's function method developed by Lindefelt & Zunger (23). We do not give the computational details here; the interested reader is referred to Reference 23 for details. This is a first-principles method whose input is the host crystal band structure and the ab initio nonlocal pseudopotential of the impurity, and whose output is the gap level and resonance states of the impurity, as well as its wave functions and charge densities in the impurity subspace. Extension of this method

(24, 25) also provides the Hellmann-Feynman forces acting on the impurity, and hence its breathing-mode relaxation. This is discussed further below.

The QBCF method provides highly accurate self-consistent solutions to the problem within the local density framework. We have previously shown the highly stable convergence properties of this method (23). We have demonstrated that when applied to problems treated by other first-principles Green's function methods (26) [the Si vacancy (23), Si : S (27), and Si : Zn (28)] the same results are produced. We have also applied QBCF to numerous other systems [O and Se in Si (27); substitutional (28–30) and interstitial (31, 32) 3d impurities in silicon; comparison of Cu, Ag, and Au impurities in silicon (33); 3d core exciton in GaP (34); and 3d impurities in GaP (9)] for which no other application with any first-principles Green's function method exists.

We focus in this review on results obtained through a self-consistent Green's function model. We do not dwell on the pioneering cluster model calculations (20, 35–38) using the multiple scattering $X\alpha$ method. These results were reviewed recently by Watkins (39), and were subjected to close comparison with the Green's function results elsewhere (9, 28, 29). I refer the reader to these references for further details.

CALCULATED MEAN-FIELD DONOR, ACCEPTOR, AND COULOMB ENERGIES

Figure 7 depicts the self-consistently calculated impurity-induced e and t energy levels for GaP : TA in the vicinity of the fundamental band gap using the occupation numbers pertinent to the transition state H(−/0) first acceptors. Along with the calculated levels (*open circles*), the mean-field energies deduced from experiment are shown (*solid circles*), and arrows indicate the correspondence between experiment and theory. Two types of impurity-induced levels in the gap appear in this figure: the upper t dangling bond hybrid (DBH) and a lower energy crystal-field resonance (CFR) of type e. Whereas the t^{DBH} is confined to the band gap region, starting in GaP : Zn as a shallow acceptor at $E_v + 0.02$ eV and ending in GaP : Cr as a deep acceptor at $E_v + 2.02$ eV, the e^{CFR} level exists as a valence band resonance for TA = Ga, Zn, Cu, Ni, and Co, and emerges first into the band gap for GaP : Fe.

The analysis of the data shows that the first acceptor transitions for the impurities Zn, Cu, Ni, Mn, and Cr involve the ionization of the t^{DBH} electron, whereas the acceptor states of Co and Fe involve ionization of the e^{CFR} electron. This switch between Ni and Co, or Fe and Mn cannot be deduced without acknowledging many-electron effects. Comparing the calculated and experimentally deduced values (9, 10) (given here in

parenthesis) for the first acceptor levels associated with the t^{DBH}, we find for GaP : Zn, GaP : Ni, GaP : Mn, and GaP : Cr the values 0.02 (0.07), 0.26 (0.14 to −0.24), 1.74 (1.59), and 2.02 (1.65) eV, respectively. Notice that for the d^9 system GaP : Zn the multiplet correction for the first acceptor state vanishes, hence the mean-field prediction should be directly comparable with experiment [after the small effective mass-binding energy ~0.05 eV (8) is added to the central cell term of 0.02 eV to correct in a rough way for electrostatic effects].

For the acceptor levels associated in our analysis of the experimental data with the e^{CFR} level, we find the values −0.20 (−0.08) for GaP : Co, and 0.13 (0.10) eV for GaP : Fe. The agreement with experiment is seen to be good except for GaP : Cr and possibly GaP : Ni, for which the calculated positions of the t levels are about 0.4 eV too high. Interestingly, the two

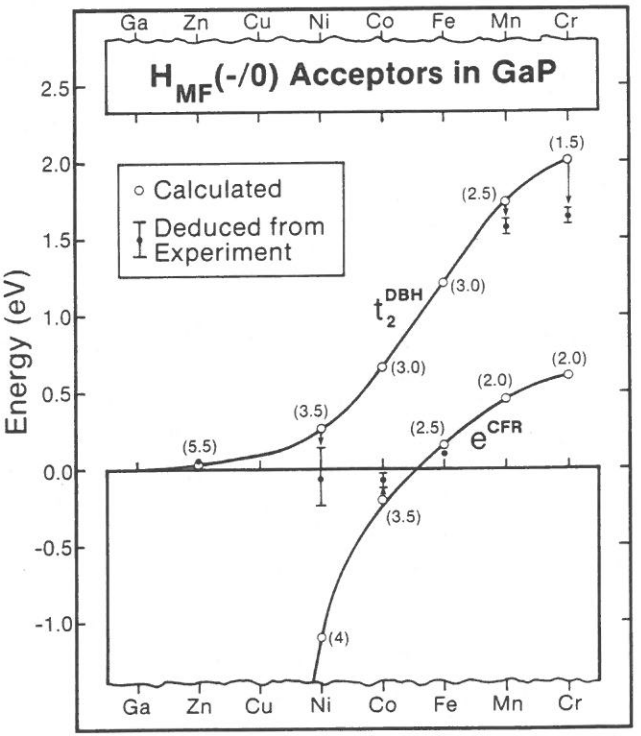

Figure 7 Calculated near-gap impurity-induced acceptors (*open circles*), compared with the mean-field values deduced from experiment (*solid circles*). Numbers in parentheses indicate the transition state occupation numbers.

impurities for which our (lattice unrelaxed) calculation shows the largest discrepancy with experiment are also known to involve lattice relaxation. GaP:Cr is known to be distorted both in its A^0 and A^- states, and GaP:Ni is known to be distorted at least in its A^- state. It is therefore tempting to associate at least part of the discrepancy between theory and experiment for these two impurities to lattice relaxations.

Turning next to double acceptor states, the calculated positions of the corresponding t^{DBH} levels compared with the experimentally deduced values (given here in parenthesis) are 1.26 (1.06 to 1.37) for GaP:Ni, and 2.58 (2.69) for GaP:Cr; whereas the value for GaP:Fe, pertinent to an ionization of the e^{CFR} level, is 1.53 (1.70) eV. The Coulomb repulsion energies $U_{MF}^{(tt)}(A^-)$ are hence 1.26–0.26 = 1 eV for GaP:Ni (experimentally deduced value is 1.23–1.29 eV); 2.58–2.02 = 0.56 eV for GaP:Cr (experimentally deduced value is 1.04 eV); and the $U_{MF}^{(ee)}(A^-)$ for GaP:Fe is 1.53–0.13 = 1.4 eV (experimentally deduced value is 1.6 eV). Our calculated second acceptor energies are hence within 0.2 eV from experiment, and the calculated Coulomb repulsion energies are between 0.2 and 0.5 eV lower than the experimental values. This suggests that the A^- state (e.g. Cr^{2+}) is stabilized by lattice distortion more than the A^{2-} state (e.g. Cr^{1+}). The largest discrepancy occurs for GaP:Cr (0.48 eV). Indeed, its A^- state is known to be distorted, whereas the A^{2-} state is a symmetric $^6A_1$ state and hence not subject to Jahn-Teller (JT) distortion.

TRENDS IN ONE-ELECTRON LEVELS FOR SUBSTITUTIONAL IMPURITIES

Figures 8 and 9 depict the calculated impurity-induced energy levels throughout the energy range for substitutional TA in Si and GaP host crystals, respectively. The overall trends are similar in these two materials. The tenfold degenerate atomic 3d level splits in the crystal into a fourfold degenerate e-type and a sixfold degenerate t-type crystal field resonance (CFR). This pair appears near the bottom of the valence band (VB) for Si:Zn and GaP:Zn as atomic-like, highly localized, d-like states. As one progresses to lighter impurities, these VB resonances delocalize and hybridize with the host p-states, and their energy and (crystal-field) splitting increase. The highly localized e^{CFR} level first penetrates the band gap for Si:Mn or GaP:Fe ("e threshold"), whereas the t^{CFR} level, repelled by the large density of host t states at $\sim E_{VB} - 1$ eV, is pinned inside the VB and never makes it to the gap.

The antibonding counterpart of t^{CFR} is the dangling bond hybrid (DBH) level t^{DBH}. It is a hybrid of the four neighboring Si (or phosphorus) dangling bonds and a smaller admixture of impurity atomic orbitals. It starts in

Si:Zn or GaP:Zn just above the VB maximum as a delocalized, predominantly p-like acceptor state. As one moves to lighter impurities the energy of t^{DBH} increases smoothly, and the state disappears for the first time into the conduction band for Si:Co ("t threshold"). Due to the larger band gap of GaP, it stays inside the gap at least up to GaP:Cr. The level is then pinned inside the conduction band by the repulsion (avoided crossing) from the large density of t host states at $\sim E_{VB} + 3.8$ eV.

Figure 10 shows the decomposition of the t^{DBH} wave function for Si:TA into its angular-momentum components $G^t_\ell(|\check{r}|)$ for the first two symmetry-allowed terms $\ell = 1$ and 2 (28). For comparison, a partial wave-function resolution is also given for the silicon vacancy (*upper left panel*, Figure 10). The significant fingerprint feature of the t^{DBH} state for all transition atom impurities in silicon is a strongly localized atomic-like $\ell = 2$ component, which has a node in the inner central cell region, and an opposite sign relative to the $\ell = 1$ wave-function component. This behavior is distinctly different from that observed in the dangling-bond wave function of the silicon vacancy, which has a delocalized $\ell = 2$ component with a node only

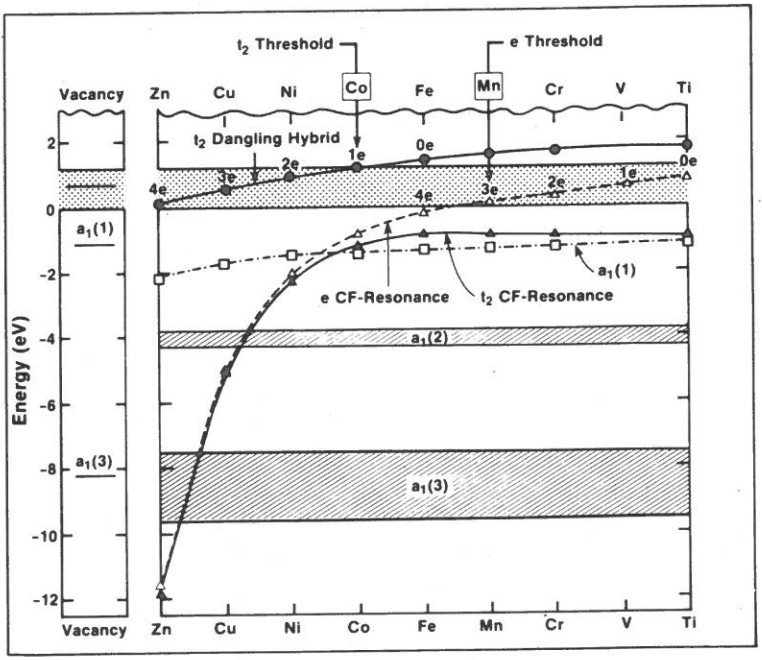

Figure 8 Calculated defect energy levels for neutral, substitutional, unrelaxed transition atom impurities in silicon. The numbers 4e, 3e, 2e, etc indicate the occupation numbers.

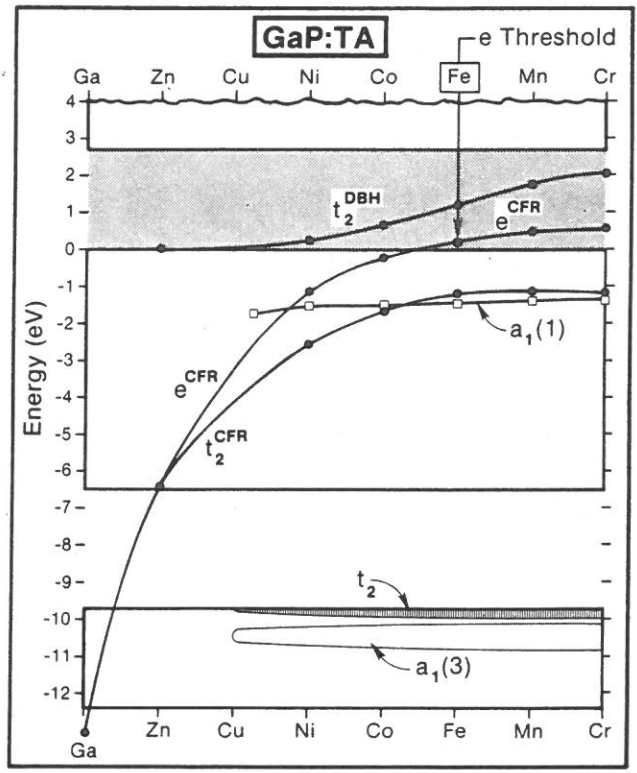

Figure 9 Calculated impurity-induced energy levels in GaP:TA, using transition state occupation numbers pertinent to first-acceptor states.

outside the central cell region, and the same sign as the $\ell = 1$ wave-function component in the inner region.

The amount of p–d hybridization in the DBH of Si:TA is shown in Figure 11, where the percentage of p and d character is depicted across the 3d series. The DBH begins at Si:Zn as a predominantly p state (coupling with p-like valence band states). It reaches a 50%–50% p–d character near the center of the 3d series, after which the p character increases a little because of the renewed availability of nearby p-like host states (this time from the conduction bands). Note that a t level such as the DBH can by no means be considered purely p-like.

STABILITY OF CHARGED STATES

It has long been known that the energy levels of different ionized states of a free ion span a large range, and that a nontransition element existing as an

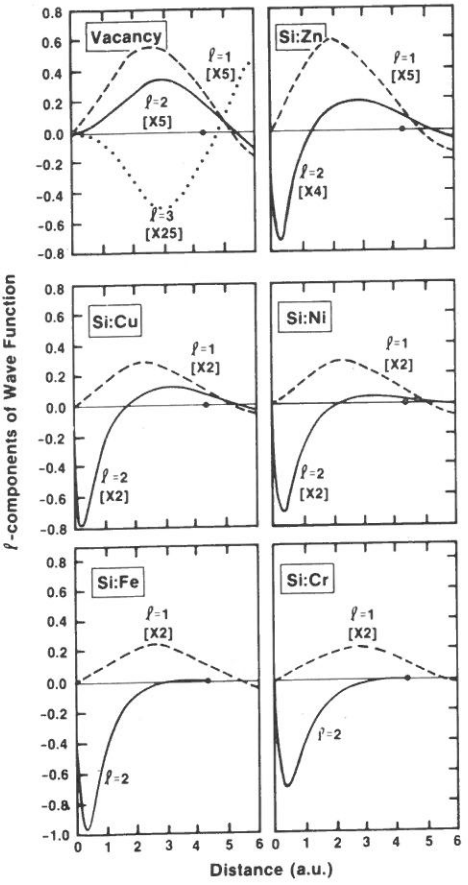

Figure 10 Angular momentum components $G_\ell^2(|\vec{r}|)$ for the t_2 dangling-bond hybrid wave functions of substitutional 3d impurities in silicon and the silicon vacancy.

impurity (Si : S, Si : P, etc) can sustain a number of ionized states in a very narrow energy range (40), often ~ 0.1 eV. Both phenomena are understandable. The weak potential perturbation of an s-p impurity in an s-p host crystal leads to defect wave functions that extend over many lattice constants. Hence, even if the intra-atomic Coulomb energy per electron (U) is not small, the orbital energy (ε_{gap}) of this extended orbit varies slowly with its occupation number (N), in proportion to the formal net charge Q_{net}. Conversely, for free atoms the localized nature of the bound orbitals and the correspondingly large free atom Coulomb energies (U^0) lead to a rapid variation in the orbital energy with the amount of electronic charge $Q(N)$ on the atom,

$$\varepsilon(N) \cong \varepsilon_0 + UQ(N). \qquad 19.$$

For example, the d^5–d^2 ionization energies of Mn are 33.7, 51.2, 72.4, and 95

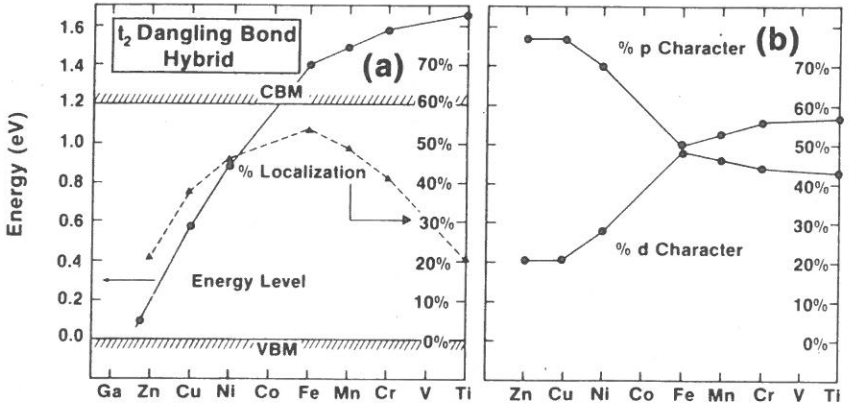

Figure 11 (*a*) Variations in percentage of localization and energy of the dangling-bond hybrid level of TA in silicon. (*b*) Variations in percentage of p- and d-character in the dangling-bond hybrid wave function.

eV, respectively, leading to $U_d^0 \approx 20$ eV/e. Surprisingly, however, when a TA exists as an impurity in a semiconductor it can sustain a large number of different charge states in a narrow energy range, despite the fact that the gap level is highly localized (e.g. 82% localization for Mn). Our calculation explains this remarkable behavior in terms of a self-regulating response (41).

The total charge $Q(N)$ in the central cell due to the impurity atom alone is calculated as

$$\int_0^{R_{cc}} [\rho_D(\vec{r}) - \rho_V(\vec{r})]\, d\vec{r}, \qquad 20.$$

where we subtract the vacancy charge $\rho_V(\vec{r})$ from the defect charge $\rho_D(\vec{r})$ to get the effective transition atom charge in the solid for each value of N. This central cell charge, denoted here as $Q(N)$, represents the total number of valence electrons an impurity atom has when it is incorporated substitutionally. It depends on the number of electrons (N) occupying the gap level. If $N = N_0$, the system as a whole has a zero formal charge ($Q_{net} = 0$) (e.g. $N_0 = 3$ for Si : Mn). We then separate $Q(N)$ into the contribution from the occupied valence bands of the perturbed system $Q_{VB}(N)$, and the contribution from the gap level alone, $Q_{gap}(N)$

$$Q(N) = Q_{VB}(N) + Q_{gap}(N); \qquad Q_{gap}(N) = \sum_{i,\alpha}^{gap} N_i q_i^\alpha. \qquad 21.$$

Figure 12*a* shows the variations in $Q_{VB}(N)$ and $Q_{gap}(N)$ with N for Si : Mn,

as obtained by performing separate self-consistent calculations for each value of N. Whereas $Q_{gap}(N)$ increases nearly linearly with N as expected, the valence band contribution $Q_{VB}(N)$ decreases with N. This leads to a variation of $Q(N)$ with N considerably slower than that of $Q_{net} = N - N_0$ or $Q_{gap}(N)$ alone. Consequently, the energy of the gap levels (ε_{gap}) (Figure 12*b*) varies very slowly with occupation, permitting many charge states to exist within a narrow range. Much like a pool below a waterfall, as electrons are poured into the gap level the valence band states rearrange themselves, "leaking out" of the central cell to partially compensate for the disturbance. This self-regulating behavior of the system in response to an external perturbation is remarkably analogous to the homeostatic control mechanisms in biology or to servomechanisms in machine operations. In such systems, despite the existence of dynamic perturbations a stability of operation within a prescribed range of operating parameters is maintained through a (often nonlinear) self-regulating feedback response. This is possible only in an open thermodynamic system such as Si:TA or GaP:TA, not in a closed system such as an isolated transition atom. In the latter case, the effect of added electrons must be absorbed by the atom itself

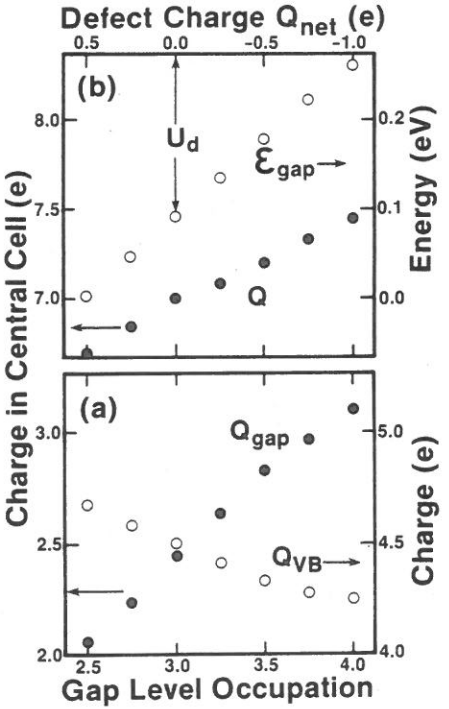

Figure 12 (*a*) The contributions of the gap level $Q_{gap}(N)$ and the valence band $Q_{VB}(N)$ to the central cell charge Q of Si:Mn as a function of the gap level occupation N. (*b*) Variation in central cell charge Q and gap level energy e_{gap} with N.

through an increased delocalization of its wave functions, and a rapid change of the ionization energy with $N - N_0 = Q_{\text{net}}$. I suspect a similar self-regulating behavior occurs in many other systems where localized states characterized by large correlation energies coexist in a similar energy domain with an itinerant manifold of states with which they can exchange charge. Surface states and valence fluctuation states are just a few possible examples.

The foregoing analysis suggests that although in the neutral charge state much of the amplitude of the impurity wave function resides on the ligands, in the ionized state some of this amplitude moves back from the ligands into the impurity sphere, compensating for most of the charge removed by ionization. Therefore, we expect the charge density at the impurity site will be nearly constant for different ionized states (as could presumably be detected through the Mössbauer isomer shift), and most of the change will be detected on the ligands (as presumably could be observed by ENDOR). These predictions await experimental testing (i.e. Mössbauer and ENDOR experimental results as a function of doping).

EFFECTIVE ELECTRONIC CONFIGURATIONS

The small U_d values for TA impurities in semiconductors have an interesting implication for the effective electronic configuration of the impurity atom. Whereas the large atomic U_d^0 values lead to the well-known preference for occupying the 4s subshell before the 3d subshell is completed, small U_d values in the solid may lead to a population inversion in the ground state.

Figure 13*a* depicts the way the atomic $a_1^1(et_2)^{n+1}$ or $a_1^2(et_2)^n$ ground-state configuration in Si : TA is redistributed when the atom is placed in the solid. Most of the atomic a_1 electrons are depleted and redistributed into the e and t_2 representations. The e representation contains nearly four electrons for Zn through Fe, then 3, 2, 1, and 0 electrons for Mn, Cr, V, and Ti, respectively, as the e^{CFR} moves into the gap and is progressively emptied. The decrease in the occupation from Zn to Fe for the t_2 representation, followed by a constant occupation of approximately four electrons thereafter, reflects the evacuation of the t_2^{CFR} level as it moves through the gap.

Figure 13*b* shows the variation with atomic number of the effective orbital configuration (ΔQ_ℓ) of the impurity atom. The total number of impurity valence electrons Z_β^v is shown for comparison (*dashed line*). The interesting result shown in this figure is that whenever the 3d subshell can also accommodate the s electrons (i.e. to the left of Ni d^8 in the periodic table) the latter are promoted into it. This trend is evident in Figure 13*b*

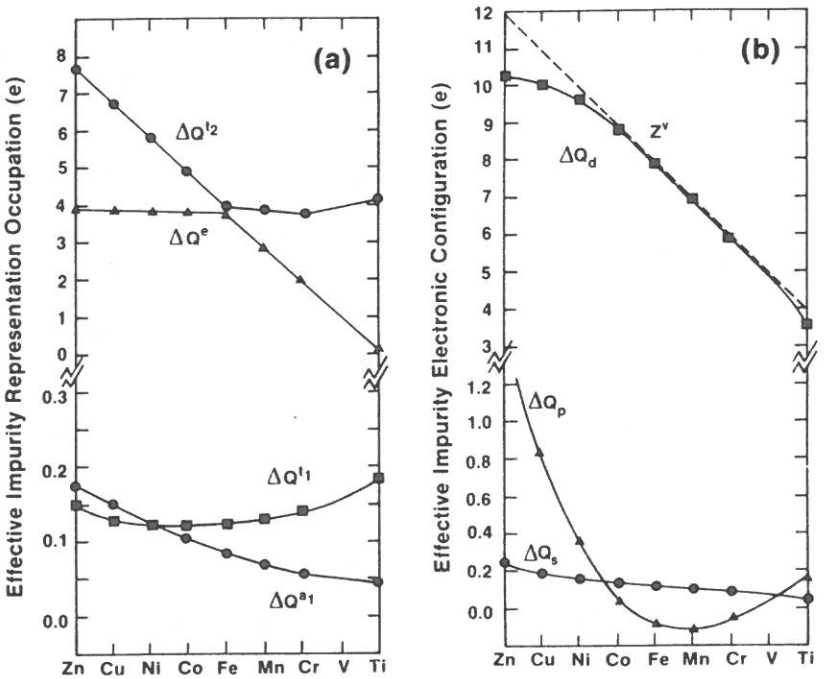

Figure 13 Electronic configuration of an effective transition atom embedded substitutionally in silicon. (*a*) Representation occupations, and (*b*) orbital occupation ΔQ_ℓ. For comparison, the total number of valence electrons Z_β^v is shown as a dashed line.

from the approach of the ΔQ_d line to the Z_β^v (*dashed*) line. Thus, Fe, Mn, Cr, and Ti have 6, 5, 4, and 2 valence d electrons in the free atom, but they have about 8, 7, 6, and 4 d electrons, respectively, when placed substitutionally in silicon. Substitutional impurities in silicon therefore tend to approach a noble metal configuration. This is in marked contradiction to the Ludwig-Woodbury model (40), which hypothesized that for substitutional 3d elements the 3d electrons would be promoted into the sp subshell to form a tetrahedral hybrid (i.e. $d^n s^2 \rightarrow d^{n-2} s^1 p^3$, or $d^{n+1} s^1 \rightarrow d^{n-2} s^1 p^3$).

INTERSTITIAL 3d IMPURITIES IN SILICON: NONMAGNETIC SOLUTIONS

In III–V and II–VI compounds 3d impurities normally occupy the cation substitutional site, but in silicon 3d impurities can also occupy the tetrahedral interstitial sites. We have applied the QBCF model to such

impurities (31, 32). We first discuss the results obtained in a non-spin-polarized fashion. The defect energy levels introduced into the band gap by the neutral interstitial impurities are shown in Figure 14. The occupation numbers (in parenthesis) have been chosen to be consistent with observed EPR data (40), i.e. to favor high-spin states. We see that Ni is predicted to be electrically inactive in the ground state because there are no partially filled gap states, and that only the one-particle levels of e-symmetry of Co and Fe are electrically active. Furthermore, the crystal-field splitting of the gap states, defined as the energy separation between the bound e- and t_2-states, increases slightly towards the lighter end of the series, reflecting a decreasing localization of the corresponding wave functions. Actually, it is the gap states of e-symmetry that become more delocalized for the lighter elements ($q_e = 0.58$, 0.55, 0.44, and 0.23 for Co, Fe, Mn, and Cr, respectively), while the t_2–gap states have roughly the same localization ($q_{t2} = 0.36$ for Mn and 0.38 for Cr). As to the amount of d-character in the

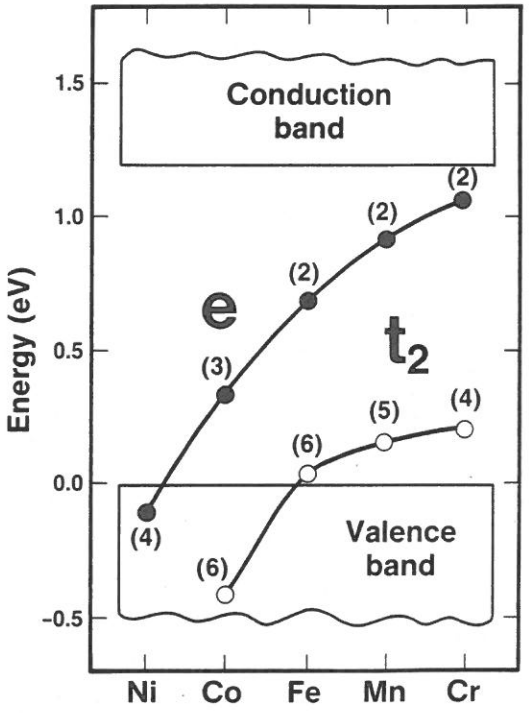

Figure 14 Defect energy levels around the band gap introduced by interstitial 3d impurities in silicon. The numbers in parenthesis are the occupation numbers.

gap-state wave functions inside the central cell (measured by $q_{i\ell}/q_i$ for i = gap and $\ell = 2$), we find that the e-states have almost 100% d-character and that the t_2-states have more than 90% d-character.

Figure 15 is an example of a wave function; it shows the gap-state wave function of e-symmetry for Si : Fe in the $\pm\langle 1,1,0\rangle$ crystal directions. The shaded area indicates the part of the wave function that corresponds to around 50% of the normalization integral. The wave function is clearly very atomic-like in the inner central cell (amost 100% atomic d-character, as mentioned earlier), but its amplitude inside the core is reduced by a factor 0.7 compared to the free atom. Around 50% of the wave function must therefore reside outside the central cell, where its amplitude is seen to be relatively small. This indicates that the wave function is rather delocalized with a large atomic-like peak near the origin. Therefore, we expect this state will show features characteristic of both a localized state, for example, multiplet splitting and hyperfine interaction, and a delocalized state, e.g. many charged states in a narrow energy region such as the band gap, and reduced g-factors. Both types of features have been observed experimentally.

We can picture an interstitial 3d impurity as an essentially spherically symmetric cloud of charge, which fills the rather empty region around the tetrahedral interstitial position on out to the nearest-neighbor atom without affecting the bonds or charge distribution in the rest of the crystal. This is very different from the impurity-induced charge rearrangement of

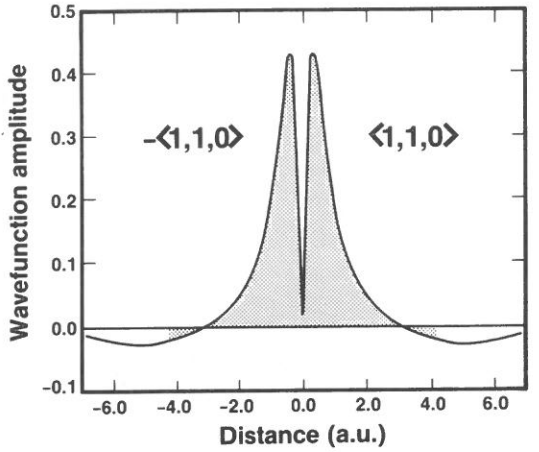

Figure 15 Bound-state wave function of e-symmetry for interstitial Si : Fe in the $\pm\langle 1,1,0\rangle$-directions. The *shaded area* corresponds to 50% of the normalization integral.

the substitutional transition atom impurities and the vacancy (28). In those cases the rearrangement is localized to the central cell, but there is a substantial pertubation (weakening) of the bonds due to the deficiency of valence electrons to repair the broken bonds (29).

By analyzing the quantities $q_{i\ell}$, which describe how the various ℓ-components of the occupied wave functions contribute to the total charge around the impurity, we find that when the atomic 3d shell is not completely filled the atomic 4s electrons go into the 3d shell, i.e. the atomic $d^n s^m$ configuration becomes effectively a $d^{n+m}s^0$ configuration in the solid. This s–d population inversion (30) as compared to the free atom was already envisioned by Ludwig & Woodbury (40) in their classical model. The s–d population inversion suggests a simple explanation of the high diffusion constant D for Ni in silicon ($D \sim 10^{-4}$ cm^2/s) relative to the lighter TA impurities ($D \sim 10^{-8}$–10^{-10} cm^2/s). In the solid the Ni atom effectively assumes a closed-shell, noble-atom-like d^{10} configuration. One would therefore expect Ni to diffuse easily through the solid.

SPIN-POLARIZED CALCULATION FOR INTERSTITIAL 3d IMPURITIES IN SILICON

The larger band gap of III–V and II–VI semiconductors relative to Si places many excited $d \rightarrow d^*$ transitions within spectroscopic reach for the compound semiconductors. This allowed us to deduce from experiment the many-electron part of such transition energies, and hence to focus our QBCF calculations on the mean-field part alone. This was not the case in Si. So far no $d \rightarrow d^*$ transitions have been observed in this system. For Si we could not use the method of FCZ to unravel the many-electron contributions from the mean-field contributions. Consequently, we adopted a different strategy for this system. We neglected spatial correlation effects and focused only on spin-correlation contributions. Katayama-Yoshida & Zunger performed spin-polarized local spin density (LSD) calculations for all impurities in Si between Sc and Cu (42). For this particular application we modified our approach as follows: (*a*) An all-electron potential, rather than a pseudopotential was used to describe the impurity core, because core-derived hyperfine interactions were of interest; (*b*) the 1s through 4f impurity-basis orbitals were optimized iteratively and nonlinearly, allowing them to adjust to the changes in the self-consistent potential obtained from the Green's function problem; (*c*) we self-consistently applied the self-interaction correction to the LSD in a fully spin-polarized manner (43); (*d*) we applied the adspace method of Williams et al (44) to obtain a variationally complete Green's function.

Figure 16 displays for Si: Fe^0 the changes $\Delta n^{\sigma}_{\Gamma,\alpha}(\varepsilon)$ in the local density of

states for spin $\sigma = \pm$ and representation $\Gamma = a_1$, e, and t_2, projected on the local orbitals $\alpha = 3d$, 4s, and 4p. We find a spin-down antibounding t_2-level in the gap at $E_v + 0.26$ eV, and a spin-up e-level above it at $E_v + 0.44$ eV. The latter contributes a local (as defined by the impurity-basis orbitals) magnetic moment of $\mu_e^{gap} = 0.384\ \mu_B$, and the former contributes $\mu_t^{gap} = -0.576\ \mu_B$. In ionic crystals the 3d impurity orbitals are isolated exclusively as gap states (45). In contrast, we find that about 80% of the 3d amplitude in the impurity orbital subspace in silicon exists as localized resonances inside the valence band, showing bonding (b) and antibonding (a) resonances (R) dominating the magnetism ($\mu^{VB} = 1.61\ \mu_B$ versus $\mu^{gap} = -0.19\ \mu_B$) and the

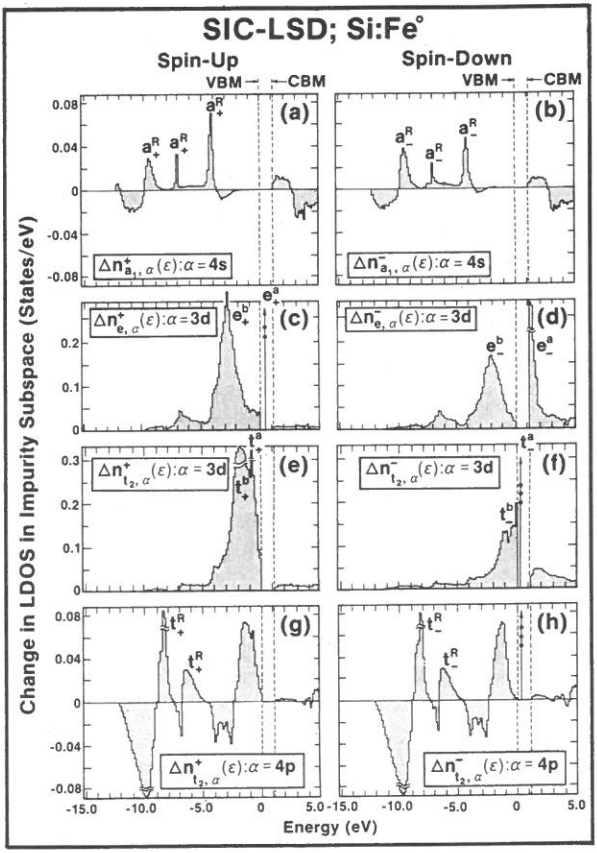

Figure 16 Changes $\Delta n^{\sigma}_{\Gamma,\alpha}(\varepsilon)$ in the local density of states (LDOS) for spin σ (+ or −) and representation $\Gamma = a_1$, e, and t_2. We show only the major orbital components: $\alpha = 4s$ for a_1, 3d for e, and 3d and 4p for t_2.

impurity charge (6.76 out of 8.0 valence electrons are in the valence band and 18 are in the core).

Few conclusions can be drawn from the results on Si:Fe^0. First, although the total magnetic moment (calculated from the phase shifts) is 2.0 μ_B, much as in the d^8 free atom, the local magnetic moment in the impurity orbital subspace is only 1.42 μ_B, and the remaining $[(1-1.42/2.0) \times 100\%]$ 29% of the spin density [as indicated by ENDOR (46)] is delocalized through the crystal, showing substantial covalency. A population analysis of the magnetic moment shows that it evolves primarily from the valence d orbitals [moment distribution of $(\text{core})^{0.00}\ 4s^{0.00}\ 4p^{0.01}\ 3d^{1.29}\ 4d^{0.10}\ 4f^{0.00}$], and is shared in a 5:1 proportion between the e and t_2 valence representations [moment distribution of $(\text{core})^{0.00}\ a_1^{0.01}\ e^{1.17}\ t_2^{0.24}$].

Second, in contrast to the dominant role of the valence band states in determining the local moment, the hyperfine field (H_{hf}) is determined largely by the core states. The contribution to the Fermi contact interaction (both direct and core polarization) from the core states is $-24.50-278.52+179.70 = -123.32$ kG for the 1s, 2s, and 3s states, respectively, and that from the 4s valence band states is only $+8.18$ kG, yielding a total contact term of -115.14 kG. The contribution from the interaction with the electron orbital magnetic moment is estimated from the g-shift to be $+5.02$ kG. Neglecting the small dipolar term, the total calculated H_{hf} of -110.12 kG is reasonably close to the observed values of $|H_{hf}| = 147.6$–152.2 kG (40) (only absolute values are known from experiment). We find a spin density at the Fe nucleus of $\delta\rho(0) = -0.220\ \text{au}^{-3}$ (where au is an atomic unit), in reasonable agreement with the observed values (46) of $|\delta\rho|(0)| = 0.282$–$0.299\ \text{au}^{-3}$. The negative H_{hf} and $\delta\rho(0)$ result from the preferred expulsion of spin-up density from the core region by the Coulomb repulsion between the core electrons and the unpaired 3d electrons.

Finally, in agreement with the nonmagnetic calculation (32) and with Ludwig & Woodbury (40), a population analysis of the orthogonalized states reveals that the s-electrons are largely transferred into d-orbitals [effective configuration: $(\text{core})^{18.0}\ 4s^{0.06}\ 4p^{-0.13}\ 3d^{8.18}\ 4d^{-0.38}\ 4f^{-0.00}$], and that in the impurity subspace iron has a small positive net charge (nuclear charge minus electronic charge of $Q_{net} = 0.276$ e), indicating a small Fe-to-Si ionic charge transfer.

The characteristics of Si:Fe^+ are different. We describe Si:Fe^+ using Slater's transition state approximation (22): We remove half an electron from either the e_+ or the t_- gap levels, placing the ionized charge in the conduction band minimum, and seek a new self-consistent solution. We find that although the highest occupied orbital of the neutral system is the e_+ state, the lowest energy ionization proceeds instead from the t_- orbital. Remarkably, upon ionizing t_- the e_+ orbital moves down in energy,

exposing the t_- as the highest occupied state of the ion, and thereby producing the observed $S = 3/2$ high-spin (weak-field) configuration as Δ_x/Δ_{CF} switches from below unity for $Si:Fe^0$ to above unity (1.4) for $Si:Fe^+$. It is important to emphasize that without the self-interaction correction the correct high-spin ground state is not obtained (47) unless the spherical approximation to the potential is involved (13). We find a donor energy $\varepsilon(t_-) - \varepsilon_{VBM} = E_v + 0.32$ eV, in good agreement with the observed value of $E_v + 0.385 \pm 0.01$ eV (48, 49). In agreement with experiment (48), no stable acceptor state (Fe^0/Fe^-) exists in the gap, as the empty e_- level that would accommodate the next electron is in the conduction band (Figure 16*d*).

The resolution to the apparent dichotomy (see Introduction) between the

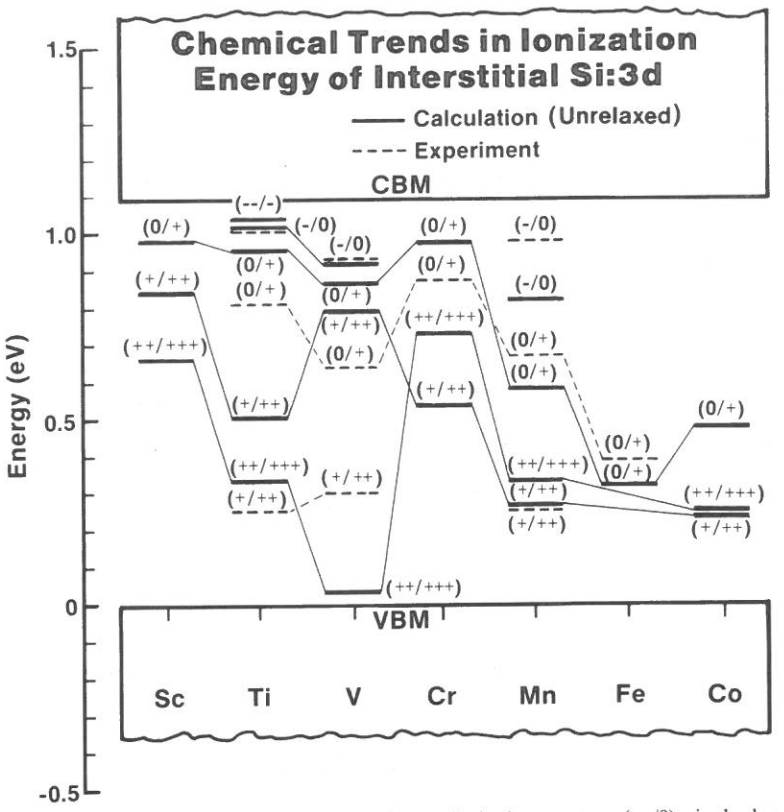

Figure 17 Calculated (spin-polarized and observed) single acceptors (−/0), single donors (0/+), double donors (+/++), and triple donors (++/+++) for interstitial 3d impurities in silicon.

covalently delocalized and atomically localized models for 3d impurities in semiconductors lies, therefore, in this self-regulating response and in the fact that different orbitals are responsible for different aspects of the localization (duality). The contact spin density and hyperfine field are decided by the hyperlocalized core states; the magnetism is largely contributed by the localized valence band resonances; the donor ionization, with its attendant high-spin configuration, and the constancy of the isomer shift (IS) are decided by a combination of delocalized gap states and the feedback (self-regulating) response of the valence band resonances to excitations of the outer states.

We have applied the same calculation technique to all other 3d impurities in silicon (50). We have calculated the acceptor $E_A(0/-)$, as well as the various donor transition energies $E_D(0/+)$, $E_{DD}(+/2+)$, and $E_{DDD}(2+/3+)$. The calculated results are compared in Figure 17 with experiment, showing generally very good agreement. Detailed results for the g-shifts (Δg) and the hyperfine coupling constants A, also showing good agreement with experiment, are discussed elsewhere (50).

POSSIBILITY OF AN EXCHANGE CORRELATION NEGATIVE EFFECTIVE U

Our foregoing discussion suggested that interelectronic coulomb repulsions are reduced in the solid because of strong screening effects, and that many-electron effects further reduce U. This opens up the interesting possibility that these effects could combine to make the effective interaction U become negative even if lattice relaxation alone were insufficient to produce $U < 0$ (51). We discuss this point next.

The mean-field portions of the acceptor transition $E_{MF}^{N,N+1}$ and of the (inverse) donor transition $E_{MF}^{N-1,N}$ can have three physical components, represented by the first three terms in Equations 22 and 23:

$$E_A(0/-) = [\Delta E_{ver}^{N,N+1} + \Delta E_R^{N,N+1} + \Delta E_{JT}^{N,N+1}] + \Delta E_{MC}^{N,N+1}, \qquad 22.$$

and

$$E_D(+/0) = [\Delta E_{ver}^{N-1,N} + \Delta E_R^{N-1,N} + \Delta E_{JT}^{N-1,N}] + \Delta E_{MC}^{N-1,N}. \qquad 23.$$

The first term represents the change in the vertical (ver) total mean-field energy attendant upon the $VB^p t^N \rightarrow VB^{p-1} t^{N+1}$ acceptor transition and the $VB^{p+1} t^{N-1} \rightarrow VB^p t^N$ donor transition when the lattice is kept unrelaxed. This term does not include any distinct many-electron corrections. However, as the wave functions are allowed to relax, it does include the changes in the polarization energies, Madelung energies, and crystal-field

splittings attendant upon the changed screening. The second and third terms, ΔE_R and ΔE_{JT}, represent the changes in the breathing-mode (R) and Jahn-Teller (JT) relaxations, respectively, in the corresponding transitions. The last term represents the many-electron correction.

We define the apparent Mott-Hubbard energy $U^{ij}(N)$ for the N-electron system in the usual way: as the energy required to remove an electron from orbital i on a neutral A^0 center (transforming it to A^+) and placing it in orbital j of a distant A^0 center (transforming it to A^-). In this definition the effective $U^{ij}(N)$ includes Coulomb repulsions, exchange attractions, and screening effects. It equals the difference between the acceptor energy E_A $(0/-)$ and the (inverse) donor energy E_D $(+/0)$, both referred to the same origin:

$$U^{(tt)}(N) \equiv E_A(0/-) - E_D(+/0) = [U_{ver} + U_R + U_{JT}] + \Delta U_{MC}, \qquad 24.$$

where $U_{ver} = E_{ver}^{N+1} + E_{ver}^{N-1} - 2E_{ver}^{N}$; $U_R = \Delta E_R^{N,N+1} - \Delta E_R^{N-1,N}$; $U_{JT} = \Delta E_{JT}^{N,N+1} - \Delta E_{JT}^{N-1,N}$; and $\Delta U_{MC} = \Delta E_{MC}^{N,N+1} - \Delta E_{MC}^{N-1,N}$. The vertical electronic term U_{ver} is positive, reflecting the increased interelectronic repulsion upon adding an electron to the system, despite changes in the screening. For most conventional impurities this is the dominant effect, hence $U(N) > 0$ and the acceptor is higher in the gap than the donor that belongs to the same center. The "Anderson negative effective U" (52) corresponds to the well-known relaxation effect in chalcogenide glasses that stabilizes the electron-rich configuration more than the electron-poor configuration (i.e. the positive U_{ver} is outweighed by the negative $U_R + U_{JT}$), leading to the metastability of this electron-poor configuration. In nonequilibrium situations, e.g. optical excitations, where no lattice relaxation occurs, this mechanism does not apply.

We wish to point out the possibility of a different type of negative effective U, an "exchange correlation negative effective U". It is conceivable that many-electron effects present also in nonequilibrium situations could outweigh (with possible help from $U_{JT} + U_R$) the electronic U_{ver}, leading to an overall $U < 0$. Transition atom impurities in semiconductors are especially relevant in this sense in three ways. First, a self-regulating response (30) of the valence band resonances reduces the MF value U_{ver} dramatically ($U_{ver} \sim 0.5$–1.0 eV) relative to the values in ionic media or free ions ($U \sim 10$–22 eV). Second, JT energies appear to be small. In the best studied case of Cr d^4 ($^5T_2$) in II–VI materials the JT energy is ~ 0.06 eV (53a); for the $^5T_2 \rightarrow {}^6A_1$ transition, $\Delta E_{JT} + E_R \cong -0.3$ to -0.4 eV (53). Third, because of their localizations, multiplet effects, including exchange splitting, are large (~ -1 eV). This suggests that even if lattice rearrangements are insufficient to produce an Anderson negative effective U on

their own (or even if lattice distortions inhibit a negative U), many-electron effects can produce an exchange correlation negative effective U if $U_{ver} < U_R + U_{JT} + \Delta U_{MC}$.

We see from Figure 3*d* that Mn is a likely impurity to show such an effect [interstitial Cr in Si is another example treated recently (50)]. For substitutional Mn in III–V materials we have three charge states: the neutral impurity $[A^0, d^4, e^2t^2, {}^5T_2]$ appearing as the 3+ oxidation state, the single positive impurity $[A^+, d^3, e^2t^1, {}^4T_1]$ appearing as the 4+ oxidation state, and the singly negative impurity $[A^-, d^5, e^2t^3m, {}^6A_1]$ appearing as the 2+ oxidation state. The dominant ground-state orbital configurations e^nt^m and the multiplet assignments were determined from the preceding analysis of the absorption data. We denote A^+, A^0, and A^- as $N-1$, N, and $N+1$, respectively (for Mn $N = 4$). We calculate from the spectra $\Delta E_{MC}^{4,5} = -1.25$ eV and $\Delta E_{MC}^{3,4} = 0.6$ eV, or $\Delta U_{MC} = -0.65$ eV. Using the estimate quoted above, $\Delta E_R^{N,N+1} + \Delta E_{JT}^{N,N+1} \cong -(0.3\text{–}0.4)$ eV from II–VI materials, and neglecting $\Delta E_R^{N-1,N} + \Delta E_{JT}^{N-1,N}$ (although, as shown in Reference 50, this might raise the effective U a little), we find that if $U_{ver} \lesssim 1$ eV the system will have an overall negative effective U, but without MC $U > 0$. MF calculations for GaP (9) predict U_{ver} values in this range. If $U < 0$ then Mn^{3+} $[A^0, {}^5T_2]$ state is never the ground state for any value of the Fermi energy E_F. This awaits experimental testing.

LATTICE RELAXATION EFFECTS

The discussion above raises the question of the extent of lattice distortions around 3d impurities. This was addressed recently by Lindefelt & Zunger (25, 31) who evaluated the breathing-mode relaxation around interstitial 3d impurities in silicon from the host crystal valence force field and the Hellmann-Feynman forces obtained in the QBCF calculations.

Figure 18 depicts the predicted breathing-mode distortions of the 82 Si atoms belonging to the 9 shells nearest the impurity. One finds a strong coupling between the displacements of the first two shells and the rest of the crystal: If the rest of the crystal is frozen, the first two shells relax only about two thirds of their unconstrained relaxation. Integrating the forces over the lattice space, we find that the energy gained by relaxation is approximately 0.4 eV for all impurities. The major relaxation pattern near the impurity consists of an outward movement of the first 1*NN* shell and an inward movement of the second (2*NN*) shell, leading to an approximately tenfold coordinated TA, a molecular species that is rare in covalent coordination chemistry of 3d elements (54). Intriguingly, all three metallic disilicide structure types, $TiSi_2$ (orthorhombic), $CrSi_2$ (hexagonal), and $MoSi_2$ (tetragonal), are known to avoid the conventional close-packed coordi-

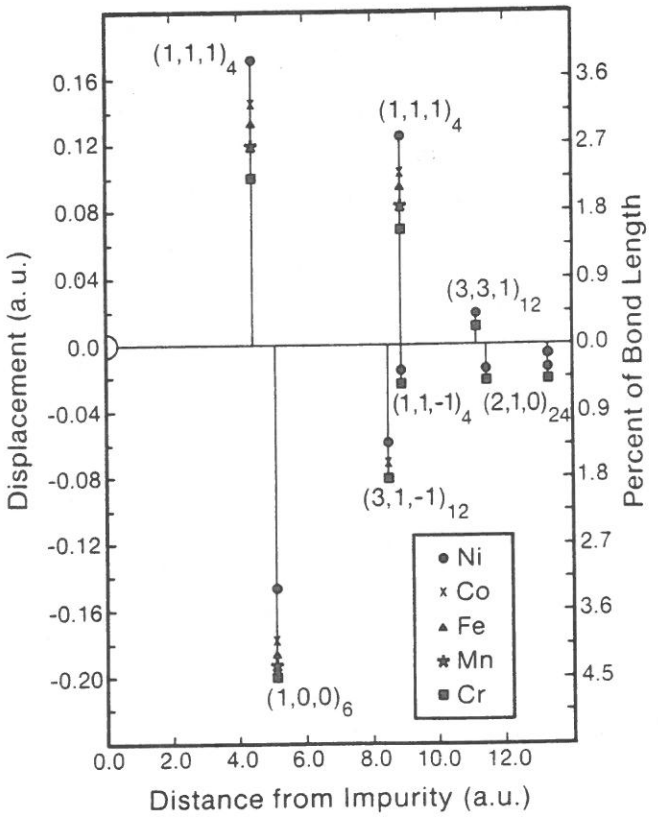

Figure 18 Predicted breathing-mode relaxations of the $(h,k,l)_I$ silicon shells (containing I atoms) surrounding the neutral tetrahedral interstitial TA impurity (at the origin).

nation of 12. Instead, they exhibit a variety of unique stacking sequences, all leading to an approximate coordination number of 10 around the transition atom. Our analysis suggests that the Si : TA system maintains a universal relaxation pattern despite the variations in the band filling, electronegativities, and ground-state configurations due to $s \rightarrow d$ population inversion. This process places all of the bonding electrons in the d shell, producing nearly identically weak bonds for all TA.

UNIVERSAL FEATURES OF BINDING ENERGIES OF 3d IMPURITIES IN SEMICONDUCTORS

The understanding gained so far on the one-electron energy level structure of 3d impurities in semiconductors and the significance of intra-atomic

Coulomb repulsions with their host crystal–dependence can now be used to unravel some universal trends (55). Constructing an effective two-level model (55), we find that the position of the impurity level should be approximately at a fixed distance from the vacuum level (vacuum-related binding energy or VRBE) in different materials, *if* the level is antibonding, as is the case for TA.

To check the idea of material invariance of the VRBE we have used our calculated levels of GaP : Fe and InP : Fe, referring all one-electron energy levels to the electrostatic potential of each host crystal at its interstitial site. The potential at the empty interstitial site was shown to be a reasonable approximation of an internal (surface-independent) vacuum level, and it produced reasonable predictions for band alignments at interfaces (56). We found that their host-related binding energies (HRBE) differ substantially ($E_c = -1.22$ and -0.28 eV for GaP : Fe and InP : Fe, respectively), but their VRBE are much closer (-2.98 and -2.86 eV, respectively). A similar insight was derived by Jaffe & Zunger (57) by analyzing the band gap anomaly in ternary chalcopyrites.

Motivated by the above considerations, we followed recent suggestions (58, 59) and referred the experimentally determined HRBE of transition atom acceptor [first $(0/-)$ and second $(-/=)$] and donor $[(0/+)]$ levels in III–V and II–VI semiconductors to an approximate vacuum level, taken as the experimentally determined (60) photothreshold (Φ) for the (110) surface (55). We neglected material variations in surface corrections because (*a*) the experimental precision for surface corrections is poorer (60) than for Φ; (*b*) Φ has proven to correlate well (61) with the valence band maximum (E_{VBM}) for a given sequence of common-cation compounds; (*c*) only relative shifts in the vacuum level from one material to another are needed. We show in Figure 19 the results for eight semiconductors for which reliable data exist (55), indicating the oxidation states that exist at each region of the gap. Note that the $1+$, $2+$, and $3+$ oxidation states shown in Figure 19 correspond to the $A^{=}$, A^{-}, and A^{0} charge states in III–V semiconductors, and to A^{-}, A^{0}, and A^{+} charge states in II–VI's. The remarkable result is that within a class of compounds the VRBE of each impurity are nearly constant, despite significant variations in HRBE.

A few chemical trends are apparent: (*a*) Shallow acceptors in CdTe and ZnTe (e.g. Cu with $E_A = E_v + 0.15$ eV) become deep acceptors in ZnS and ZnSe (around $E_v + 1.3$ eV and $E_v + 0.7$ eV for Cu in ZnS and ZnSe, respectively) merely because the VBM in the latter systems recedes. This is why CdTe can be made low resistivity p-type by cation substitution, whereas sulphides cannot (57). However, isovalent substitutional elements that lack a deep level (e.g. Li, Na) can form shallow acceptors in II–VI's, yielding similar HRBE and different VRBE. The same is true of Mn

acceptors in III–V's: they are deep in GaP but shallow in GaAs because $\Phi_{GaP} > \Phi_{GaAs}$.

(*b*) Cr, Co, and Ni impurities that exist as deep donors in ZnS, ZnSe, and CdSe, but were not observed in CdTe are actually predicted here to be inside the CdTe valence band.

(*c*) Iron impurity forms a midgap (semi-insulating) level in InP, but Cr is needed to form a midgap level in GaAs (despite the similarities in band gaps) because the VBM of InP is lower than that of GaAs.

(*d*) Impurities in CdTe and ZnTe have similar HRBE, hence only ZnTe is shown since their Φ's are nearly identical. Other common-anion semiconductors (e.g. InP and GaP) show variations in deep level HRBE because their Φ's are different.

(*e*) The failure to detect a V acceptor in InP, despite its existence in GaAs, is consistent with the prediction that it lies just above the conduction band minimum (CMB). However, an experimental search for the V acceptor level in GaP would be important to shed light on its position in GaAs.

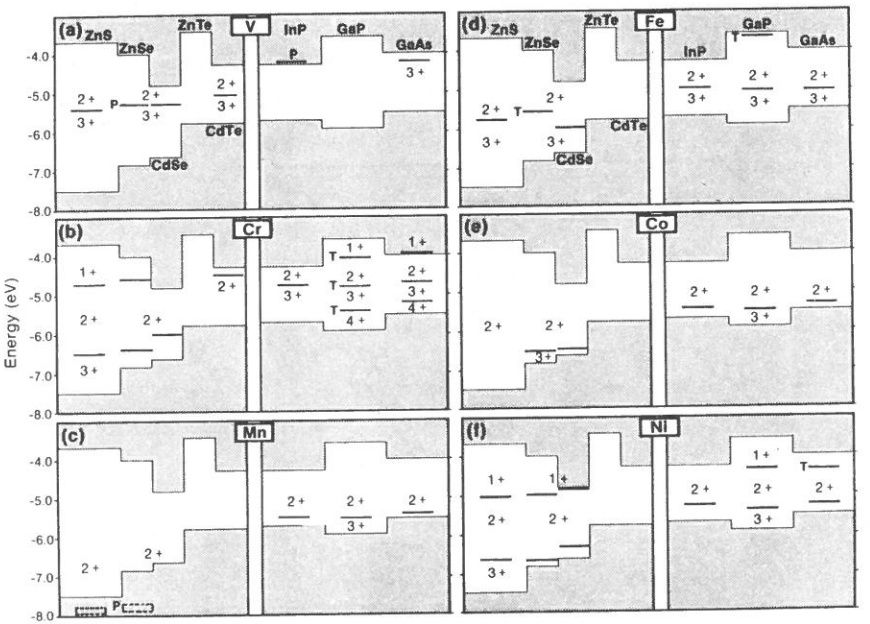

Figure 19 Vacuum-related binding energies of six 3d impurities in eight host semiconductors, showing that the regions of stability of the 1+, 2+. and 3+ oxidation states are very similar in materials of the same class. *T* equals tentative experimental value; *P* is predicted.

(*f*) Cr in GaP can appear in the 1+ oxidation state, but it does not exist in GaAs and InP (although it could be forced into the gap by applying pressure) since the conduction band minima of the latter materials are lower than in GaP.

(*g*) We predict that the VRBE of transition atom impurities in mixed alloys (e.g. ZnS_xSe_{1-x} or $GaAs_xP_{1-x}$) will follow the variations with x in Φ and not the HRBE (e.g. the CBM or any conduction band in particular). This was recently confirmed experimentally (62) for GaAsP : Cu.

Figure 20 shows the universal trends in the VRBE of donors (M^{2+}/M^{3+}) in II–VI's, and of acceptors (M^{3+}/M^{2+}) in III–V semiconductors. (Similar trends are obtained for acceptors in II–VI's except that the jump is between Cr and Mn). The overall trends, including the local minima in Mn, parallel those in free-ion ionization energies: The jump is larger in the more ionic II–VI systems because the impurity Mott-Hubbard coulomb repulsion energies (U) are larger ($U \sim$ 10–20 eV in free ions, 2–3 eV in II–VI's, and 1–2 eV in III–V's). The overall width of the distribution of VRBE is dictated by the host covalency. We hope that the universality of VRBE can be used to predict the approximate location of unknown deep centers in crystals and alloys from the knowledge of Φ and the level position in related semiconductors and their alloys.

The vacuum pinning rule (55) has a significant and unexpected implication on a completely different area of semiconductor research, namely the

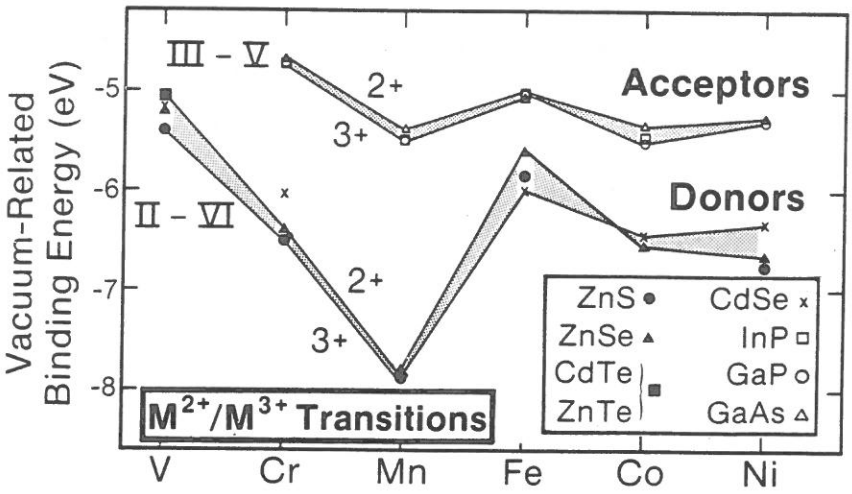

Figure 20 Universal binding energy curve for deep acceptors in III–VI's and deep donors in II–VI's.

determination of valence band offsets in semiconductor interfaces (56, 61). The vacuum pinning rule suggests that if one experimentally determines the concentration dependence $E_{imp}(x)$ of a deep cation impurity level E_{imp} in an alloy $A_xB_{1-x}C$ (where x is the concentration), then the slope α of this dependence $E_{imp}(x) = E_{imp}(0) + \alpha x$ equals the difference $\Phi(1) - \Phi(0)$ in the intrinsic positions of the valence band maxima in the two binary materials AC and BC. This difference (*intrinsic* valence band offset) is an important quantity in numerous heterojunction problems (56) and can be determined independently by interface photoemission studies (however, in this case including interface-specific effects). Experimental determination of α from photoluminescence studies of cation impurities in bulk alloys can then provide an independent measure of the intrinsic band offset $\Phi(1) - \Phi(0)$. Let us indicate here a few examples: (*a*) Samuelson et al (62) determined for $GaAs_{1-x}P_x$: Cu that $\alpha = 0.43 \pm 0.03$ eV. The vacuum rule predicted (55) that $\Phi_{GaP} - \Phi_{GaAs} = 0.43 \pm 0.03$ eV, in excellent agreement with the measured difference in the photothreshold values 0.45 ± 0.03 eV of these materials (63). A recent measurement (64) for $GaAs_{1-x}P_x$: Fe has similarly revealed the same α value as for Cu, substantiating the prediction (55) that for deep cation impurities α measures an intrinsic property of the host alloys, not that of the impurity. (*b*) A recent experiment (65) for self-activated defects in ZnS_xSe_{1-x} alloys showed four defect states with an average α of 0.61 ± 0.04 eV. This is close to our previous (55) prediction of $\Phi_{ZnS} - \Phi_{ZnSe} = 0.65 \pm 0.03$ eV. (*c*) Recent experiments (66) on $A\ell_xGa_{1-x}As$: Fe and $A\ell_xGa_{1-x}As$: Cu (67) revealed $\alpha = 0.45$ eV in both cases. According to our theory this means that the valence band offsets are $\Phi_{A\ell As} - \Phi_{GaAs} = 0.45$ eV. Until recently, this result conflicted with the accepted value of the valence band offsets of AℓAs-GaAs: 0.19 eV measured by photoemission (Reference 63 and references therein) or ~ 0 eV, predicted theoretically (61). However, very careful recent experiments on the AℓAs-GaAs interfaces (68) now reveal a valence band offset of 0.45 ± 0.05 eV, in excellent agreement with the vacuum pinning rule.

The success of this rule in predicting intrinsic valence band offsets offers some new insights into a previous rule, suggested by tight-binding impurity calculations (69). This model of Hjalmarson et al predicted that the slope α of the variation of deep cation impurity levels with alloy composition in $A_xB_{1-x}C$ equals the difference in the cation vacancy levels in the end-point materials AC and BC. Using, for example, the calculated energies of the t_2 gap levels of the Ga vacancy in GaAs ($E_v + 0.06$ eV, Reference 70) and GaP ($E_v + 0.15$ eV, Reference 71), this rule predicts $\alpha = 0.15 - 0.06 = 0.09$ eV, almost five times smaller than the observed value (62) of 0.43 ± 0.03 eV. For $A\ell_xGa_{1-x}As$, the vacancy pinning rule predicts $\alpha = 0$, in sharp disagree-

ment with experiment (66, 67) and with our vacuum pinning rule (55). This substantiates previous criticism of the vacancy pinning rule (72).

SUMMARY

Cation-substitutional transition atom impurities in semiconductors have a universal energy level scheme; that is a bonding, d-like, t^{CFR} level, an antibonding t^{DBH} dangling-bond hybrid, and a nonbonding, d-like, e^{CFR} level between them. The impurity states that evolve from the atomic 3d orbitals are e^{CFR} and t^{CFR}, not e^{CFR} and t^{DBH}, as hypothesized by Ludwig & Woodbury. This pattern of energy levels is consistent with a simple three-level scheme, showing that the physically relevant (level-confining) gap is the tt* gap ($\sim$3–4 eV), not the optical gap. Some acceptor transitions (e.g. Zn, Cu, Ni, Mn, and Cr in III–V materials) evolve from the ionization of the t^{DBH} orbital, while others (Co and Fe) evolve from the ionization of the e^{CFR} orbital.

Despite substantial covalency, the impurity-induced states are rather localized inside the central cell, leading to significant proportions of electron–electron multiplet corrections to the acceptor and donor ionization energies and to the Mott-Hubbard Coulomb repulsion energies. One-electron theory can, at best, reproduce the first component. Many-electron effects decrease in their relative importance for heavier 3d elements (e.g. from Ni to Cu, reaching zero in Zn).

Mott-Hubbard Coulomb repulsion energies (U) are of the order of 0.2–0.3 eV in Si:TA, 1–1.6 eV in III–V's, and 2–3 eV in II–VI's. They are significantly larger for e^{CFR} electrons than for t^{DBH} electrons. They are predicted by theory to within an accuracy of 0.2 eV. The large values of U invalidate the use of Koopman's theorem in these systems. Excitation energies between different $d \rightarrow d^*$ configurations are predicted to within $\sim$0.3 eV, where theory systematically overestimates the interconfigurational crystal-field splittings for such transitions.

The significant reduction in the Coulomb energies relative to their values in free ions ($\sim$2 eV) is a consequence of a charge self-regulating response whereby valence band resonances are reorganized spatially to accommodate population changes in the gap levels. This self-regulating response explains the seemingly paradoxical dual nature of deep TA impurities in semiconductors: They are sufficiently localized to retain large enough multiplet corrections to produce high-spin ground states, yet they are sufficiently delocalized to accommodate numerous charge states within a narrow energy range. This mechanism is speculated to be the dominant factor leading to small relaxation potentials in biological (Fe-containing) electron transporting systems.

Consistent with the self-regulating response and the smallness of U, we find very short (2–4 au) screening lengths in the system, a fast recovery of the impurity-induced charge fluctuation, a partial $s \rightarrow d$ configuration crossover, and a screening overshoot. The net charge on most TA impurities is small (0.1–0.3 eV), indicating the rather minor role of ionicity. The self-regulating response further suggests that although the A^0 centers have significant charge on the ligands, upon ionization (forming A^+) charge flows from the ligands to the impurity. Hence, both A^0 and A^+ have comparable effective charges at the impurity core.

The reduction of U by screening effects (a mean-field phenomena) and the further decrease in U by many-electron effects lead to the possibility of formation of an exchange correlation (XC) negative effective U. This can occur when relaxation effects alone are insufficient to make U negative, but when combined with XC effects U becomes negative. We propose that this effect exists in GaAs : Mn and Si : Cr_i.

The impurity-induced levels in GaP : TA are generically related to the (concentrated-limit) band structure of TA phosphides, much as the levels of Si : TA are related to the band structure of TA silicides. The binding energies of the impurity-induced levels decrease with the impurity's atomic number at a rate that is 5–10 times slower than for free ions. This results from the formation of new transition atom phosphide bonds.

The one-electron energy level scheme that evolves from our calculation suggests a special universality in the binding energies of the antibonding gap levels. When referred to the vacuum level (through the intrinsic work function), all binding energies of the same 3d impurity in various host crystals belonging to the same class of semiconductors are approximately the same. Thus, knowing the binding energy in one host crystal is sufficient to approximate its value in another one.

The self-interaction correction is necessary to obtain the correct high-spin ground state of a number of interstitial 3d impurities in Si. Spin-polarized calculations for Si : 3d reproduce the correct trends in the g-shifts, the hyperfine coupling constants, and the acceptor and donor energies.

Calculation of breathing-mode relaxation in Si : 3d shows a universal pattern : The first shell of neighbors moves out and the second shell moves in, resulting in an effectively tenfold coordinated 3d atoms much as in a number of bulk silicides.

Acknowledgments

It is a great pleasure to thank my collaborators, Adalberto Fazzio, Marilia Caldas, Hiroshi Katayama-Yoshida, Ulf Lindefelt, and Vijay Singh, for fruitful and enjoyable collaborations.

Literature Cited

1. Liechti, C. A. 1977. *Conf. Ser. Inst. Phys.* 33:227
2. Zucca, R., Welch, B. M., Asbeck, P. M., Eden, R. C., Long, S. I. 1980. *Semi-Insulating III–V Materials*, ed. E. J. Rees, p. 33. Kent: Shiva. 361 pp.
3. Bergh, F. A., Dean, P. J. 1976. *Light Emitting Diodes*. London: Oxford Univ. Press
4. Casey, H. C., Panish, M. B. 1978. *Heterostructure Lasers*. New York: Academic
5. Liechti, C. A. 1976. *IEEE Trans. Microwave Theory Tech.* 24:279
6. Hovel, H. J. 1975. *Semiconductors and Semimetals*, ed. R. K. Williamson, A. C. Beer, Vol. 11. New York: Academic
7. Kaufmann, U., Schneider, J. 1983. *Adv. Electron. Electron Phys.* 58:81; Schneider, J. 1983. *Defects in Semiconductors II*, ed. S. Mahajan, J. W. Corbett. New York: North-Holland. 225 pp.
8. Madelung, O., Schulz, M., Weiss, H., eds. 1982. *Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology*, Vol. 17. Berlin: Springer-Verlag. 200 pp.
9. Singh, V. A., Zunger, A. 1985. *Phys. Rev. B*. In press
10. Fazzio, A., Caldas, M., Zunger, A. 1984. *Phys. Rev. B* 30:3430; 1984. 29:5999
11. We use the local spin density formalism of Kohn, W., Sham, L. J. 1965. *Phys. Rev. A* 140:1133; and the Perdew-Zunger form (Perdew, J. P., Zunger, A. 1981. *Phys. Rev. B* 23:5048) of the Ceperly-Alder correlation (Ceperley, D. M., Alder, B. J. 1980. *Phys. Rev. Lett.* 45:566)
12. Brandow, B. H. 1977. *Adv. Phys.* 26:651
13. Wuthruich, K., Schulman, R. G. 1970. *Phys. Today* p. 43; Schejter, A., Aviriam, I., Goldkorn, T. 1982. *Electron Transport and Oxygen Utilization*, ed. Chien Ho. Amsterdam: North-Holland. 95 pp.
14. Griffith, J. S. 1970. *The Theory of Transition Metal Ions in Crystals*. New York: Academic; Ballhausen, C. J. 1962. *Introduction to Ligand Field Theory*. New York: McGraw-Hill. 295 pp.
15. Sugano, S., Tanabe, Y., Kamimura, H. 1970. *Multiplets of Transition-Metal Ions in Crystals*. New York: Academic. 331 pp.
16. Mott, J. F. 1974. *Metal Insulator Transitions*. London: Taylor & Francis. 278 pp.
17. Allen, J. W. 1980. *Semi-Insulating III–V Materials*, ed. G. J. Rees. Kent: Shiva. 261 pp.
18. Kohn, W., Onffroy, J. R. 1973. *Phys. Rev. B* 8:2485
19. Baranowski, J. M., Allen, J. W., Pearson, G. L. 1967. *Phys. Rev.* 160:627
20. Hemstreet, L. A., Dimmock, J. O. 1979. *Phys. Rev. B* 20:1527; Hemstreet, L. A. 1977. *Phys. Rev. B* 15:834
21. Fazzio, A., Zunger, A. 1984. *Solid State Commun.* 52:265
22. Slater, J. C. 1974. *The Self-Consistent Field for Molecules and Solids*, pp. 51. New York: McGraw-Hill. 583 pp.
23. Lindefelt, U., Zunger, A. 1982. *Phys. Rev. B* 26:846; 1981. *Phys. Rev. B* 24:5913
24. Lindefelt, U. 1983. *Phys. Rev. B* 28:4510
25. Lindefelt, U., Zunger, A. 1984. *Phys. Rev. B* 30:1102

26a. Bernholc, J., Lipari, N. O., Pantelides, S. T. 1980. *Phys. Rev. B* 21:3545

26b. Baraff, G. A., Schluter, M. 1980. *Phys. Rev. B* 19:4965

27. Singh, V. A., Zunger, A., Lindefelt, U. 1983. *Phys. Rev. B* 27:4909; 1983. 27:1420
28. Zunger, A., Lindefelt, U. 1983. *Phys. Rev. B* 27:1191
29. Zunger, A., Lindefelt, U. 1983. *Physica* 117B:185
30. Zunger, A., Lindefelt, U. 1983. *Solid State Commun.* 45:343
31. Lindefelt, U., Zunger, A. 1984. *J. Phys. C* 17:6047
32. Zunger, A., Lindefelt, U. 1982. *Phys. Rev. B* 26:5989
33. Fazzio, A., Caldas, M., Zunger, A. *Phys. Rev. B*. In press
34. Zunger, A. 1983. *Phys. Rev. Lett.* 50:1215
35. Gemma, N. 1984. *J. Phys. C* 17:2333
36. Fazzio, A., Leite, J. R. 1980. *Phys. Rev. B* 21:4710; Fazzio, A., Dal Pino, A. Jr. 1983. *Rev. Bras. Fis. Semiconductors* 1:436
37. Cartling, B. G. 1975. *J. Phys. C* 8:3171
38. De Leo, G. G., Watkins, G. D., Fowler, W. B. 1982. *Phys. Rev. B* 25:4962
39. Watkins, G. D. 1983. *Physica* 177B:9
40. Woodbury, H. H., Ludwig, G. W. 1960. *Phys. Rev.* 117:102; *Phys. Rev. Lett.* 5:96; Ludwig, G. W., Woodbury, H. H. 1962. *Solid State Physics*, ed. F. Seitz, H. Ehrenreich, D. Turnbull. New York: Academic 13:263
41. Haldane, F. D. M., Anderson, P. W. 1976. *Phys. Rev. B* 13:2553
42. Katayama-Yoshida, H., Zunger, A. 1984. *Phys. Rev. Lett.* 53:1256
43. Perdew, J. P., Zunger, A. 1981. *Phys. Rev. B* 23:5048
44. Williams, A. R., Feibelman, P. J., Lang, N. D. 1982. *Phys. Rev. B* 26:5433

45. Chase, D. B., McClure, D. S. 1976. *J. Chem. Phys.* 64:74; Harrison, J. G., Lin, C. C., Ching, W. Y. 1981. *Phys. Rev. B* 24:6060
46. Greulich-Weber, S., Niklas, J. R., Weber, E. R., Spaeth, J. M. *Phys. Rev. B.* In press
47. Zunger, A. 1983. *Phys. Rev. B* 28:3628
48. Feichtinger, H., Czaputz, R. 1983. *Phys. Status Solidi* 79:K143
49. Wunstel, K., Wagner, P. 1982. *Appl. Phys.* 27:207
50. Katayama-Yoshida, H., Zunger, A. Unpublished
51. Caldas, M., Fazzio, A., Zunger, A. 1985. *Proc. 13th Int. Conf. Defects in Semiconductors*, p. 1035
52. Anderson, P. W. 1975. *Phys. Rev. Lett.* 34:953
53a. Kaminska, M., Baranowski, J. M., Uba, S. M., Vallin, J. T. 1979. *J. Phys. C* 12:2197
53b. Godlewski, M., Kaminska, M. 1980. *J. Phys. C* 13:6537
54. Wells, A. F. 1975. *Structural Inorganic Chemistry*. Oxford: Clarendon. p. 789. 4th ed.
55. Caldas, M., Fazzio, A., Zunger, A. 1984. *Appl. Phys. Lett.* 45:671
56. Frensley, W. R., Kroemer, H. 1977. *Phys. Rev. B* 16:2642
57. Jaffe, J. E., Zunger, A. 1984. *Phys. Rev. B* 29:1882
58. Ledebo, L. A., Ridley, B. K. 1982. *J. Phys. C* 15:1961
59. Rojo, P., Leyral, P., Nouailhat, A., Guillot, G., Lambert, B., et al. 1984. *J. Appl. Phys.* 55:395
60. Swank, R. K. 1967. *Phys. Rev.* 153:844; Fischer, T. E. 1966. *Phys. Rev.* 142:519; Gobeli, G. W., Allen, F. G. 1965. *Phys. Rev.* 137:245
61. Harrison, W. A. 1980. *Electronic Structure and the Properties of Solids.* San Francisco: Freeman. 254 pp.
62. Samuelson, L., Nilsson, S., Wang, Z. G., Grimmeiss, H. G. 1984. *Phys. Rev. Lett.* 45:671
63. Zunger, A. 1985. *Phys. Rev. Lett.* 54:848
64. Samuelson, L. 1985. Proc. Int. Conf. Defects in Semiconductors, *J. Elect. Materials*, p. 101
65. Grimmeiss, H. G., Meijer, E., Mack, R., Müller, G. O. 1984. *J. Appl. Phys.* 56:2762
66. Wang, Z. G., Ledebo, L. A., Grimmeiss, H. G. 1984. *J. Appl. Phys.* 56:2762
67. Jansson, L., Wang, Z. G., Ledebo, L. A., Grimmeiss, H. G. 1983. Il. Nuovo Cimento 2D:1718
68. Wang, W. I., Stern, F. Unpublished results
69. Hjalmarson, H. P., Vogl, P., Wolford, D. J., Dow, J. D. 1980. *Phys. Rev. Lett.* 44:810
70. Bachelet, G. B., Baraff, G. A., Schluter, M. 1981. *Phys. Rev. B* 24:915
71. Scheffler, M., Bernholc, J., Lipari, N. O., Pantelides, S. T. 1984. *Phys. Rev. B* 29:3269
72. Singh, V. A., Lindefelt, U., Zunger, A. 1982. *Phys. Rev. B* 25:2781