Lehoczky, and Dr. S. Zwerdling for useful discussions of this problem. I would also like to thank Dr. R. J. Wagner for very helpful discussions of his and Professor Ewald's experiment, and Dr. J. C. Phillips for a most informative discussion of the ionicity theory.

*Research conducted under the McDonnell Douglas Independent Research and Development Program.

[1]S. Groves and W. Paul, Phys. Rev. Letters 11, 194 (1963).
[2]L. Liu and D. Brust, Phys. Rev. Letters 20, 651 (1968).
[3]L. Liu and D. Brust, Phys. Rev. 173, 777 (1968).
[4]D. Sherrington and W. Kohn, Rev. Mod. Phys. 40, 767 (1968).
[5]B. I. Halperin and T. M. Rice, Rev. Mod. Phys. 40, 755 (1968).
[6]J. G. Broerman, Phys. Rev. Letters 25, 1658 (1970).
[7]L. Liu and E. Tosatti, Phys. Rev. Letters 23, 772 (1969).
[8]J. G. Broerman, Phys. Rev. 183, 754 (1969).
[9]J. G. Broerman, Phys. Rev. Letters 24, 450 (1970).
[10]L. Liu and E. Tosatti, Phys. Rev. B 2, 1926 (1970).
[11]J. G. Broerman, Phys. Rev. B 1, 4568 (1970).
[12]J. G. Broerman, Phys. Rev. B 2, 1818 (1970).
[13]J. G. Broerman, L. Liu, and K. N. Pathak, Phys. Rev. B 4, 664 (1971).
[14]J. G. Broerman, J. Phys. Chem. Solids 32, 1263 (1971).
[15]D. Sherrington and W. Kohn, Phys. Rev. Letters 21, 153 (1968).
[16]R. E. Lindquist and A. W. Ewald, Phys. Rev. 135, A191 (1964).
[17]E. D. Hinkley and A. W. Ewald, Phys. Rev. 134, A1260 (1964).
[18]C. F. Lavine and A. W. Ewald, J. Phys. Chem. Solids 32, 1121 (1971).
[19]C. R. Whitsett, Phys. Rev. 138, A829 (1965).
[20]R. R. Galazka, D. G. Seiler, and W. M. Becker, in Proceedings of the Conference on Semimetals and Narrow Gap Semiconductors, Dallas, 1970 (unpublished).
[21]V. I. Ivanov-Omskii, B. T. Kolomietz, V. K. Ogorodnikov, and K. P. Smekalova, Fiz. Tekh. Poluprov. 4, 264 (1970) [Sov. Phys. Semicond. 4, 214 (1970)].
[22]R. J. Wagner and A. W. Ewald, J. Phys. Chem. Solids 32, 697 (1971).
[23]H. Ehrenreich and M. H. Cohen, Phys. Rev. 115, 786 (1959).
[24]P. Nozières and D. Pines, Nuovo Cimento 9, 470 (1958).
[25]P. Nozières and D. Pines, Phys. Rev. 109, 762 (1958).
[26]J. McDougall and E. C. Stoner, Phil. Trans. Roy. Soc. London A237, 350 (1938).
[27]B. L. Booth and A. W. Ewald, Phys. Rev. 168, 805 (1968).
[28]J. C. Phillips, Phys. Rev. Letters 20, 550 (1968).
[29]J. A. Van Vechten, Phys. Rev. 182, 891 (1969).
[30]E. Cohen and A. K. W. A. van Lieschert, Z. Physik. Chem. Leipzig 173, 1 (1935); 173, 32 (1935); 173, 67 (1935).

PHYSICAL REVIEW B VOLUME 5, NUMBER 2 15 JANUARY 1972

Concentration and Temperature Dependence of Impurity-to-Band Activation Energies

G. F. Neumark
Philips Laboratories, Briarcliff Manor, New York 10510
(Received 2 March 1971)

The present paper investigates the cause of the experimentally well-known reduction of impurity activation (ionization) energies in the intermediate doping range. It is shown that this many-body problem can be reasonably approximated by a one-electron screened-impurity approach. Application of this result to literature data on GaP(Zn) gives agreement with the experimental results provided that screening by ionized impurities is included; screening by free carriers alone is insufficient. The inclusion of the ionized impurity screening thus for the first time provides a *quantitative* explanation for the observed reductions in activation energies. A corollary of the screening effect is that the activation energy decreases with increasing temperature; use of this temperature dependence clarifies previous discrepancies between Hall and neutron activation values for the Zn concentration in GaP.

I. INTRODUCTION

It has been known since the work of Pearson and Bardeen[1] in 1949 that impurity activation energies decrease at high impurity concentrations, i.e., the average separation between the impurity levels and the band decreases with increasing impurity concentration. In the present paper, we shall consider only the case of "intermediate" doping, where this separation has not decreased too far, i.e., the impurity band has not yet appreciably merged with the conduction band.[2] (The high doping range is already reasonably well understood.[2,3]) In the intermediate range, it has been shown by Fritzsche[4] that, in

fact, the conductivity (σ) cannot be described by a single-impurity activation energy (ϵ), but has at least three separate and distinct contributions:

$$\sigma = \sum_{i=1}^{3} C_i \, e^{-\epsilon_i / kT} \quad . \tag{1}$$

The energy which dominates at the higher temperatures is referred to[4,5] as ϵ_1 and is attributed[5] to activation from the impurity level to the band. The other two energies have been attributed,[5] respectively, to activation to a band with two carriers per impurity and to hopping between ionized impurities in compensated material.

Here we consider only the problem of calculating ϵ_1, i.e., the activation between the impurity level and the conduction (or valence) band. A first-principles calculation of ϵ_1 involves complicated many-body effects. Even for a hypothetical array of perfectly ordered impurities, one has to consider interaction among positive, negative, and neutral impurities and free carriers. Moreover, in the frequently encountered case of "substantial" compensation ($\geq 10\%$), all these species are generally present in comparable numbers. In addition, for any real system the impurities will not form a regular array, but a random one. It is thus not surprising that extensive early work in the 1950's (see, for example, Debye and Conwell[6]) never fully explained the observed reductions in activation energy. However, the situation has since been improved by the availability of more advanced many-body techniques.[2] Use of these techniques has defined conditions for appropriate simplification of the complex many-body situation,[2] but the results have not yet been applied to the problem of determining ϵ_1.

Essentially, Bonch-Bruevich[2] has derived conditions under which the many-body problem can be reduced, approximately, to a far simpler one-electron screened-impurity approach. Moreover, these conditions apply in the range of present interest. In fact, *free-carrier* screening has already been suggested as the cause of the reduction in activation energy.[7] However, it was then realized[6] that (for Ge) this type of screening alone does not account for the magnitude of the observed reductions in ϵ_1. This latter conclusion has been confirmed in the present work also for GaP (Sec. III C). We show here that this discrepancy can be resolved, and that screening *does* give a dominant contribution to the reduction of ϵ_1. This follows *provided* one includes screening, not only by free carriers, but also that caused by *ionized impurities*. Such ionic screening has previously been included by Harvey[8] in a thermodynamic analysis of activation energies in Ge and by Morgan[9] in an analysis of line shapes in GaAs, and has given good results in both cases. In addition, we reexamine the conditions for applicability of ionic screening, and consequently propose a somewhat different formulation from that given by either Harvey[8] or Morgan.[9] Of further interest, and importance, is the result that screening causes the activation energy to decrease with increasing temperature.

The results of our analysis are applied to p-type GaP (Zn) doped in the 10^{17}–10^{18}-cm^{-3} range. The electronic properties of this material are currently of prime interest for injection-luminescence applications,[10] and recent detailed Hall measurements[11,12] have shown that this doping range exhibits the aforementioned concentration dependence of activation energies. Moreover, the work of Casey *et al.*[11] gives both the usual Hall determination of impurity concentration and neutron-activation data for the Zn concentration. It thus provides a check on the reliability of the Hall method in determining impurity concentrations. Such a check was not available from prior work on Ge and Si. In fact, the neutron data is used to confirm the prediction that the activation energy decreases with temperature; satisfactory agreement can then be obtained with the Hall results provided this temperature dependence is included.

The reduction of the many-electron approach to the one-electron screened-impurity case is given in Sec. II A, and the screening by ionized impurities is discussed in Sec. II B. The comparison to experimental data[11] on GaP (Zn) is given in Sec. III. A discussion of the results, and of their significance as regards Hall analyses, is given in Sec. IV.

II. DERIVATION OF EQUATIONS

A. Simplification of Many-Body Problem

The first assumption in most[2,3] treatments of the semiconductor-impurity problem is the validity of the effective-mass approximation. It has in the past been suggested[13] that the concentration dependence of activation energy may be caused by a lowering of the band edge due to a higher dielectric constant in impure material. This change in dielectric constant results from the large orbits of carriers bound to impurities. Since this effect is not included in the effective-mass formalism, we carried out a separate test of its importance. We found that for the GaP case ($N_{\text{neutral}} \lesssim 10^{18}$ cm^{-3}) it is negligible: Carrying out the calculation analogously to Castellan and Seitz,[13] one obtains a reduction in activation energy of $\lesssim 1$ meV; moreover, an extension using an expression for the change in dielectric constant derived by Hanamura[14] by the random-phase-approximation (RPA) method gave the same result. Since these conclusions correlate with those of Debye and Conwell[6] on Ge, we assume that this effect is generally negligible, i.e., that the

effective-mass approach is applicable.

The many-body impurity treatment, within the effective-mass approximation, is defined by the Hamiltonian

$$H=\sum_i\left(\frac{p_i^2}{2m^*}\right)-\sum_{i,I}\frac{e^2}{\kappa\,|\,r_i-R_I\,|}+\frac{1}{2}\sum_{i\neq j}\frac{e^2}{\kappa\,|\,r_i-r_j\,|}\;, \qquad (2)$$

where the r's are the electron coordinates, the R_I's are the impurity positions, κ is the dielectric constant, p is the momentum, and m^* is the effective mass. Randomness effects are implicitly included, provided the R_I's are distributed at random.

A simplification of Eq. (2) has been achieved by Bonch-Bruevich,[2] who uses Green's-function techniques to show that for nondegenerate statistics and low electron concentration (n) such that

$$(e^2/\kappa kT)n^{1/3}\ll 1\;, \qquad (3)$$

one obtains a one-electron many-impurity equation

$$H=(p^2/2m^*)+\sum_I V(r-R_I)\;, \qquad (4)$$

where $V(r)$ is a *screened* potential,

$$V(r)=-(e^2/\kappa r)e^{-qr}\;, \qquad (5)$$

with q the inverse screening length. This result shows that the electron-electron interactions effectively cancel the unscreened part of the electron-impurity interactions. Moreover, it has been shown by Glasko and Mironov[15] that for nondegenerate statistics and under the condition

$$\frac{4\pi(\hbar e)^2 n}{\kappa(kT)^2 m^*}\lesssim 4\;, \qquad (6)$$

the screening is given by the Debye approximation.[16] The result for screening only by free electrons (no ionized impurities) is well known[16]:

$$q=(4\pi e^2 n/\kappa kT)^{1/2}\;. \qquad (7)$$

It can be noted that Eq. (4) still retains electron–many-impurity interactions (and also the effects of a random impurity distribution, provided R_I is random). It is readily apparent, and has already been pointed out by Bonch-Bruevich,[2] that the summation over impurities can be eliminated if the screening radius ($1/q$) is small compared to the average interimpurity distance. In compensated material this requirement refers to neutral impurities and in n-type material leads to the condition

$$(N_D-N_A-n)^{1/3}\ll q\;, \qquad (8)$$

where N_D is the donor and N_A the acceptor concentration. It can be seen that Eq. (8) will be satisfied as $n\to(N_D-N_A)$, but this situation will prevail only at high temperatures, whereas low temperatures are of interest for the evaluation of ϵ_1. We will, therefore, evaluate Eq. (8) assuming $(N_D-N_A)\gg n$. With this condition and use of Eq. (7), one obtains the requirement

$$(4\pi e^2 n/\kappa kT)^{1/2}\gtrsim 3(N_D-N_A)^{1/3}\;. \qquad (8a)$$

Ranges of validity of Eqs. (3), (6), and (8a) are given in Table I for typical semiconductor parameters. It is apparent that Eq. (8a) is the most restrictive. Although inclusion of ionized-impurity screening will relax the requirement of Eq. (8a) somewhat, this condition is still frequently not satisfied (in fact, it is not satisfied for the GaP situation which we consider in Sec. III). Thus an accurate treatment requires use of both *screened*-impurity potentials and also a *summation* over such potentials.

There have in fact already been a number of papers[17–20] which use a sum of potentials. Three cases have been treated: (i) abrupt[17] (square-well or δ-function) potentials; (ii) unscreened Coulomb electron-impurity potentials[18,19] [the second right-hand term of Eq. (2)]; (iii) a sum of the unscreened electron-impurity potentials plus a constant electron-electron potential[20] [the second and third right-hand terms of Eq. (2) but with the latter term constant]. Qualitatively all these papers give a similar result: With increasing impurity concentration the impurity band broadens and its peak shifts, first moving closer to the band and then merging with the band. Quantitatively the agreement for the broadening and peak shift derived from the different potentials is not good, which precludes direct application of these results to the screened potential. However, if we assume that case (i) above (i.e., the short-range δ potential) is more relevant to the relatively short-range screened Coulomb potential than cases (ii) and (iii) of unscreened Coulomb potentials, one then obtains a rough measure of the importance of summing over potentials, from the results of this case (i). The work of the related papers[17] of Lax and Phillips and of Frisch and Lloyd provides an exact analysis for a one-dimensional δ-function potential in terms of the product $[N(1/K_0)]$ of impurity density per cm (N) and the reciprocal of the bound-state wave-function range (K_0). For three-dimensional hydrogenic impurities, this product corresponds, approximately, to $a_0\,|\,N_D-N_A\,|^{1/3}$, where a_0 is the Bohr radius. The results[17] show a totally negligible shift of the maximum of the distribution towards the band ($<10^{-4}\%$) and only relatively slight broadening, up to $(N/K_0)=0.1$, and in fact only a slight shift ($\lesssim 5\%$) up to $(N/K_0)=0.25$. On the other hand, calculations for GaP (Zn) based on the data of Casey *et al.*[11] and for Ge(As) based on the data of Debye and Conwell[6] show a 50% reduction in activation energy at $a_0\,|\,N_D-N_A\,|^{1/3}$

TABLE I. Concentration limits (cm^{-3}) obtained from Eqs. (3), (6), (8a), and (8b) (using $\kappa = 11$).

	$T = 300\,°K$	$T = 50\,°K$
	Upper limit on n	
Eq. (3) [using $n \le \frac{1}{3}(\kappa kT/e^2)^3$]	2.6×10^{18}	1.2×10^{16}
Eq. (6) $m^* = m$	2.1×10^{19}	5.9×10^{17}
Eq. (6) $m^* = \frac{1}{2}m$	1.1×10^{19}	3.0×10^{17}
	Lower limit on n	
Eq. (8a) $N_D - N_A = 10^{17}$	[Eq. (8a) not	5.3×10^{16}
Eq. (8a) $N_D - N_A = 10^{18}$	applicable]	2.5×10^{17}
	Upper limit on $N_D - N_A$	
Eq. (8b) ($a_0 = 11$ Å)	6×10^{18}	

~0.1–0.2. Thus the effect of *summing* over potentials is unable to account for observed reductions in activation energies. It thus seems plausible to assume that for the range

$$a_0 \left| N_D - N_A \right|^{1/3} \lesssim 0.2 \tag{8b}$$

the main energy reduction is due to screening, i.e., to first order it is not necessary to sum over potentials, and

$$H \approx (p^2/2m^*) - (e^2/\kappa r)e^{-qr} . \tag{9}$$

We thus use Eq. (8b) as a replacement for the more restrictive inequality (8a); validity limits of Eq. (8b) are also given in Table I.

It can still be noted that as a consequence of the random impurity location there will be a range of q values, leading to a broadening of the energy-level distribution. We neglect any such effect in the quantitative treatment; as discussed in Sec. IV, this appears reasonably justified.

B. Screened-Single-Impurity Problem

The solution for the screened hydrogenic Hamiltonian, Eq. (9), gives the impurity activation energy ($\equiv E_A$) as a function of the inverse screening length q. Here it is to be noted that the activation energy will depend on temperature [since $E_A = f(q)$, $q = f(n)$, and $n = f(T)$]. We thus use E_A to denote the activation energy at any one temperature; the quantity ϵ_1 is being used to denote experimentally determined activation energies which have been assumed independent of temperature.

The ground state of the Hamiltonian, Eq. (9), has been evaluated in a convenient dimensionless form by Krieger[21]: The ratio (E_A/E_0) is given as a function of $a_0 q$, where E_0 is the energy and a_0 the Bohr radius of the unscreened impurity [i.e., for $q = 0$ in Eq. (9)]. By carrying out an empirical fit to Krieger's[21] results (as presented in his Fig. 1), one obtains

$$E_A/E_0 \approx 1.00 - 1.81 a_0 q + 0.81 (a_0 q)^2 , \tag{10}$$

valid to $\Delta(E_A/E_0) \le \pm 0.02$.

Evaluation of Eq. (10) with q given by Eq. (7) and appropriate values of a_0 and E_0 (Sec. III) shows that for the GaP (Zn) case of interest the screening by free carriers alone gives a negligible contribution at low temperatures (see Sec. III, Fig. 2). The same conclusion has also been reached by Debye and Conwell[6] for Ge. However, it is well known, for instance from the usual Debye[16] treatment, that screening is determined by an appropriate sum over *all* ionized species.

The difficulty in applying ionized-impurity screening to semiconductors is that the problem now deviates from the usual Debye theory in which the charged species are implicitly assumed to be mobile; in semiconductors the charge carriers satisfy this condition, but at the usual measurement temperatures the ions are spatially fixed. In the earlier[8,9] work, Harvey[8] discussed the problem of this ionic mobility, but proceeded with the calculations as if the ions were mobile at the measurement temperature. Morgan[9] assumed that due to the lack of mobility the closest distance between the ions was given by the average interionic separation, and used this distance as the ionic "radius" in the standard Debye[16] formulation for ions large compared to the screening radius. In the present work, we assume standard screening to apply above the temperature ($\equiv T_m$) at which the majority[22] ion is still mobile, and consider two alternate approximate treatments at lower temperatures. Both these approximations use fixed ionic positions (not done by Harvey[8]) and preferential neutralization of appropriately located donors (for n-type material) once the carriers start freezing out (not used in the Morgan[9] approach). Moreover, both give screening intermediate between Harvey's[8] and Morgan's.[9]

The first treatment, discussed in more detail in the Appendix, assumes that the ions screen as long as the majority ion is still mobile[22] (to the temperature T_m), and that, subsequently, the free carriers neutralize the ions such as to maintain a self-consistent potential. The result for n-type material is

$$q^2 = \frac{4\pi e^2}{\kappa k}\left(\frac{n}{T} + \frac{N_A}{T_m} + \frac{N_A + n}{T_m}\right) . \tag{11}$$

The first term is the contribution of the free carriers, the second is from the acceptors, all of which are ionized, and the third term corresponds to the ionized donors. It can be noted that Eq. (11) also follows if one assumes additivity of electron and ion screening lengths, with the ions screening as if they were at a temperature T_m.

The second approximate treatment is that given

by Falicov and Cuevas[23] in a calculation of ionized-impurity scattering. Here it is assumed that the electrons preferentially neutralize donors which are far from acceptor sites, such that the remaining donor ions are preferentially close to the acceptor sites, i.e., the system tends to the configuration of minimum electrostatic energy. For an exponential fall-off in the resultant ion-pair correlation, and for an electron concentration negligible compared to the ionized-acceptor concentration, one obtains [Falicov and Cuevas,[23] Eq. (2.11)]

$$q_{FC}=[8\pi(N_D-N_A)]^{1/3}\,. \qquad (12)$$

Combination of Eqs. (11) and (12) with Eq. (10) gives the respective energy changes for n material:

$$\left(\frac{E_A}{E_0}\right)\approx 1.00-1.81a_0\left(\frac{4\pi e^2}{\kappa k}\right)^{1/2}\left(\frac{n}{T}+\frac{2N_A+n}{T_m}\right)^{1/2}+0.81a_0^2\left(\frac{4\pi e^2}{\kappa k}\right)\left(\frac{n}{T}+\frac{2N_A+n}{T_m}\right), \qquad (13)$$

$$\left(\frac{E_A}{E_0}\right)_{FC}\approx 1.00-11.4a_0(N_D-N_A)^{1/3}+32.0a_0^2(N_D-N_A)^{2/3}\,. \qquad (14)$$

For p-type material, n is replaced by the hole concentration p, and N_A by N_D. It can be noted that at low temperatures the condition

$$(n/T)\ll(2N_A/T_m) \qquad (15)$$

will in general apply, and it then follows from Eq. (13) that in this range the activation energy is independent of temperature. As the temperature increases and Eq. (15) is no longer valid, the activation energy will decrease as n increases. (There will also be some high-temperature increase in E_A once n saturates, but this is not of interest in the usual measurement range of Hall data.) No such temperature dependence of activation energy is shown by Eq. (14); however, as stated earlier, this equation is valid only under the condition

$$n\ll N_A\,. \qquad (16)$$

At higher temperatures, beyond the validity range of Eq. (16), one would expect an improved Falicov-Cuevas treatment to also give a decreasing activation energy. Thus, the present treatment predicts an activation energy which is independent of temperature at low temperatures, but which subsequently decreases at higher temperatures.

III. COMPARISON TO EXPERIMENT

A. Reevaluation of Hall Data

Here, we are interested in a comparison to data[11,12] on p-type GaP (Zn). It is apparent from Eqs. (13) and (14) that for a comparison to experimental results (in p-type material) one requires values of E_A, N_D, N_A, and of hole concentration p, all as a function of temperature. These quantities are most frequently obtained from Hall measurements. An analysis of such data gives the quantity p directly [to within the Hall factor r which is the ratio of Hall mobility (μ_H) to drift mobility (μ_D), and which is of order unity], but E_A, N_D, and N_A must be obtained by analysis of p as a function of temperature, via the standard[6,11,12] equation:

$$\frac{p(p+N_D)}{N_A-N_D-p}=\frac{N_0(Tm^*/m)^{3/2}}{g}\,e^{-E_A/kT}\,, \qquad (17)$$

where $N_0=2(2\pi mk/h^2)^{3/2}$, g is the degeneracy factor, and m is the free-electron mass.

For E_A independent of temperature, i.e., at low temperatures, within the validity range of Eqs. (15) and (16), and with the additional (low-temperature) condition of

$$p\ll N_A-N_D\,, \qquad (18)$$

Eq. (17) gives a linear plot of log $(p/T^{3/2})$ (or, approximately, log p) vs $1/T$, with a slope[24] of (E_A/k). The slope of such a plot gives E_A, and N_A/N_D can be obtained from the intercept I in terms of g and m^*:

$$I=[(N_A/N_D)-1][N_0(m^*)^{3/2}/g]\,. \qquad (19)$$

For acceptors in GaP, g is customarily taken as 4, but m^* is less well known. We have therefore used two values[25] in our analysis: $m^*/m=0.6$ and $m^*/m=0.9$.

In attempting to apply this analyis to measured[11,12] samples, we found that there was only *one* sample which is sufficiently impure to give a reduction in E_0 greater than 10%, and which simultaneously gave a long linear low-temperature plot. {It must be noted that the activation energies ϵ_2 and ϵ_3 [Eq. (1)] become important at low temperatures.} This was sample No. 3 of Casey *et al.*[11] (hence referred to as CEW), and we have reanalyzed this sample. The low-temperature part of the curve gave $E_A=28$ meV, and $N_D/N_A=0.27$ and 0.39 for $m^*/m=0.6$ and 0.9, respectively.

A complete evaluation of sample parameters also requires a value of N_A [or equivalently, (N_A-N_D)]. Usually this is derived by curve fitting with the assumption of a temperature-independent activation

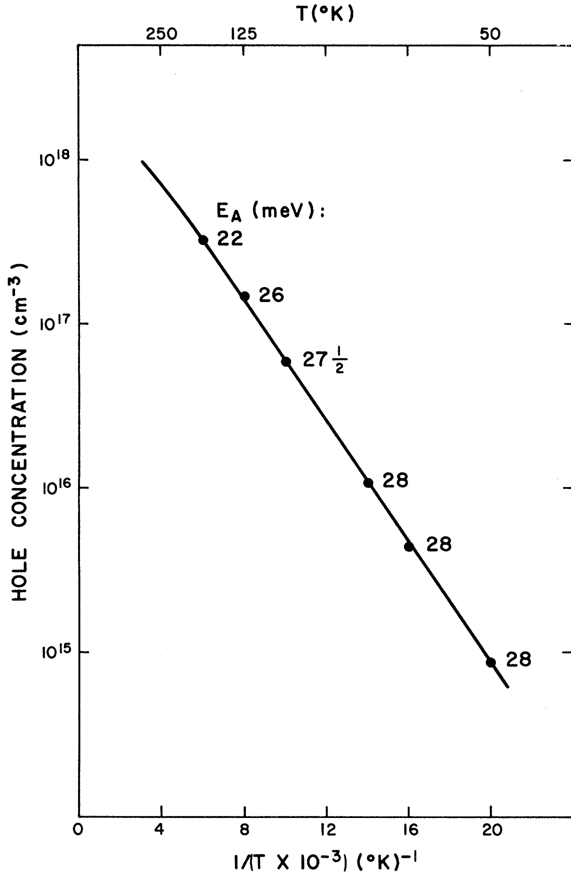

FIG. 1. Plot of hole concentration vs reciprocal temperature on a semilogarithmic scale. The solid line is the data for CEW (Ref. 11) sample No. 3, and the points are calculated by Eq. (14) with $g=4$, $(m^*/m)=0.9$, $N_A=1.6\times10^{18}$ cm^{-3}, $N_D=6\times10^{17}$ cm^{-3}, and E_A as given in the figure.

energy, However, as discussed in Sec. II B, this assumption is invalid except at low temperatures, within the validity range of Eqs. (15) and (16). We have therefore reanalyzed the data, with $E_A=F(T)$, as follows: (i) The values of (N_A-N_D) are estimated using the high-temperature part of the curves, assuming these to be close to saturation. Reasonable values were found to be $N_A-N_D=1.1\times10^{18}$ cm^{-3}, $N_A-N_D=1.0\times10^{18}$ cm^{-3} for $m^*/m=0.6$ and 0.9, respectively. [A slightly higher value is required for lower m^*, as follows from Eq. (17).] (ii) With N_A and N_D now determined, E_A at any one temperature is obtained from the observed[11] hole concentration [by Eq. (17)]. The results of the above analysis are shown in Fig. 1 for $m^*/m=0.9$. The results for $m^*/m=0.6$ are similar.

B. Comparison to Neutron-Activation A_A Value

The analysis discussed in the Sec. III A has given a zinc concentration (N_A) of $(1.5-1.6)\times10^{18}$ cm^{-3}. These values are much closer to the CEW[11] neutron-activation results of $(0.7-1.0)\times10^{18}$ cm^{-3} than the value of 3.5×10^{18} cm^{-3} obtained from their Hall analysis,[26] which assumed E_A independent of T. These N_A results are summarized in Table II. That the present values are closer to the neutron values is not surprising, since with a reduced high-temperature activation energy, a lower value of N_A will give the same value of p. As to the residual discrepancy between our values and the neutron data, we feel that this is within error limits, considering the following three factors: (i) The first is possible errors in the neutron activation. (ii) The Hall factor (μ_H/μ_D) is unknown. As already pointed out by CEW, the value for p-type GaP may well be lower than the assumed value of unity, possibly closer to 0.7 as in p-type Ge. Such a value would reduce the hole concentrations, and consequently N_A, by a corresponding factor. (iii) There may be additional deep acceptors, as indicated, for instance, for CEW sample No. 1 (see their Fig. 5).

C. Screening Calculations

Evaluation of the reduction in activation energy via Eqs. (13) and (14) requires values of a number of parameters of two classes: (i) those which depend on the particular sample (N_A, N_D, p) and (ii) those which are independent of the sample (κ, E_0, a_0, T_m). For the first class, the impurity concentrations N_A and N_D have been evaluated in Sec. III A, with the hole concentration p known from the CEW Hall data. As for the remaining parameters, we used $\kappa=11.1$ (Ref. 27), $E_0=64$ meV (Ref. 28), and $a_0=10-12$ Å (Ref. 29). For T_m, the lower bound on the temperature range over which the Zn ion is still mobile, it is known[30] that Zn-O pairs can form at temperatures as low as 770–870 °K, so $T_m=800$ °K appears as a conservative estimate.

The activation energy E_A has been calculated from Eqs. (13) and (14) using the above parameter values. The results, in terms of the reduction in activation energy (E_0-E_A), are given in Table III. Values of (E_0-E_A) obtained from the analysis of the Hall data (Sec. III A) are also given. The results for one case ($m^*/m=0.9$, $a_0=12$ Å) are shown

TABLE II. N_A of CEW (Ref. 11) sample No. 3.

	N_A (cm^{-3})
CEW Hall analysis	3.5×10^{18}
CEW neutron activation:	
nondestructive	$1\ \times10^{18}$
destructive	0.7×10^{18}
Present reanalysis of CEW data	
$m^*/m=0.6$	1.5×10^{18}
$m^*/m=0.9$	1.6×10^{18}

TABLE III. Activation energies for CEW (Ref. 11) sample No. 3 (E_0-E_A) (meV).

	$m^*/m=0.6$, $N_A=1.5\times10^{18}$ cm^{-3}, $N_D=4\times10^{17}$ cm^{-3}			$m^*/m=0.9$, $N_A=1.6\times10^{18}$ cm^{-3}, $N_D=6\times10^{17}$ cm^{-3}		
	Calculated from Eq. (13)		Reanalysis of	Calculated from Eq. (13)		Reanalysis of
T	$a_0=10$ Å	$a_0=12$ Å	CEW data	$a_0=10$ Å	$a_0=12$ Å	CEW data
50	16	19	36	19	22	36
71	17	20	36.5	20	23	36
100	20	23	38	22	26	36.5
125	23	27	41	25	29	38
167	27.5	32	45	30	34	42
	Calculated from Eq. (14)			Calculated from Eq. (14)		
	31	35.5		30	35	

$T_m=\infty$ °K corresponding to screening by only free carriers. These results are also shown in Fig. 2. It can be seen that free-carrier screening is indeed negligible at low temperatures, whereas if the Zn ions do have some mobility down to 400 °K, Eq. (13) also gives very good agreement.

IV. DISCUSSION AND CONCLUSIONS

The theoretical treatment of Sec. II has predicted a number of features connected with the impurity activation energy in the intermediate doping range, with these predictions checked against experimental results in Sec. III. To gauge the extent of the agreement, we enumerate the following points: (i) The dominant effect on the reduction of activation energy with impurity concentration is electrostatic screening, i.e., the effective Hamiltonian is approximately given by Eq. (9); (ii) the screening due to ionized impurities must be added to that of the free carriers; (iii) although it is well known[2] that there will be some distribution in activation energies (a broadening), due to a summation over potentials and due to the randomness in the screening effect, this broadening is assumed relatively slight in the parameter range of interest; (iv) the screening results in an activation energy which decreases with increasing temperature over an appreciable temperature range for usual temperatures used in Hall measurements; (v) the "intermediate" doping range is considered to extend from impurity concentrations such that the activation energy is decreased from its value at infinite dilution (E_0) through a range limited at the upper end by the most restrictive of inequalities (3), (6), or (8b) (Table I).

In interpreting the comparison to the experimental data, it should first be pointed out that the analyzed sample (CEW No. 3) does satisfy the criteria of the intermediate doping range [point (v), above].

The quantitative evaluation of screening with inclusion of ionic screening (Sec. III C, Fig. 2) shows as well in Fig. 2. It can be seen that the theoretical results of Eq. (14) give very good agreement, within their validity range, with the data. The results of Eq. (13), with $T_m=800$ °K, are somewhat less satisfactory, but even so can account for 50–80% of the observed reduction.

For comparison, two further calculations were carried out with Eq. (13): (i) the $m^*/m=0.9$, $a_0=12$ Å case with $T_m=400$ °K, and (ii) the case

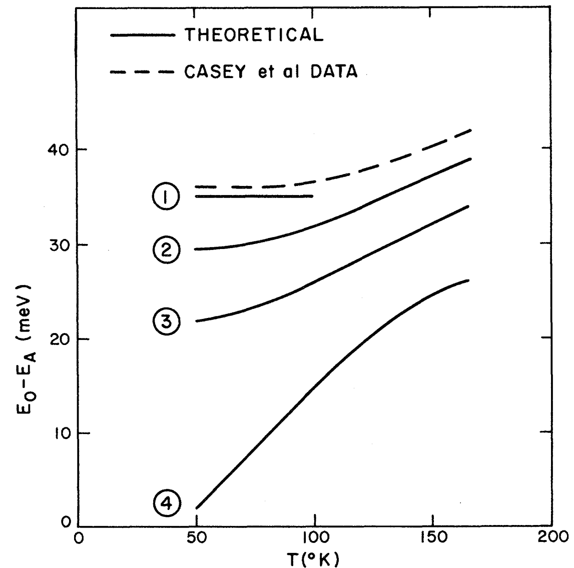

FIG. 2. Plot of the reduction in activation energy (E_0-E_A) vs temperature. The dashed line is obtained from E_A as given by the data of CEW (Ref. 11) sample No. 3, assuming $m^*/m=0.9$, and using the analysis given in Sec. II A (i.e., using the E_A values of Fig. 1). The solid lines are theoretical, evaluated for $a_0=12$ Å, and the other parameters are as discussed in the text; curve 1 is obtained from Eq. (14), and curves 2–4 from Eq. (13) with $T_m=400$, 800, and ∞ °K, respectively. The last curve 4 corresponds to screening by only free carriers.

that the Falicov-Cuevas approach [Eqs. (12) and (14)] gives quite satisfactory agreement with the observed activation energies, and even the more conservative approach of Eqs. (11) and (13) with $T_m = 800\,^\circ$K can account for 50–80% of the observed reduction in activation energy. Moreover, if ionic mobility sufficient for screening (where only slight motion is required) does persist to lower temperatures (e.g., 400 °K—see Fig. 2), both screening approaches give roughly equally good agreement. This thus confirms the applicability of the approximations leading to Eq. (9) [point (i), above]. It can also be seen that screening by free carriers alone is negligible at low temperatures (curve 4 of Fig. 2, $T_m = \infty\,^\circ$K), and that satisfactory agreement is obtained only with inclusion of ionic screening [point (ii)]. And, since good agreement is obtained without consideration of broadening effects, their neglect appears justified [point (iii)].

The results discussed in Sec. III B show that use of an activation energy which decreases with increasing temperature [point (iv)] clears up an annoying discrepancy between the Hall and neutron-activation data on CEW[11] sample No. 3. Moreover, any broadening of the distribution would give the inverse dependence, i.e., an activation energy which increases with temperature (the shallow levels empty first). It can thus again be concluded that broadening effects are relatively minor [point (iii)].

It can also be noted that, as mentioned in Sec. III A, only one sample with $E_0 > \epsilon_1 > 0$ out of all those measured by CEW[11] and van der Does de Bye and Peters[12] (a total of 10) showed a long (greater than about one order of magnitude in p) linear range of log p vs $1/T$. This somewhat puzzling fact can easily be accounted for by a temperature-dependent activation energy [point (iv)].

Over-all, one can thus conclude that a shift in the center of the level distribution due to screening provides the *main* contribution to the decrease of activation energy with concentration in the intermediate doping range, *provided* that ionic species are included in the screening, and also that a consequent temperature dependence of the activation energy gives an improved understanding of experimental results.

Finally, in view of the dependence of activation energy on temperature, some comments on the reliability of analyses of Hall data as a function of temperature appear in order. (a) For $\epsilon_1 \approx E_0$, the usual approach [Eq. (17) with E_A constant] is expected to be fully valid. (b) For $\epsilon_1 < E_0$, but for samples showing long linear regions, an analysis such as carried out in Sec. III A should be quite satisfactory. However, in this case $E_A(T)$ is not constant. From this it also follows that now one cannot determine m^* by a least-squares fit to the carrier-concentration–vs–temperature data, as has sometimes been done in the literature. (c) For $\epsilon_1 < E_0$ and short linear regions, we doubt that much reliable information, other than the carrier concentration, can be obtained.

Note added in proof. Recent pair emission results of J. A. W. van der Does de Bye, A. T. Vink, A. J. Bosman, and R. C. Peters, J. Luminescence 3, 185 (1970), give $a_0 \approx 15$ Å, i.e., a value larger than the 10–12 Å assumed here. Such a value would lead to improved agreement with experiment for the theoretical values of E_A as given by Eq. (13), and a slight overestimate by Eq. (14).

ACKNOWLEDGMENTS

The author gratefully acknowledges helpful discussions with, and comments by, Dr. R. N. Bhargava, Dr. S. K. Kurtz, Dr. T. N. Morgan, Dr. F. K. du Pré, and Dr. P. J. Price.

APPENDIX

In order to derive Eq. (11), we will first summarize the standard[16] Debye-Hückel approach in a form appropriate to n-type semiconductors. In our case these equations are unmodified for $T \geq T_m$.

$$\nabla^2 V = -4\pi\rho/\kappa \ , \tag{A1}$$

$$\rho = e(N_+ - N_- - n) \ , \tag{A2}$$

$$(N_\pm, n_-) = (N^I, n)\, e^{\mp eV/kT} \approx (N^I, n)\,[1 \mp (eV/kT)] \ , \tag{A3}$$

where ρ is the net charge, N_+ is the excess positive ions (ionized donors), N_- is the excess negative ions (ionized acceptors), n_- is the excess electrons, and N^I is the average concentration of appropriate ionic species. (Note that $N_A^I = N_A$ for n-type semiconductors.)

Combination of Eqs. (A1)–(A3) gives

$$\nabla^2 V = -(4\pi e^2/\kappa kT)\,(N_D^I + N_A + n)\,V, \tag{A4}$$

with the solution

$$V = (e/\kappa r)\, e^{-qr} \ , \tag{A5}$$

$$q^2 = (4\pi e^2/\kappa kT)(N_D^I + N_A + n) \ . \tag{A6}$$

We now deviate from the standard approach, and assume that ionic mobility ceases at some temperature T_m (see below for an analysis of this assumption). Since T_m will generally be high, we also assume that all impurities are ionized at $T = T_m$ [i.e. $N_D^I(T_m) = N_D$], and remain so down to some temperature T_i. For $T_m > T > T_i$, one then obtains from Eq. (A3)

$$(N_+ - N_-) \approx (N_D - N_A) - (N_D + N_A)\,(eV_m/kT_m) \ , \tag{A7}$$

where V_m is the potential at $T = T_m$, and this leads

to a modified Eq. (A4)

$$\nabla^2 V = -(4\pi e^2/\kappa k)\left[(N_D+N_A)(V_m/T_m)+n(V/T)\right] . \quad \text{(A8)}$$

For temperatures below T_i, carrier capture takes place, and we assume that at some temperature close to T_i this is sufficient to modify the average ionic potential such that V_m approaches the potential (V) required to satisfy Eq. (A3) for the free carriers, i.e., such that Eq. (A8) becomes

$$\nabla^2 V = -(4\pi e^2/\kappa k)\left[(N_D^I+N_A)/T_m+n/T\right]V . \quad \text{(A9)}$$

This equation has the same functional dependence on voltage as the standard equation (A4) and with use of $N_D^I=(N_A+n)$ gives the screening of Eq. (11).

As to the approximation that ionic mobility ceases abruptly at $T=T_m$, this depends on use of an activated process for diffusion, with the concomitant rapid decrease in mobility. As a specific example, we again consider GaP(Zn) (similar qualitative behavior would, of course, be expected for other cases). Reliable values of the parameters required for the drift velocity (v_d) of Zn in GaP are not yet available; however, approximate values can be estimated from the literature, assuming that the relevant drift velocity is that of interstitial Zn (and using $v_d=\mu E$, $\mu=eD/kT$):

diffusion coefficient $(D)\approx 10^7 e^{-2.5/kT}$ cm^2/sec (Ref. 31),

internal fields $(E)\approx 10^4$ V/cm (from estimates of Ref. 32).

Results for the diffusion coefficient and for the Zn drift velocity are as follows:

T(°K)	Diff. coeff. (cm^2/sec)	v_d(Å/h)
600	1×10^{-14}	700
550	1×10^{-16}	8
500	6×10^{-19}	0.05

Since it seems unreasonable that a sample could equilibrate with $v_d \lesssim 1$ Å/h, effective ionic mobility for these parameters ceases between ~550 and ~525 °K, i.e., within a quite narrow temperature interval. Slightly different parameters (but in a reasonable range) would be expected to give a similar narrow temperature interval for no further motion, thus justifying use of an arbitrary parameter T_m. [Such slightly different values could, however, easily lead to values of T_m as given in the text as examples (400, 800 °K).]

[1]G. L. Pearson and J. Bardeen, Phys. Rev. 75, 865 (1949).

[2]V. L. Bonch-Bruevich, *Semiconductors and Semimetals*, Vol. I, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), p. 101.

[3]See, e.g., B. I. Halperin and M. Lax, Phys. Rev. 148, 722 (1966).

[4]H. Fritzsche, Phys. Rev. 99, 406 (1955).

[5]See, e.g., N. F. Mott and E. A. Davis, Phil. Mag. 17, 1269 (1968).

[6]P. P. Debye and E. M. Conwell, Phys. Rev. 93, 693 (1954).

[7]L. Pincherle, Proc. Phys. Soc. (London) A64, 663 (1951).

[8]W. W. Harvey, Phys. Rev. 123, 1666 (1961).

[9]T. N. Morgan, Phys. Rev. 139, A343 (1965).

[10]R. H. Saul, J. Armstrong, and W. H. Hackett, Jr., Appl. Phys. Letters 15, 229 (1969); R. H. Saul and W. H. Hackett, Jr., J. Electrochem. Soc. 117, 921 (1970).

[11]H. C. Casey, Jr., F. Ermanis, and K. B. Wolfstirn, J. Appl. Phys. 40, 2945 (1969).

[12]J. A. W. van der Does de Bye and R. C. Peters, Philips Res. Rept. 24, 210 (1969).

[13]G. W. Castellan and F. Seitz, *Semiconducting Materials* (Butterworths, London, 1951), p. 8.

[14]E. Hanamura, J. Phys. Soc. Japan 28, 120 (1970).

[15]V. B. Glasko and A. G. Mironov, Fiz. Tverd. Tela 4, 336 (1962) [Sov. Phys. Solid State 4, 241 (1962)].

[16]See, e.g., R. M. Fuoss and F. Accascina, *Electrolytic Conductance* (Interscience, New York, 1959).

[17]See, e.g., M. Lax and J. C. Phillips, Phys. Rev. 110, 41 (1958); H. L. Frisch and S. P. Lloyd, *ibid.* 120, 1175 (1960).

[18]T. Matsubara and Y. Toyozawa, Progr. Theoret. Phys. (Kyoto) 26, 739 (1961).

[19]N. Majlis, Proc. Phys. Soc. (London) 90, 811 (1967).

[20]M. von Ortenberg, Phys. Letters 28A, 715 (1969).

[21]J. B. Krieger, Phys. Rev. 178, 1337 (1969).

[22]In considering the similar problem of ion mobility for ion pairing, H. Reiss [J. Chem. Phys. 25, 400 (1956)] was able to show that for equal numbers of positive and negative ions the same equations apply for one ion mobile as for all ions mobile. The same should thus hold with mobile majority ions.

[23]L. M. Falicov and M. Cuevas, Phys. Rev. 164, 1025 (1967).

[24]If N_D is very low such that $N_D \ll p \ll N_A - N_D$, a linear low-temperature region can also be obtained with a slope of $(E_A/2k)$, but we believe such low values of N_D are unlikely in the present "state of the art" of GaP.

[25]Effective-mass ratios of 0.43–0.54 are obtained by Casey *et al.* (Ref. 11) for their pure samples; values of 0.6–1.3 are obtained by van der Does de Bye and Peters (Ref. 12), and a theoretical value of ≈ 0.9 is obtained by R. A. Faulkner as quoted by J. S. Jayson, R. N. Bhargava, and R. W. Dixon, J. Appl. Phys. 41, 4972 (1970).

[26]CEW obtain an alternate $N_A = 2.5\times10^{18}$ cm^{-3} (their Table V) by analyzing the curve with use of a low N_D value (Ref. 24). We believe the required $N_D = 1\times10^{14}$ is unreasonably low.

[27]A. S. Barker, Jr., Phys. Rev. 165, 917 (1968).

[28]P. J. Dean, C. J. Frosch, and C. H. Henry, J. Appl. Phys. 39, 5631 (1968).

[29]K. Weiser, Phys. Rev. Letters 21, 1337 (1968).

[30]A. Onton and M. R. Lorenz, Appl. Phys. Letters 12, 115 (1968).

[31]Calculated from a diffusion activation energy of 2.5 eV given by L. L. Chang and G. L. Pearson, J. Appl. Phys. 35, 374 (1964) and from an estimate for the rate of interstitial diffusion of $\sim 10^{-4}$ cm^2/sec at $\sim 1200\,^\circ$K given by L. L. Chang and G. L. Pearson, J. Appl. Phys. 35, 1960 (1964). It can be noted that the 2.5 eV was obtained at high temperatures, with a lower activation energy at lower temperatures, but S. F. Nygren and G. L. Pearson [J. Electrochem. Soc. 116, 648 (1969)] have since attributed the low-temperature results to diffusion along dislocations; such diffusion is obviously of no relevance to the present work.

[32]D. Redfield, Phys. Rev. 130, 914 (1963).

PHYSICAL REVIEW B VOLUME 5, NUMBER 2 15 JANUARY 1972

Identification of Γ Transitions in the E_0' Region of Germanium by Piezoreflectance Measurements

D. D. Sell and E. O. Kane
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 7 July 1971)

The reflectance of germanium and its logarithmic derivative $(dR/dE)/R$ in the 2.8- to 4.0-eV spectral region have been reexamined as a function of temperature (2 to 300 K) and uniaxial stress. Two relatively sharp structures at 3.00 and 3.19 eV are seen in the unstressed derivative data at 2 K. The observed splittings and polarizations of these two structures for large uniaxial stress along [001] or [111] directions agree well with the behavior calculated for $\Gamma_8^v \to \Gamma_7^c$ and $\Gamma_8^v \to \Gamma_8^c$ transitions, respectively. The data are incompatible with a Δ symmetry assignment for these two structures. A third broader structure is seen near 3.5 eV, the expected energy for a $\Gamma_7^v \to \Gamma_8^c$ transition; however, the temperature independence and strength of this structure indicate that it is not a Γ transition.

I. INTRODUCTION

Structures commonly denoted as E_0, E_1, E_0', and E_2 have been observed in a large number of semiconductors in groups IV, III-V, and II-VI.[1] Of these, the E_0 (at Γ) and E_1 (along Λ) have been studied extensively and are reasonably well understood. In comparison, the E_0' and E_2 regions have received relatively little attention and the nature of these transitions has not been thoroughly established. In this paper we are concerned with the E_0' region in germanium. E_0' structure near 3.1 eV was seen in the early reflectance data of Philipp and Taft[2] and the transmission data of Cardona and Harbeke.[3] In later work Potter[4] (oblique-angle reflectance) and Ghosh[5] (electroreflectance) reported additional structures in the 2.6–3.2-eV region. Zucca and Shen[6] and Braunstein and Welkowski[7] studied the reflectance in this region by the wavelength-modulation technique. In general, the observed structure is weak and relatively difficult to study. The data obtained by the different techniques, and the interpretations of these data, are not completely compatible. These results do not provide the evidence needed to identify the observed structure with particular electronic transitions. Band calculations associate the E_0' structure with transitions at,[8] or near,[9] Γ ($\Gamma_5^v \to \Gamma_5^c$)[10] or possibly with $\Delta_5 \to \Delta_1$ transitions along an extended region of the Δ-symmetry line.[11,12]

Recently, Fischer and co-workers[13] have presented direct evidence from transverse electroreflectance and photoemission experiments which identifies a portion of the E_0' structure in germanium as $\Gamma_5^v \to \Gamma_5^c$ transitions. Their results prompted us to reexamine the reflectance in this region at low temperatures as a function of large uniaxial dc stress. The splittings and polarization behavior which we observe for [111] and [001] stress provide conclusive proof that their assignments are correct. At 2 K the 3.00- and 3.19-eV structures are $\Gamma_8^v \to \Gamma_7^c$ and $\Gamma_8^v \to \Gamma_8^c$ transitions, respectively. Our data are incompatible with a Δ-symmetry assignment for these transitions.

These results are important for two reasons: First, the transition energy provides an additional parameter for band calculations which determines the position of the second conduction band at the center of the zone. Second, in view of the similarity of semiconductor spectra, our results for germanium suggest that similar results may be observed in other materials.

II. SYMMETRY ANALYSIS OF Γ TRANSITIONS FOR [001] AND [111] STRESS

In this section we consider the polarization selection rules for optical transitions between Γ_5(p-like) states subjected to uniaxial stress along the [001] or [111] crystallographic directions. At first, this problem appears to be a rather trivial extension of earlier stress studies[14] of E_0 structure ($\Gamma_5 \to \Gamma_1$). In fact, the problem of rotating from