

verted to the latter by annealing.

The authors would like to thank Dr. Peter Mark

of Princeton University for his help and comments in preparation of this paper.

\* Work sponsored in part by NASA Grant No. NGR 19-002-027 and in part by Air Force Contract No. F 33601-69-C-0550.

† Present address: Stanford Research Institute, Menlo Park, Calif. 94025.

<sup>1</sup>M. A. Lampert, Phys. Rev. **103**, 1648 (1956).

<sup>2</sup>M. A. Lampert and P. Mark, *Current Injection in Solids* (Academic, New York, 1970).

<sup>3</sup>H. T. Henderson and K. L. Ashley, Phys. Rev. **186**, 811 (1969).

PHYSICAL REVIEW B

VOLUME 5, NUMBER 8

15 APRIL 1972

## Band Gaps of Semiconductor Alloys\*

D. Stroud†

*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14850*

(Received 6 August 1971)

It is shown, on the basis of perturbation theory and a simple assumption about the matrix elements of the random scattering potential, that the effect of alloy disorder is inevitably to reduce the band gap of a semiconductor alloy below its virtual-crystal value. The conclusion, which was previously suggested by Van Vechten and Bergstresser (VB) on the basis of a more intuitive argument, is in agreement with existing theory and calculation. In contrast to VB, however, the present treatment predicts that the magnitude of the disorder-induced shift will vary from one direct gap to another in the same alloy, and that the uppermost gaps may be larger than a straight-line interpolation of the gaps of the constituents. This prediction appears to be corroborated by the electroreflectance studies of Thompson *et al.* on Ga(As<sub>x</sub>P<sub>1-x</sub>).

The band gaps of a number of semiconductor alloys are smaller than those predicted by a virtual-crystal calculation, in which the randomly varying alloy potential is approximated by a periodic lattice of average atomic potentials. Indeed, when the band gap is a direct gap which is of the same symmetry in both limiting pure semiconductors, it seems invariably to lie below a straight-line interpolation of the band gaps of the constituents.<sup>1</sup> This is strikingly evident in Table I, which is derived from a recent paper by Van Vechten and Bergstresser (VB).<sup>2</sup> VB suggested that the increased scattering due to the disorder in the actual crystal potential brought about a decrease in the magnitude of the various direct band gaps, and proposed a simple formula, based on Fermi's "golden rule," to estimate the magnitude of this effect.

It is the purpose of the present Comment to improve upon VB's intuitive argument, and simultaneously to draw a connection between their work and the coherent-potential treatment<sup>3</sup> of disordered alloys. Using a simple general argument based on perturbation theory, we show that the effect of alloy disorder is indeed to reduce the band gap *below* its virtual-crystal value. We also derive an approximate expression for the magnitude of the shift. The resulting formula has a structure somewhat similar to that of VB, except that the relevant parameters now have an unambiguous microscopic interpretation. Furthermore, the present treatment

suggests, in contrast to VB's simple estimate, that the shift should vary in magnitude from one direct gap to another in the same material, and that the highest gaps might lie *above* a straight-line interpolation. At the end of this Comment, we shall cite experimental evidence which seems to favor the present point of view.

The essence of the argument is as follows: We wish to calculate the disorder-induced shift in energy of  $E_v$  and  $E_c$ , the uppermost state of the virtual-crystal valence band and the lowest of the conduction band of a semiconductor alloy. According to second-order perturbation theory, both of these states will be pushed down in energy by the states in the conduction band and pushed up by states in the valence band. But by definition  $E_c$  is closer to the conduction band than  $E_v$ . It will therefore be pushed down more than  $E_v$ , and pushed up less, and the gap will be reduced below its virtual-crystal value.

To make the argument precise, we consider a substitutional semiconductor alloy  $A_xB_{1-x}$  whose atoms  $A$  and  $B$  are randomly distributed on the  $N$  sites of a periodic lattice,<sup>4</sup> and whose one-electron Hamiltonian is

$$H = K + \sum_l v_l . \quad (1)$$

Here  $v_l = v_l^A$  or  $v_l^B$  is the atomic potential associated with the  $l$ th atomic site, and  $K$  is the kinetic energy. The shift  $\Delta E_{n\vec{k}}$  of virtual-crystal eigenstate  $|m\vec{k}\rangle$

TABLE I. Deviations of band gaps from virtual-crystal and straight-line behavior.  $\Delta E_g^{\max}$  is the maximum deviation of the smallest direct band gap from its virtual-crystal value in a number of semiconductor alloy systems.  $(E_g - E_{11n})^{\max}$  is the maximum deviation of the gap from a straight line. All data are based on a similar table from Ref. 2, and none given in that reference is omitted here. Where the reference gives two values an average has been taken. The accuracy of the experimental measurements is given as  $\pm 0.02$  eV or worse. Accuracy of the virtual-crystal calculation is hard to assess, but probably is no better than  $\pm 0.02$  eV, also. The value of  $\Delta E_g^{\max}$  for  $\text{Zn}(\text{S}_x\text{Se}_{1-x})$  is thus zero to within experimental and calculational accuracy. For several of the Ga-based alloys listed here, the minimum direct gap is larger in energy than the indirect band gap over part of the concentration range. The theorem proven in text is therefore not rigorously valid over that part of the range. Experimentally, however,  $\Delta E_g$  is negative for all concentrations, so that the theorem is more than satisfied.

Alloy system	$\Delta E_g^{\max}$ (eV)	$(E_g - E_{11n})^{\max}$ (eV)
$\text{Ga}(\text{As}_x\text{P}_{1-x})$	-0.00	-0.05
$\text{In}(\text{As}_x\text{P}_{1-x})$	-0.02	-0.06
$(\text{Ga}_x\text{In}_{1-x})\text{P}$	-0.08	-0.11
$(\text{Ga}_x\text{In}_{1-x})\text{As}$	-0.04	-0.11
$\text{In}(\text{As}_x\text{Sb}_{1-x})$	$\leq -0.14$	$\leq -0.14$
$(\text{Ga}_x\text{Al}_{1-x})\text{As}$	$\sim -0.05$	$\sim -0.05$
$(\text{Ga}_x\text{In}_{1-x})\text{P}$	-0.12	-0.22
$\text{Zn}(\text{S}_x\text{Se}_{1-x})$	$\sim +0.03$	$\sim 0.00$
$\text{Zn}(\text{Se}_x\text{Te}_{1-x})$	-0.33	-0.32
$\text{Zn}(\text{S}_x\text{Te}_{1-x})$	-0.53	-0.60
$(\text{Ag}_x\text{Cu}_{1-x})\text{I}$	-0.11	-0.15
$\text{Cu}(\text{I}_x\text{Br}_{1-x})$	-0.27	-0.11

from its virtual-crystal eigenvalue  $E_{n\vec{k}}$  can be calculated from knowledge of the electron self-energy operator  $\Sigma(z)$ , which is defined in terms of energy  $z$  by

$$\langle (z - H)^{-1} \rangle = [z - H_{\text{vc}} - \Sigma(z)]^{-1}. \quad (2)$$

Here  $\langle \rangle$  denotes an average over all configurations of  $A$  and  $B$  atoms, and the virtual-crystal Hamiltonian  $H_{\text{vc}} = \langle H \rangle$ .

If the form factors of the  $A$  and  $B$  atoms differ only slightly, it may be sufficient to calculate  $\Sigma(z)$  only to lowest order in  $\Delta_i = v_i^B - v_i^A$ . Such an expression may be obtained, for example, by expanding the approximate coherent-potential<sup>3</sup> formula for  $\Sigma(z)$  in powers of  $\Delta_i$ . To the same order in the  $\Delta_i$ 's,  $\Delta E_{n\vec{k}}$  is given by

$$\begin{aligned} \Delta E_{n\vec{k}} &= \text{Re} \langle n\vec{k} | \Sigma(E_{n\vec{k}}) | n\vec{k} \rangle \\ &= x(1-x) \text{Re} \sum_i \langle n\vec{k} | \Delta_i G_{\text{vc}} \Delta_i | n\vec{k} \rangle, \end{aligned} \quad (3)$$

where  $G_{\text{vc}} = (z - H_{\text{vc}})^{-1}$ .<sup>5</sup> In all cases of interest, the sum is trivial, resulting in a factor of  $N$ . Hence (3) becomes

$$\Delta E_{n\vec{k}} = x(1-x) \mathcal{O} \left( \frac{1}{N} \sum_{n'\vec{k}'} \frac{|M_{nn'}; \vec{k}\vec{k}'|^2}{E_{n\vec{k}} - E_{n'\vec{k}'}} \right), \quad (4)$$

where  $|M_{nn'}; \vec{k}\vec{k}'|^2 = N^2 |\langle n\vec{k} | \Delta_i | n'\vec{k}' \rangle|^2$ . Formula (4) is obtained by inserting the complete set of states  $\{|n'\vec{k}'\rangle\}$  into (3) and then multiplying and dividing by  $N$  in order to express  $\Delta E_{n\vec{k}}$  in terms of quantities of order unity. Introducing  $\bar{M}_{n\vec{k}}$  and  $F(E)$  by

$$\begin{aligned} \mathcal{O} \left( N^{-1} \sum_{n'\vec{k}'} \frac{|M_{nn'}; \vec{k}\vec{k}'|^2}{E_{n\vec{k}} - E_{n'\vec{k}'}} \right) \\ \equiv |\bar{M}_{n\vec{k}}|^2 \mathcal{O} \left( N^{-1} \sum_{n'\vec{k}'} \frac{1}{E_{n\vec{k}} - E_{n'\vec{k}'}} \right) \\ \equiv |\bar{M}_{n\vec{k}}|^2 F(E_{n\vec{k}}), \end{aligned} \quad (5)$$

where  $\mathcal{O}$  means "principal value of," we find that  $\Delta E_{n\vec{k}}$  is determined to lowest order in  $\Delta_i$  by an effective matrix element  $\bar{M}_{n\vec{k}}$  and a function  $F(E)$  which depends only on the energy  $E_{n\vec{k}}$ .  $F(E)$ , moreover, is just the Hilbert transform of the virtual-crystal density of states  $\rho_{\text{vc}}(E)$ ,<sup>6</sup>

$$F(E) = \mathcal{O} \int \frac{\rho_{\text{vc}}(E') dE'}{E - E'}. \quad (6)$$

We now assume, by analogy with the approach successfully used to calculate the interband absorption of semiconductors, that  $\bar{M}_{n\vec{k}}$  is a slowly varying function of  $n$  and  $\vec{k}$ , which may be taken as a constant for states near the band gap. It then follows from (4) and (5) that the shift in the band gap due to disorder scattering is just

$$\Delta E_g = \Delta E_c - \Delta E_v = x(1-x) |\bar{M}|^2 [F(E_c) - F(E_v)]. \quad (7)$$

$|\bar{M}|$  corresponds to VB's electronegativity difference  $C_{AB}$ , while  $F(E_v) - F(E_c)$  is the analog of their inverse bandwidth parameter  $1/A$ .

Figure 1(b) shows a schematic of  $F(E)$  for a typical semiconductor alloy with density of states as shown in Fig. 1(a). The main point of the figure is that

$$F(E_c) < F(E_v). \quad (8)$$

From the definitions of  $E_c$  and  $E_v$  and Eq. (6), it is readily shown that Eq. (8) always holds for any semiconductor alloy. (This is the mathematical statement of the remarks in the third paragraph of this Comment.) The other factors in (7) are all positive. Therefore, if perturbation theory is sufficient and  $\bar{M}_{n\vec{k}}$  is constant, the band gap of any semiconductor alloy must invariably be smaller than its virtual-crystal value, as was to be shown.

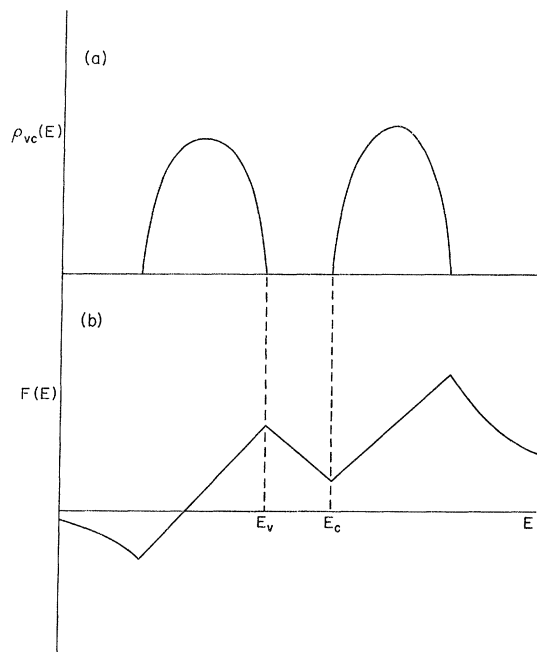


FIG. 1. (a) Schematic of virtual-crystal density of states  $\rho_{vc}(E)$  of a typical semiconductor alloy. (b) Function  $F(E)$ , the Hilbert transform of  $\rho_{vc}(E)$ .  $F(E_c) < F(E_v)$  always.

It remains to consider the accuracy of second-order perturbation theory. Equation (7) will be adequate if  $\overline{MF}(E) \ll 1$  for all  $E$  of interest. In a semiconductor alloy,  $F(E)$  is of the order of the

inverse valence bandwidth, namely,  $1 \text{ Ry}^{-1}$ . For alloys of Table I,  $\overline{M}$  is of the order of the difference between the form factors of the constituents, namely,  $0.01$ – $0.05 \text{ Ry}$ . Thus  $\overline{MF}(E) \ll 1$ , as required.

The applications of Eq. (8) in semiconductor alloys are not restricted to the smallest band gaps. To calculate  $\Delta E_g$  for higher band gaps (e. g., the gap between virtual-crystal states  $E'_v$  in the valence band and  $E'_c$  in the conduction band), we must take into account the variation of  $F(E'_c) - F(E'_v)$  and, to a lesser extent, of  $|\overline{M}|^2$ , with  $E'_v$  and  $E'_c$ . To judge from Fig. 1,  $F(E'_c) - F(E'_v)$  may be expected to change from negative to positive for sufficiently large  $E'_c - E'_v$ . It may therefore be anticipated that  $\Delta E_g$  will be *positive* for some of the highest band gaps.

Evidence that precisely this situation prevails in  $\text{Ga}(\text{As}_x\text{P}_{1-x})$  may be found in the extensive electroreflectance studies of Thompson *et al.*<sup>7</sup> The lowest gaps ( $E_0, E_0 + \Delta$ ) exhibit distinct upward bowing (i. e., lie below a straight-line interpolation of the band gaps of the constituents). The highest gaps ( $E_2, E_2 + \delta$ ), on the other hand, lie almost on a straight line, or even slightly *above* a straight line. When the bowing that would be present even in the virtual-crystal approximation is subtracted out,<sup>2</sup> it appears that  $\Delta E_g$  for the uppermost gap is indeed *positive*. Thus the experimental trend is evidently in accord with expectations based on Eq. (8) and Fig. 1.

The author is grateful to Dr. Robert Besserman for rekindling his interest in this problem. He also wishes to acknowledge useful discussions with Professor J. W. Wilkins and Dr. J. A. Van Vechten.

\*Supported by the Atomic Energy Commission under Contract No. AT(30-1)-3699, Technical Report No. NYO-3699-61.

†Present address: Department of Physics, Ohio State University, Columbus, Ohio 43210.

<sup>1</sup>See, for example, A. G. Thompson and J. C. Woolley, *Can. J. Phys.* **45**, 255 (1967). It was previously suggested by M. Cardona [*Phys. Rev.* **129**, 69 (1963)] that this reduction was due to alloy disorder. Cardona, however, mistakenly assumed that the virtual-crystal approximation would lead to a linear variation of the band gap.

<sup>2</sup>J. A. Van Vechten and T. K. Bergstresser, *Phys. Rev. B* **1**, 3351 (1970).

<sup>3</sup>See, for example, D. Stroud and H. Ehrenreich, *Phys. Rev. B* **2**, 3197 (1970); B. Velický, E. S. Kirkpatrick,

and H. Ehrenreich, *Phys. Rev.* **175**, 747 (1968).

<sup>4</sup>The following arguments are immediately applicable to pseudobinary alloys of the type  $(A_xB_{1-x})C$ , in which case  $N$  is simply the number of  $A$  and  $B$  sites.

<sup>5</sup>We neglect the lifetime broadening of  $E_c$  and  $E_v$ , which also arises from disorder scattering but which is very small near the band edges in crystalline semiconductor alloys.

<sup>6</sup>Strictly speaking,  $F(E)$  is infinite for an unbounded density of states  $\rho_{vc}(E)$ . Since, however, states  $|n'\vec{k}'\rangle$  corresponding to nonzero matrix elements  $M_{m'}; \vec{k}'$  presumably fall into a bounded region of energy, this nominal difficulty is of no physical significance.

<sup>7</sup>A. G. Thompson, M. Cardona, K. L. Shaklee, and J. C. Woolley, *Phys. Rev.* **146**, 601 (1966).