PHYSICAL REVIEW B VOLUME 1, NUMBER 8 15 APRIL 1970

Electronic Structures of Semiconductor Alloys

J. A. VAN VECHTEN*†
Bell Telephone Laboratories, Murry Hill, New Jersey 07974

AND

T. K. BERGSTRESSER‡§
The James Franck Institute, The University of Chicago, Chicago, Illinois 60637

(Received 30 October 1969)

The problem of the band structure of semiconductor alloy systems is treated by both the dielectric two-band method and by the use of an empirical (local) pseudopotential. With both methods, calculations are made in the virtual-crystal approximation assuming linear dependence on alloy concentration of the lattice constant and the parameters of the two methods. Contrary to some previous assertions, both methods predict, in general, a nonlinear dependence of the interband gaps on concentration. An estimate is also made of the effects of second-order perturbations to the virtual-crystal approximation, i.e., the effect of disorder. Of particular interest are the lowest direct and indirect energy gaps and the deviations of these from linearity. The treatment is confined to alloys of compounds having the formula A^NB^{8-N}, but quaternary and more complicated alloys may be treated as easily as the ternary alloys to which most previous experimental work has been confined. Results are compared to experiment and to the empirical formula of Thompson and Woolley. We find that, with one free parameter, the dielectric method gives good agreement with experiment, but that the local-pseudopotential method apparently does not yield satisfactory results for this problem.

I. INTRODUCTION

IT has been shown experimentally that the lowest direct band gap $E_0(\Gamma_{15}{}^v-\Gamma_1{}^c)$ in many semiconducting alloys $(AB)_x(CD)_{1-x}$ of compounds AB and CD has an approximately quadratic dependence on the mole fraction of one compound x,[1-3]

$$E_0(x)=a+bx+cx^2. \tag{1}$$

The nonlinear or bowing parameter $c=c(AB,CD)$ is four times the deviation of E_0 from linearity at the equimolar concentration $x=0.5$. The other two parameters a and b are determined by the values of E_0 observed in the pure semiconducting compounds. Although empirical rules have been proposed[1] to predict the variation of c with AB and CD, these rules have not been found to have general validity. Until now, this problem has not received the theoretical attention that it warrants. For the practical problem of achieving a desired color of luminescence or laser emission or in determining the alloy which will have the largest possible direct gap, the deviation from linearity is not a negligible effect, although it is typically about 5% of the total variation.

Thomposon and Woolley[1] invoked the effect of microscopic inhomogeneity in the crystal potential with substitutional alloying, which goes as $x(1-x)$, to justify the quadratic variation of their formula (1). However, even in the virtual-crystal approximation,[4] in which one neglects this aperiodic part of the crystal potential, one does not expect the variation of the energy gap to be linear in composition because the energy gap is not a linear function of the assumed potential for any of the commonly used methods of band-structure calculation.

II. DIELECTRIC MODEL CALCULATION

A dielectric method of calculating the band structures[5,6] of tetrahedrally coordinated semiconductors based on the Phillips's theory of electronegativity differences[7] has recently been introduced. For 19 tetrahedrally coordinated binary semiconductors, the method requires three times fewer parameters[6] and is as accurate as previous empirical pseudopotential results.[8] A summary of the dielectric method of calculating band structures presented in Ref. 6 is given in the Appendix, together with a table of the parameters used for the pure compounds.

We have applied this dielectric method to the alloy problem by adjusting the parameters of the theory (the average homopolar and ionic energy gaps[7,9] E_h and C, the average valence band to d-band f sum D_{av},[6,9] which is a parameter characterizing nonlocal effects, and the lattice constant a) so as to agree with experiment in the pure compounds from which the alloy is formed. We then assume, in the spirit of the virtual-crystal approximation, that these parameters vary

---

[*] Portions of this work completed while supported as a Fannie and John Hertz Foundation Fellow.

[†] Present address: Naval Research Laboratory, Washington, D. C., 20390.

[‡] Supported in part by the U. S. Army Research Center (Durham).

[§] Present address: Department of Physics, University of Wisconsin, Madison, Wisc. 53706.

[1] A. G. Thompson and J. C. Woolley, Can. J. Phys. **45**, 255 (1967).

[2] M. Cardona, Phys. Rev. **129**, 69 (1963).

[3] S. Larach, R. E. Schrader, and C. F. Stocker, Phys. Rev. **108**, 587 (1957).

[4] L. Nordheim, Ann. Physik **9**, 607 and 641 (1931).

[5] J. C. Phillips and J. A. Van Vechten, Phys. Rev. Letters **22**, 705 (1969).

[6] J. A. Van Vechten, Phys. Rev. **187**, 1007 (1969).

[7] J. C. Phillips, Phys. Rev. Letters **20**, 550 (1968).

[8] M. L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966).

[9] J. A. Van Vechten, Phys. Rev. **182**, 891 (1969).

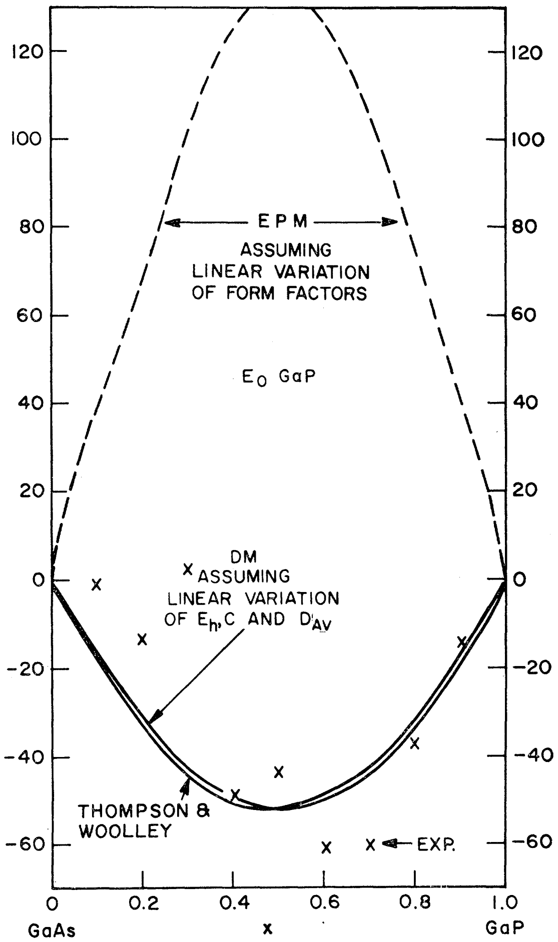

FIG. 1. Deviation from linearity of E_0 gap in GaAs-P as observed in electroreflectance (Ref. 11) and as calculated by the dielectric method (DM), the local empirical pseudopotential method (EMP) and the empirical formula (1) of Thompson and Woolley.

linearly with composition.[10] The resulting deviation from linearity of the calculated E_0 gap for GaAs-P is shown in Fig. 1. This alloy system has been the subject of an extensive study by the electroreflectance method.[11] The experimental points and the empirical formula of Thompson and Woolley adjusted to fit the data are also shown. Note that our method has used no free parameters to fit the bowing effect and that our curve nowhere deviates from the empirical curve by more than 2 meV. This is especially surprising because, in the calculation shown in Fig. 1, we completely neglected any effect of disorder in the crystal potential. Indeed, the experimental uncertainty is greater than 2 meV. We must conclude that the agreement as to the magnitude of the bowing is somewhat fortuitous but that the calculated curve has essentially the same shape as the empirical parabola.

There exists no rigorous quantitative theory of the effect of disorder in semiconductor alloys. It is widely believed that any aperiodic potential will cause a tailing of the density-of-states spectrum from the square-root Van Hove singularities into the forbidden region at the band edges. This effect has been found in second- and higher-order perturbation calculations.[12,13] Such an effect would be observable and would allow one to distinguish between bowing due to the variation of the average-crystal potential, the virtual-crystal effect, and the bowing due to the aperiodicity. The tailing effect would also explain some of the discrepancies between various experimental determinations of the bowing parameter c.

However, for the substitutional alloys considered here, which do not involve donor or acceptor impurities or an amorphous structure, examination of the optical-absorption spectrum[3,14] does not reveal evidence for this inhomogeneity broadening. It appears that the effect of disorder is to narrow the band gap without altering the shape of the singularity in the spectrum.[15]

We will here propose a simple method to estimate the size of the bowing effect due to disorder. We propose that the total bowing parameter c should be the sum of the intrinsic bowing c_i found in the virtual-crystal approximation and the extrinsic bowing c_e due to the effects of aperiodicity:

$$c=c_i+c_e. \tag{2}$$

The aperiodic potential is a short-range effect. Under the assumption that the distribution of substituted atoms is perfectly random on the one sublattice, the range of the effect is one unit cell. Thus, the aperiodic potential cannot distinguish the various points in one

---

[10] As discussed in Refs. 6 and 9, we assume that both the average homopolar gap E_h and the homopolar value of each individual interband gap at a particular point in k space are functions of nearest-neighbor distance only and, thus, lattice constant only. In this calculation, we retain these same functions and vary the E_h values between the compounds by assuming that the lattice constant varies linearly with composition. Thus, the variation of E_h is not strictly linear, but, because the difference of the lattice constants of most compounds which form alloys is quite small, the difference from linearity is not significant.

[11] A. G. Thompson, M. Cardona, K. L. Shaklee, and J. C. Woolley, Phys. Rev. **146**, 601 (1966).

[12] N. F. Mott, Advan. Phys. **16**, 49 (1967); M. Cutler and N. F. Mott, Phys. Rev. **181**, 1336 (1969).

[13] R. H. Paramenter, Phys. Rev. **97**, 587 (1955).

[14] J. C. Woolley, M. B. Thomas, and A. G. Thompson, Can. J. Phys. **46**, 157 (1968); J. C. Woolley (private communication) has pointed out that the doping level of the samples may have obscured the presence of a tail in conduction-band density of states. However, these results do preclude the presence of a tail at the valence-band maximum. We know of no reason to believe that such perturbations could cause a tail in the conduction band but not the valence band. Of course, the breakdown of the phonon selection rule due to the mass difference between alloyed elements causes a qualitative change in any indirect edge near the threshold. This effect has been observed by P. J. Dean, G. Kaminsky, and R. B. Zetterstrom [Phys. Rev. **181**, 1149 (1969)] in the lowest indirect transition in P-rich Ga(As,P) alloys. Beyond this effect, one might also interpret this data as evidence for tailing of the band edge, but this would be represented by an absorption strength of a few tenths of an inverse centimeter. We do not consider such extremely weak effects.

[15] Evidence that a sharp band edge also obtains in amorphous but unstrained Ge has been presented by T. M. Donovan, W. E. Spicer, and J. M. Bennett, Phys. Rev. Letters **22**, 1058 (1969).

Brillouin zone and the c_e for a given pair of bands is approximately constant, unless the band character changes over the zone. The average- or virtual-crystal potential, being periodic throughout the crystal, does affect the bandgap at various points in the zone differently.

It is also clear from the above considerations that c_e must be describable in terms of some average potential difference between the substituted atoms. The details of the potential cannot be important because they are too short ranged to be distinguishable. The averaged potentials we choose to consider are those of the Phillips's electronegativity theory,[7,9] E_h and C. If we assume that, in the alloy, all bond lengths are equal so that we have a perfect zinc-blende (or wurtzite) lattice, then the homopolar part of the potential E_h, which is a function of nearest-neighbor distance only,[6] is given exactly in the virtual-crystal approximation and the only fluctuation in the potential arises from differences in the antisymmetric part C. If the bond lengths are not all the same, then there are fluctuations in E_h also. At the opposite extreme, we might assume that all bond lengths are the same as in the pure compound, i.e., that in $\mathrm{GaAs}_{1-x}\mathrm{P}_x$ the Ga-As bond lengths are all the same as in pure GaAs. In that case, the total potential would fluctuate by an amount proportional to the difference in the total mean energy gap $E_g=(E_h^2+C^2)^{1/2}$ in the pure compounds. In the "virtual lattice" or equal-bond-length approximation, the magnitude of the fluctuation is proportional to the valence of the substituted element. This is easily seen by considering the definition[7,9] of C as the screened Coulombic potential difference of the two in

$$C(A,B)=be^2\left(\frac{Z_A}{r_A}-\frac{Z_B}{r_B}\right)e^{-k_s(r_A+r_B)/2}. \tag{3}$$

In (3), k_s is the Thomas-Fermi screening wave number, r_A is the covalent radius[9] of atom A, and Z_A is its valence number (1–7). b is a factor inserted to account for the fact that the Thomas-Fermi approximation overestimates dielectric screening at short distances. It has been found[9] that this factor is constant to within ±10% at a value of 1.5 for zinc-blende and wurtzite crystals.

From (3), we see that in an alloy $MF_{1-x}G_x$ the magnitude of the fluctuations of the actual potential in the virtual-lattice approximation is

$$C_{FG}=be^2Z\left|1/r_F-1/r_G\right|e^{-k_sR}, \tag{4}$$

where $Z=Z_F=Z_G$ and $R=\frac{1}{2}[r_M+(1-x)r_F+xr_G]$. We have calculated values for C_{FG} for all alloy systems for which data is available to us assuming $b=1.5$ and $x=0.5$ (see Table I).

Note that, if we make the equal-bond-length assumption, the magnitude of the potential fluctuations are generally larger but not as strongly dependent on the valence Z of the substituted element. To our knowledge, there is no x-ray evidence that could distinguish which of these two assumptions is more nearly correct. However, our results tend to support the assumption that the fluctuations of the potential is proportional to Z and, thus, support the virtual-lattice approximation.

TABLE I. Comparison of theory and experiment for the bowing parameter of the E_0 minimum direct gap. C_{FG} is the electronegativity difference of the alloyed elements, c_i is the parameter for the intrinsic (virtual-crystal) bowing, c_e is for the extrinsic (disorder) bowing, and $c=c_i+c_e$. Only Refs. a and b involve differential-type measurements for which the estimated uncertainty is ±0.08 ev (see text). For other data, the uncertainty is greater.

Alloy MF-G	C_{FG} (eV)	c_i (eV)	c_e (eV)	c_{calc} (eV)	c_{exp} (eV)
GaAs-P[a]	0.31	0.21	0.09	0.30	0.21
InAs-P[b]	0.28	0.15	0.08	0.23	0.20, 0.26
Ga-InSb[c]	0.48	0.12	0.24	0.36	0.43
Ga-InAs[d,e]	0.53	0.28	0.29	0.57	0.33, 0.56
InAs-Sb[f]	0.81	0.03	0.67	0.70	≥0.58
Ga-AlAs[g]	0.18	0	0.03	0.03	~0.2
Ga-InP[h]	0.56	0.39	0.31	0.70	0.88
ZnS-Se[i]	0.37	0.14	0.14	0.28	~0
ZnSe-Te[i]	1.06	−0.04	1.14	1.10	~1.28
ZnS-Te[i]	1.44	0.28	2.12	2.40	~2.40
Ag-CuI[j]	0.33	0.14	0.11	0.25	0.58
CuI-Br[j]	1.23	−0.65	1.54	0.89	0.44

[a] Reference 11.

[b] A. G. Thompson, J. E. Rowe, and M. Rubenstein, J. Appl. Phys. **40**, 3280 (1969).

[c] J. C. Woolley and J. A. Evans, Proc. Phys. Phys. Soc. (London) **78**, 354 (1961).

[d] Reference 14.

[e] E. F. Hockings, I. Kudman, T. Seidel, C. M. Schmelz, and E. F. Steigmeier, J. Appl. Phys. **37**, 2879 (1966).

[f] J. C. Woolley and J. Warner, Can. J. Phys. **42**, 1879 (1964).

[g] H. C. Casey, Jr., and M. B. Panish, J. Appl. Phys. **40**, 4910 (1969).

[h] M. R. Lorentz, W. Reuter, W. P. Dumke, R. J. Chicotka, G. D. Pettit, and J. M. Woodall, Appl. Phys. Letters **13**, 421 (1968).

[i] Reference 3.

[j] Reference 2.

Having determined the magnitude of the potential fluctuations C_{FG}, we make a simple estimate of their effect on the bandgap by assuming that c_e is proportional to the probability that the fluctuations scatters an electron, which, according to the Golden Rule No. 2, is proportional to the square of the fluctuation divided by some typical bandwidth. Thus,

$$c_e=C_{FG}^2/A\,, \tag{5}$$

where we shall take the bandwidth parameter A to be a constant for all compounds considered and shall fit it to one datum. We do this because it is not clear exactly what the correct bandwidth ought to be and because the accuracy of the experimental data available at present does not seem to warrant more elaborate fitting.

The one datum that we use to determine A is that for which the largest c_e is expected (and observed), i.e., the alloy system for which C_{FG} is largest. This is the ZnS-Te system. For this system, the total observed bowing parameter[3] c is approximately 2.40 eV, while our calculated $c_i=0.28$ eV, so that $c_e\cong 2.12$ eV and $A\cong 0.98$ eV.

TABLE II. Comparison of theory and experiment for E_1 and E_2 direct gaps. These transitions connect the same two bands so the effect of disorder c_e is the same as in Table I. All data is from the electroreflectance measurements of Refs. 11 and 16 so that the estimated uncertainty is ±0.08 eV.

Alloy	$c_{i,1}$ (eV)	c_e (eV)	$c_{1\text{calc}}$ (eV)	$c_{1\text{exp}}$ (eV)	$c_{i,2}$ (eV)	$c_{2\text{calc}}$ (eV)	$c_{2\text{exp}}$ (eV)
GaAs-P	0.10	0.09	0.19	0.25	0.02	0.11	0.03
InAs-P	0.07	0.08	0.15	0.26	0.02	0.10	0.27
Ga-InSb	0.07	0.24	0.31	0.33	0.05	0.29	0.24
Ga-InAs	0.13	0.29	0.42	0.50	0.06	0.35	0.27
InAs-Sb	0.02	0.67	0.69	0.55	0.05	0.74	~0.6

III. COMPARISON WITH EXPERIMENT

The result of our calculation using these parameters is compared with all the experimental data available to us for the E_0 lowest direct gap in Table I and for the higher critical-point direct gaps,[16] E_1 and E_2, connecting the same two bands in Table II. Table I contains data from various types of experiment, while Table II contains only electroreflectance data. It is fairly generally conceded that differential techniques such as electroreflectance[17,18] are the most accurate for determining critical-point bandgaps. For the electroreflectance method, the accuracy of the determination of the critical-point bandgap is limited by the uncertainty in the line shape. Until a theory of this line shape is available, the uncertainty in the determination of the bowing parameter c by careful electroreflectance measurements[19] will remain about ±0.08 eV, i.e., ±20 meV uncertainty in the edge. The uncertainty of the determinations by other methods is substantially greater than this and is difficult to estimate.

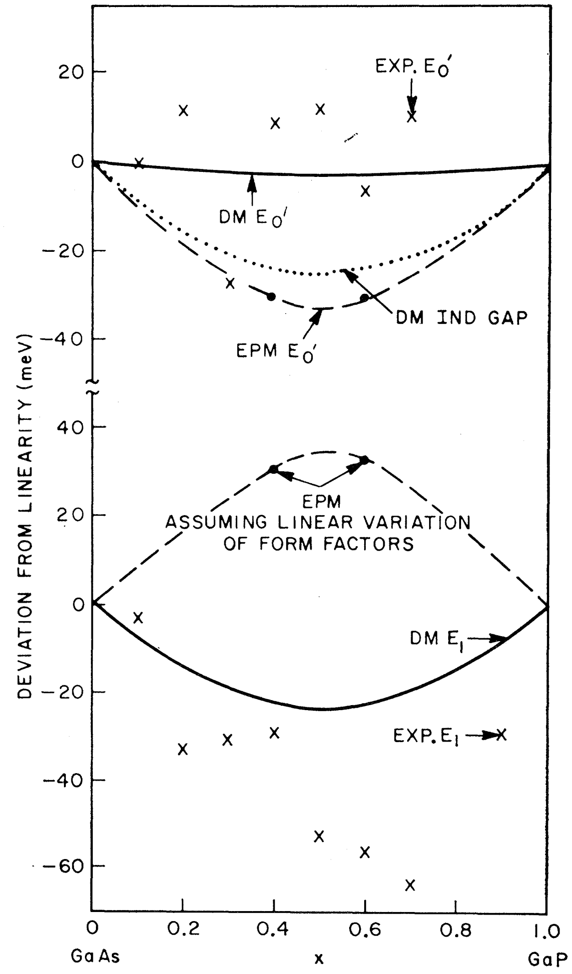

FIG. 2. Deviation from linearity of lowest indirect gap E_1 and E_0' direct gaps in GaAs-P as observed by electroreflectance (Ref. 11) and as calculated.

With these limits of experimental errors in mind, we conclude that the agreement between our theory and experiment in Table I and Table II is quite satisfactory for all alloys investigated with the possible exception of the noble-metal halides.

IV. LOCAL PSEUDOPOTENTIAL CALCULATION

In order to use the (local) empirical pseudopotential method,[8] the simplest approximation to make, other than to interpolate linearly the energy levels, is to assume the form factors to be linear functions of composition. To be specific, let us take the system Ga(As,P), where the effects of disorder are expected to be small (c_e=0.09 eV from Table I). In addition to the previously discussed results, Fig. 1 shows the result of a calculation of E_0 as a function of composition x when the form factors for GaAs and GaP recently obtained by Walter and Cohen[20] are interpolated linearly. The calculated value of c has the wrong sign and is too large by more than a factor of 2. The calculated results of the dielectric and pseudopotential methods for other gaps in the Ga(As,P) system are shown in Fig. 2. Again, the pseudopotential results deviate greatly from experiment. Furthermore, it is clear that any effect of disorder that might be contemplated to correct the results for E_0 would leave E_1 in disagreement and vice versa, assuming as we have that disorder affects E_0 and E_1 are approximately equally. Similarly, poor results have been found in other III–V alloy systems.

The result that one obtains for c is surprisingly sensitive to the form-factor values one chooses at the endpoints $x=0$ and $x=1$. To illustrate with the Ga(As,P) system; the use of the form-factor values of Cohen and Bergstresser[8] yields once again an upward bowing of E_0, but with one-half of the magnitude found with the Walter-Cohen form factors. This seems to be closer to experiment, but such a consideration is a poor

---

[16] S. S. Vishnubhatia, B. Eyglunent, and J. C. Woolley, Can. J. Phys. **47**, 1661 (1969).

[17] B. O. Seraphin and R. B. Hess, Phys. Rev. Letters **14**, 138 (1965).

[18] M. Cardona, K. L. Shaklee, and F. H. Pollak, Phys. Rev. **154**, 696 (1967).

[19] D. E. Aspnes (private communication).

[20] J. P. Walter and M. L. Cohen, Phys. Rev. **183**, 763 (1969).

way to choose between two sets of form factors. The latter form factors have been evaluated using the latest experimental data and yield band structures and a fit to reflectivity that is certainly better than the older Cohen-Bergstresser determination. A third set of pseudopotential form factors for GaAs and GaP have been used by Stukel, Theiman, and Collins[21] to make a similar alloy calculation. These were chosen to fit the band structure calculated by the OPW method for the pure compounds. These form factors give, for E_0, a curve which does not bow at all. This is the essential point to this discussion: A bias of ~ 0.01 Ry to the form-factor values which is unrelated to the alloy problem, e.g., a bias caused by setting $V(q)=0$ for large q or forcing a set of local form factors to fit an intrinsically nonlocal band structure, will yield an unknown and potentially unacceptably large error in the determination of the bowing parameter c.

In principle, the local pseudopotential form-factor values should not vary strictly linearly with composition. To estimate this effect, one can take a dielectric screening point of view, where the form factor of one atomic species $V(q)$ is given by

$$V(q)=V^b(q)/\epsilon(q). \qquad (6)$$

Here, $\epsilon(q)$ is the dielectric screening function, which depends on the composition of the alloy; and $V^b(q)$ is a bare atomic form factor, which does not depend on composition other than through the normalization factor of $(1/a^3)$, where a is the lattice constant. A nonlinear variation of the form-factor values $V(G)$, where G is a reciprocal-lattice vector, will result from the nonlinear variation of the normalization volume and dielectric screening function, and from the variation in magnitude of the reciprocal-lattice vectors coupled with the wave-vector dependence of the bare form factors and of the dielectric screening function. It may also be observed that the form factors might depend on the crystalline environment through environment-dependent core shifts.[22,23]

The effect of normalization volume is simple to calculate. Using once again the Walter-Cohen form factors, we find a change in the bowing parameter, $\Delta c=+0.07$. In order to bring c into correspondence with experiment, $\Delta c=+0.7$ is needed. The nonlinear variation considering the concentration and wave-vector dependence of the dielectric screening function was found to be small in Ga(As,P) and several other systems when evaluated by any of these three methods: (1) use of the Hartree screening function (more appropriate to metals than to semiconductors), (2) use of Srinivasan's screening function[24] with $\epsilon(0)$ varying linearly, (3) use of Srinivasan's screening function with E_g varying linearly. The last method yielded $\Delta c=0.00$.

The remaining effect, the dependence of the form factor on the wave vector coupled with the change in magnitude of reciprocal-lattice vectors, is potentially large, but only where there is a dependence of the lattice constant on concentration. The previously mentioned form-factor determinations yield only the values at the reciprocal-lattice vectors, so an interpolation is needed. We do not have a satisfactory method for such interpolation; more investigation is needed in this area. However, several different methods were tried, including fits to several different analytic expressions, and including using the slope and curvature of the Animalu-Heine[25] form factors at the appropriate values of the wave vector, taking into account the differences in screening and atomic volume, and with and without an arbitrary removal of the kink at $q=2k_F$. In no case did a decrease of the upward bowing occur. This negative result is specific to the Ga(As,P) system, and, in fact, it is specific to the endpoint form factors assumed. Several other systems were tried with no consistent success. To pursue this further, one would need to determine a nonarbitrary method of interpolation.

V. CONCLUSIONS

Because the bowing in the pseudopotential calculations is so dependent on the choice of form factors for the pure compounds, we conclude that this method is unsuited for calculating the bowing for two reasons. First, the method contains too many empirically determined parameters, six form-factor values for each compound.[26] Some of them are sufficiently close to being linearly dependent on the rest in their effect on the band structure, that different sets of form factors giving quite similar band structures can give a quite different bowing parameter. This problem is not nearly so great with the dielectric method because, when the homopolar values of the energy gap are determined from the nearest-neighbor distance alone, there remain only two parameters, the average antisymmetric potential or electronegativity difference C and the nonlocal parameter D_{av}, to be determined for each compound. This may be done quite unambiguously.

The second reason that the empirical pseudopotential method does not give accurate results for the dependence of bandgap on alloy concentration is that it makes the local approximation, i.e., the form factors are constants independent of wave number. Thus, s- and p-like electron wave functions are treated equivalently. For the purpose of calculating the band structure of a given

---

[21] D. J. Stukel and T. C. Collins (unpublished).

[22] P. J. Lin and J. C. Phillips, Phil. Mag. **14**, 257 (1965).

[23] O. P. Gupta, Phys. Rev. **174**, 668 (1968).

[24] G. Srinivasan, Phys. Rev. **178**, 1244 (1969).

[25] A. O. E. Animalu and V. Heine, Phil. Mag. **12**, 1249 (1965); A. O. E. Animalu, Solid State Theory Group, Cavendish Laboratory, Cambridge, England, Technical Report No. 4 (unpublished).

[26] This number is often reduced to 3 per compound by fixing the three symmetric form factors to be the average of the symmetric form factors of the group-IV element of the row of the Periodic Table to which the constituent elements belong. This corresponds to our assumption that the average symmetric potential E_a is a function of nearest-neighbor distance only.

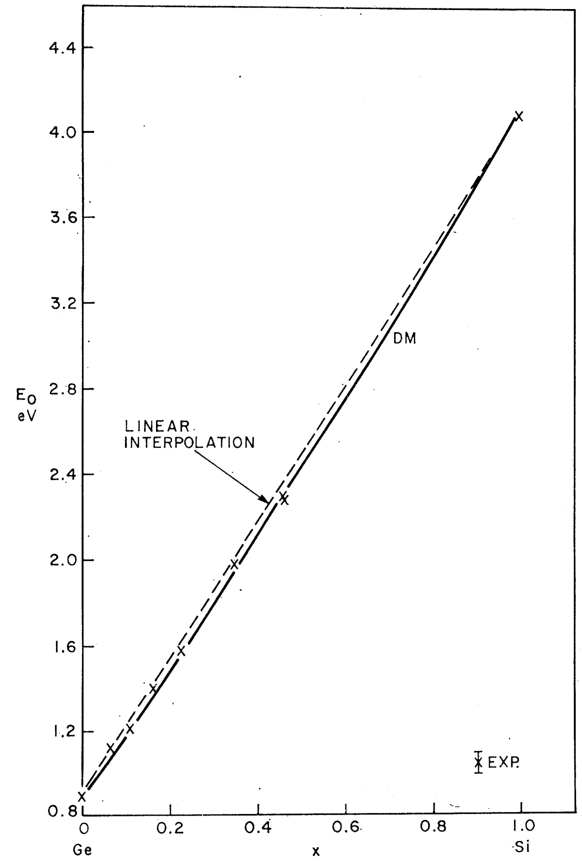

FIG. 3. Variation of E_0 gap in Ge-Si alloy system as observed in electroreflectance and as calculated.

compound, this approximation causes no great difficulty because the s-like states are far more sensitive to small changes in the form factors than are the p-like states.[8] Although in principle the six form factors are overdetermined by the data, in practice one can adjust the values to simulate nonlocal effects on the s-like states without disturbing the p-like states significantly. However, these empirical adjustments are nonphysical and vary from one compound to another, so that one does not obtain the adjustment appropriate for the alloy by averaging the adjusted form factors of the endpoint compounds.

TABLE III. Multiplication table giving calculated c_i values for E_0 gap of all combinations of eight III-V semiconductors. Not all combinations will form alloys. Minus sign indicates upward bowing. In all cases, our predicted c_e would outweigh the negative c_i's.

	AlAs	AlSb	GaP	GaAs	GaSb	InP	InAs	InSb
AlAs	···	188	−104	0	744	180	684	1464
AlSb	188	···	−4	−292	162	−296	−60	572
GaP	−104	−4	···	208	840	392	972	1640
GaAs	0	−292	208	···	156	36	280	636
GaSb	744	162	840	156	···	60	−76	124
InP	180	−296	392	36	60	···	148	424
InAs	684	−60	972	280	−76	148	···	32
InSb	1464	572	1640	636	124	424	32	···

TABLE IV. Dielectric-method predictions for the critical composition x_c for the alloy system $A_{1-x}B_x$ at which the lowest direct and indirect gaps will be equal in energy. Also shown are the values of the gap before spin-orbit splitting and experimental values of x_c. For those compounds marked with an asterisk, very little empirical information concerning the band structure is available. Therefore, we have had to estimate the parameter for these compounds from Ref. 6. As empirical refinements have not been made for these alloys, the values must be considered tentative.

A	B	$x_{c\,\text{calc}}$	$E_{0c\,\text{calc}}$	$x_{c\,\text{expt}}$
GaAs	GaP	0.475	2.08	0.45,[a,b] 0.50[c]
GaAs	AlAs*	0.32	2.00	0.34,[d] 0.37[a]
GaAs	AlSb	0.43	2.10	
GaAs	AlN*	0.59	4.81	
InP	GaP	0.68	2.25	0.8,[e] 0.75[f]
InP	AlP*	0.37	2.55	
InP	AlAs*	0.53	2.18	
InP	AlN*	0.71	5.87	
InAs	AlAs*	0.60	1.83	
InAs	AlSb	0.72	1.95	
InAs	GaP	0.77	2.07	
GaSb	AlSb	0.325	1.44	
GaSb	GaP	0.45	1.57	
GaSb	AlAs*	0.24	1.33	
InSb	AlAs*	0.62	1.64	
InSb	AlSb	0.73	1.82	
InSb	GaP	0.78	1.95	
GaN*	GaP	0.76	3.23	
GaN*	AlN*	0.45	7.23	

[a] H. C. Casey, Jr., and M. B. Panish, J. Appl. Phys. **40**, 4910 (1969).
[b] M. G. Craford, G. E. Stillman, J. A. Rossi, and N. Holonyak, Jr., Phys. Rev. **168**, 867 (1968).
[c] J. W. Allen and J. W. Hodby, Proc. Phys. Soc. (London) **82**, 315 (1963).
[d] H. Nelson and H. Kressel, Appl. Phys. Letters **15**, 7 (1969).
[e] M. R. Lorenz, W. Renter, W. P. Dumke, R. J. Chicotka, G. D. Pettit, and J. M. Woodall, Appl. Phys. Letters **13**, 421 (1968).
[f] H. C. Casey, Jr. (private communication).

The dielectric method does not suffer from the above difficulty because it does not make the local approximation. Nonlocal effects are explicitly introduced through the D_{av} parameter which describes the lowering of s-like conduction-band states from their value in the absence of such effects (see appendix and Ref. 6). We have noted that the value of c_i calculated by the dielectric method is generally several times more sensitive to the choice of the endpoint D_{av} values than to their C values.

It should be noted that the dielectric method was specifically designed to represent the chemical trends in the bandgaps of these semiconductors.[6] Therefore, it is not surprising that this method can "interpolate" these trends between pure compounds. It is ideally suited to the alloy problem and, we believe, accurate enough to be useful in the design of semiconductor alloys. We present in Table III a multiplication table for the calculated c_i values of the E_0 gap in the eight most common III-V semiconductors. These are given irrespective of whether or not an alloy actually may be formed of the two compounds. In Table IV we tabulate our calculated values for the critical composition x_c at which the $A_{1-x}B_x$ has its lowest direct and indirect gaps at the same energy and the calculated value of that energy neglecting disorder and spin-orbit splitting. [These effects lower the energy but do not affect the value of x_c because the valence-band maximum is invariably the same (Γ_{15}) and the conduction-band minimum is always a singlet belonging to the same band.]

In Fig. 3, we compare the electroreflectance data of Kline, Pollak, and Cardona[27] for the E_0 gap of the Si-Ge alloy system with the result of a slightly modified version of our calculation. The modification is necessary because, by symmetry, $C=0$ for both Si and Ge, but this is clearly not so for the alloy. Thus, we assume that the equimolar alloy is an ordered zinc-blende compound SiGe having $C=0.25$ eV. Thus, we divide the alloy range into two parts and calculate for Ge and SiGe and then for SiGe and Si. It is gratifying that experiment and theory agree within experimental error in a system where the bands cross and linear variation exceeds 3 eV.

Note added in proof. The spectrum of GaN has recently been observed from 3 to 12 eV by Kosicki, Powell, and Burgiel.[28] Peaks observed at higher energies are in good agreement with the predictions of Ref. 6, but the E_0 gap was observed at 3.40 eV. Maruska and Tietjen[29] have obtained similar results. These results indicate that the empirical refinement appropriate to GaN is that in Table VI D_{av} should be increased from 1.110 to 1.138. Then in Table IV, the calculated values of x_c and E_{0c} should be 0.80 and 3.0 eV for GaN-P and 0.58 and 7.1 eV for Ga-AlN. Thanks are due to Dr. Kosicki and Dr. Maruska for communicating their results to us prior to publication.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Professor J. C. Woolley for many helpful discussions and suggestions. One of us (J. A. V. V.) has also benefitted from discussions with Professor M. Cardona, Dr. B. I. Halperin, and Dr. J. C. Phillips, and the other (T. K. B.) has benefitted from conversations with Professor M. H. Cohen and Dr. B. Velický.

APPENDIX

The result of the dielectric method analysis of the tetrahedrally coordinated $A^N B^{8-N}$ crystals[6] may be expressed in seven postulates:

(1) Any direct energy gap E_i in the absence of d-state perturbations (see Ref. 7) is given by

$$E_i = E_{i,h}[1+(C/E_{i,h})^2]^{1/2}, \qquad \text{(A1)}$$

where $E_{i,h}$ is the value of the gap for the corresponding homopolar crystal. (All homopolar variables are denoted by a subscript h.)

TABLE V. Structural parameters of dielectric method as derived in Ref. 6.

Parameter	Value for Si	Logarithmic derivative
$(\mathrm{I.P.})_h$	5.17	−1.3077
Vac-X_4	8.63	−1.43
$(E_0)_h$	4.10	−2.75
$(E_1)_h$	3.60	−2.22
$(E_2)_h$	4.50	−2.3821
$(E_0')_h$	3.40	−1.92
E_1'	5.90	−1.67
ΔE_0	12.80	−5.07
ΔE_1	4.976	−4.97
$(X_3-X_1)/C=0.071$		

(2) The $E_{i,h}$'s, and all other homopolar variables, are assumed to be simple power-law functions of the nearest-neighbor distance d, only determined by the value obtained in Si:

$$E_{i,h} = E_{i,h}|_{\mathrm{Si}}(d/d_{\mathrm{Si}})^{s_i}, \qquad \text{(A2)}$$

where the index s_i is a parameter.

(3) The ionization potential, i.e., the energy difference between the top of the valence band (at Γ) and the vacuum level, is given by

$$\mathrm{I.P.} = \mathrm{I.P.}_h[1+(C/\mathrm{I.P.}_h)^2]^{1/2}, \qquad \text{(A3)}$$

where the homopolar value $\mathrm{I.P.}_h$ is scaled with d as per (2).

(4) The energy of the top of the valence band at the symmetry point X, the X_5 state, relative to the vacuum level is independent of C, i.e., ionicity, and is given as a

TABLE VI. Values of the parameters of certain of the compounds discussed in the text. The parameters shown are: the nearest-neighbor distance d, the electronegativity difference C, the refined value of D_{av}, the (negative) absolute energy of the X_4 (X_5) level, and the assumed value of the E_0 gap (without spin-orbit splitting) which determines the refined value of D_{av}. For compounds marked by an asterisk, the values are tentative and were estimated in Ref. 6.

Crystal	d (a_0)	C (eV)	D_{av} refined	E_{X_4} (eV)	E_0 (eV)
Si	4.444	0	1.0	8.540	4.100
Ge	4.630	0	1.267	7.940	0.887
GaAs	4.626	2.90	1.235	8.149	1.551
GaP	4.460	3.30	1.152	9.116	2.770
GaSb	5.006	2.10	1.306	7.435	0.998
InSb	5.302	2.30	1.417	6.687	0.463
InAs	4.940	2.74	1.354	7.375	0.557
InP	4.802	3.339	1.270	7.594	1.391
AlSb	5.020	2.90	1.171	7.874	2.465
ZnS	4.426	6.20	1.155	8.380	3.815
ZnSe	4.638	5.60	1.20	8.14	2.904
ZnTe	4.982	4.475	1.219	7.375	2.561
AgI*	5.296	5.70	1.213	6.687	3.476
CuI*	4.944	5.50	1.148	7.375	4.023
CuBr*	4.704	6.90	1.205	6.428	3.399
GeSi*	4.536	0.25	1.122	8.24	2.450
AlAs*	4.632	2.688	1.121	8.38	2.982
AlP*	4.460	3.135	1.0	8.63	5.129
AlN*	3.576	7.30	1.0	11.66	10.432
GaN*	3.674	7.64	1.110	11.320	4.804

---

[27] J. S. Klein, F. H. Pollak, and M. Cardona, Helv. Phys. Acta **41**, 968 (1968).

[28] B. B. Kosicki, R. J. Powell, and J. C. Burgiel (unpublished).

[29] H. P. Maruska and J. J. Tietjen, Appl. Phys. Letters **15**, 327 (1969); and (private communication).

function of d only as per (2):

$$E_X = E_X|_{\rm Si}(d/d_{\rm Si})^{s_x} = \text{Vac.} - E_{X_5}. \qquad \text{(A4)}$$

(5) The energy of the top of the valence band at the symmetry point L, the L_3 state, is midway between the values for Γ and X:

$$E_{L_3} = \tfrac{1}{2}(\text{I.P.} + E_{X_5}). \qquad \text{(A5)}$$

(6) The splitting of the conduction-band X levels, X_1 and X_3, is proportional to C:

$$E_{X_3} - E_{X_1} = \text{const} \times C. \qquad \text{(A6)}$$

(7) The perturbative effect of the d band on the s-like levels of greatest interest, $\Gamma_{1,c}$ and $L_{1,c}$, is expressed by decreasing the $E_0(\Gamma_{15,v} \to \Gamma_{1,c})$ and $E_1(L_{3,v} \to L_{1,c})$ energy gaps from the values indicated by (1) according to

$$E_i = [E_{i,h} - (D_{\rm av} - 1)\Delta E_i][1 + (C/E_{i,h})^2]^{1/2}, \qquad \text{(A7)}$$

where here $i = 0$, 1 and ΔE_i is a parameter which is a function of d only as per (2). The values of the 19 parameters of this formulation are shown in Table V. In Table VI, we give the empirically determined parameters used in this paper for the pure compounds from which the alloys are formed.

PHYSICAL REVIEW B VOLUME 1, NUMBER 8 15 APRIL 1970

Electroabsorption in Semiconductors: The Excitonic Absorption Edge

John D. Dow*
Joseph Henry Laboratory of Physics, Princeton University, Princeton, New Jersey 08540
and
RCA Laboratories, Princeton, New Jersey 08540

AND

David Redfield
RCA Laboratories, Princeton, New Jersey 08540

(Received 26 September 1969)

Numerical calculations of the optical-absorption coefficient for direct, excitonic transitions in a uniform applied electric field are presented. The electron-hole scattering is treated within the effective-mass approximation and leads to an absorption coefficient which differs markedly in size and shape from the Franz-Keldysh absorption spectrum. A detailed numerical study of the shape of the absorption-edge spectrum at photon energies somewhat below the zero-field absorption threshold suggests that for small field strengths the dominant asymptotic form of the absorption coefficient is $\exp(-C_0|E-E_0'|/f)$, where $f = |e|Fa/R$ is the electric field strength in units of exciton Rydbergs per electron-exciton Bohr radius. This result contradicts the existing belief that the electron-hole interaction does not alter the asymptotic form of the Franz-Keldysh shape: $\exp(-C_0'|E-E_0'|^{3/2}/f)$. Physical arguments are presented to show why the exciton effects should be important. A discussion is presented of the interrelationships among the present treatment of electroabsorption and various one-electron, exciton, and many-body formalisms.

I. INTRODUCTION

In 1958 Franz[1] and Keldysh[2] independently developed the theory of direct band-to-band optical transitions in semiconductors in a uniform applied electric field. The primary prediction of their theory, that the optical-absorption edge would broaden and shift toward lower energy in an electric field, was verified experimentally a year later by Böer, Hänsche, and Kummel[3] and others.[4] The introduction of modulation techniques by Seraphin[5] in 1964 greatly increased the interest in electroreflection and electroabsorption as diagnostic tools in analysis of the energy-band structure. This field has naturally expanded rapidly and comprehensive reviews of the experimental[6] and theoretical[7] developments are now becoming available.

It has long been recognized that the Franz-Keldysh theories and their extensions included no correlation

---

* Research sponsored in part by the U. S. Air Force Office of Scientific Research under Contract No. AF49(638)1545.

[1] W. Franz, Z. Naturforsch. **13a**, 484 (1958).

[2] L. V. Keldysh, Zh. Eksperim. i Teor. Fiz. **34**, 1138 (1958) [English transl.: Soviet Phys.—JETP **7**, 788 (1958)].

[3] K. W. Boer, H. J. Hansche, and V. Kummel, Z. Physik **155**, 170 (1959).

[4] R. Williams, Phys. Rev. **117**, 1487 (1960); **126**, 442 (1962); V. S. Vavilov and K. I. Britsyn, Fiz. Tverd. Tela **2**, 1936 (1960) [English transl.: Soviet Phys.—Solid State **2**, 1746 (1969)]; L. V. Keldysh, V. S. Vavilov, and K. I. Britsyn, in *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961), p. 824.

[5] B. O. Seraphin, in *Proceedings of the Seventh International Conference on the Physics of Semiconductors*, edited by M. Hulin (Dunod Cie., Paris, 1964), p. 165.

[6] B. O. Seraphin, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. Beer (Academic Press Inc., New York, to be published), Vol. VI.

[7] D. E. Aspnes and N. Bottka, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. Beer (Academic Press Inc., New York, to be published), Vol. VI.