

## Broadening of Impurity Levels in Silicon\*†

MELVIN LAX,† *Physics Department, Syracuse University, Syracuse, New York, and Crystal Branch, Solid State Division, United States Naval Research Laboratory, Washington, D. C.*

AND

ELIAS BURSTEIN, *Crystal Branch, Solid State Division, United States Naval Research Laboratory, Washington, D. C.*

(Received July 12, 1955)

Sharp absorption lines have been observed in *p*- and *n*-type silicon whose absolute and relative positions lead to their interpretation as optical transitions between bound states of trapped holes or electrons that are approximately hydrogenic in character. These lines have a finite breadth of the order of 0.001 electron-volt at liquid helium temperatures. This breadth is determined by the zero-point vibrations of the lattice. At higher temperatures, the squared breadth increases in proportion to the mean squared amplitude of oscillation of those lattice modes that contribute significantly to the broadening. The theory indicates that the modes of importance have wavelengths of the order of the Bohr radius of the trapped carrier state. These are rather long wave-

length acoustic modes whose energy  $\hbar\omega$  corresponds to 80°K. Thus the squared broadening is expected to increase by a factor of two from helium to nitrogen temperatures—and this increase is confirmed experimentally. This confirms the hypothesis that the broadening is due to the interaction of the trapped electron with the acoustic lattice vibrations. The form of the electron-lattice interaction is taken to be that of the Bardeen-Shockley deformation potential, and its strength is determined from the experimental mobility. Thus the theory contains no adjustable parameters. In absolute magnitude, the theoretical line breadth turns out to be several times too large. Possible reasons for the discrepancy are discussed.

### 1. INTRODUCTION

EARLY work on the broadening of absorption or luminescence lines is based on a classical or semiclassical application of the Franck-Condon principle.<sup>1</sup> Williams and Hebb<sup>2</sup> have made a quantum mechanical treatment of luminescence in thallium activated potassium chloride using a single configurational coordinate. Huang and Rhys<sup>3</sup> in treating *F*-center absorption made the first quantum mechanical treatment of a many-coordinate problem. Their technique is limited, however, by the requirement that all the normal coordinates have the same frequency.

The present problem—broadening of impurity levels in homopolar semiconductors—cannot be treated by any of the previously available methods: (1) The criterion previously developed for the validity of the semiclassical Franck-Condon principle<sup>1</sup> will be shown to be violated [see Eq. (5.37)]. (2) The orbits of impurity trapped electrons are so large that interaction takes place with many ( $\sim 10^3$ ) atoms<sup>4</sup> and a Williams-type treatment based on a single real coordinate is inadequate.

(3) Acoustical modes will be more important for broadening at all temperatures below the Debye temperature. The spread in frequencies in these modes prohibits the use of the Huang-Rhys analysis.

A generalization of the Huang-Rhys method to the case of modes with an arbitrary frequency distribution was therefore developed.<sup>1</sup> The relationship between this generalized procedure, the Huang-Rhys method and the more recent modification introduced by O'Rourke<sup>5</sup> will be discussed in Sec. 2. In this section, we shall attempt to summarize the physical assumptions made in carrying out the calculation.

The chief approximation made in the course of our calculation has been that we have dealt with the broadening of impurity levels of a hydrogenic type. A hydrogenic energy level scheme will not be quite right unless one has a single nondegenerate band whose band edge is at  $k=0$ , so that a simple effective mass theory is valid. It is now clear from cyclotron resonance experiments<sup>6</sup> that the valence band of both silicon and germanium is degenerate, and the conduction bands have minima at points other than  $k=0$ . At the time the calculations reported here were made, however, this information was not available, and the simplest possible assumptions were made. It seems worth while to report these calculations in their original form because they illustrate the essential mechanism for broadening without leading to undue complications of a purely computational nature.

The results will be compared with experiments on boron-doped silicon (see Figs. 1 and 2) because the most reliable and complete experimental information is avail-

\* A summary of the results contained here was reported to the American Physical Society in March, 1953 [Phys. Rev. **91**, 208 (1953)].

† Supported by the Department of the Navy (Office of Naval Research) and the Department of the Air Force (Office of Scientific Research). Publication of this document, for any purpose, by the United States Government, is authorized.

‡ Present address: Bell Telephone Laboratories, Murray Hill, New Jersey, on leave of absence from Syracuse University.

<sup>1</sup> For references to earlier work and a critical analysis of the Franck-Condon principle see M. Lax, J. Chem. Phys. **20**, 1752 (1952) hereafter referred to as I. See also reference 14.

<sup>2</sup> F. E. Williams and M. H. Hebb, Phys. Rev. **84**, 1181 (1951).

<sup>3</sup> K. Huang and A. Rhys, Proc. Roy. Soc. (London) **A204**, 406 (1950).

<sup>4</sup> The number of atoms in a Bohr sphere is  $(4\pi a^3/3)(2n_0) \approx 500$  atoms where  $n_0 = 2.5 \times 10^{22}$  primitive cells per  $\text{cm}^3$ , 2 = number of atoms per primitive cell, and  $a$ , the Bohr radius of the impurity state is shown in Eq. (1.2) to be about 13.5 Å for *p*-type silicon doped with boron.

<sup>5</sup> R. C. O'Rourke, Phys. Rev. **91**, 265 (1953).

<sup>6</sup> Dresselhaus, Kip, and Kittel **92**, 827 (1953); Lax, Zeiger, Dexter, and Rosenblum, Phys. Rev. **93**, 1418 (1954); R. N. Dexter and B. Lax, Phys. Rev. **96**, 223 (1954); Dexter, Lax, Kip, and Dresselhaus, Phys. Rev. **96**, 222 (1954).

able for this case<sup>7</sup>: transitions to excited impurity states have been observed, and a level scheme has been worked out. The unusually good agreement of the empirical level scheme with the hydrogen model is slightly fortuitous in view of the degeneracy of the valence band in silicon. However, this agreement made it easy to assign quantum numbers to the level scheme. If the correct wave functions for acceptor impurity states were available, they could be inserted into (3.1) to yield an estimate for the broadening in the presence of degenerate bands. However, only donor states have been investigated theoretically in any detail, and for these the energy levels are better known than the wave functions.<sup>8</sup>

The basic physical assumption made in this paper is that the broadening of the impurity levels is produced by the lattice vibrations through the electron-lattice interaction. Using the Born-Oppenheimer viewpoint, we may say that the energy of a trapped electron (or hole) is a function of the nuclear coordinates. The energy difference between a ground and excited impurity state will depend on the nuclear configuration. The frequency of the absorbed radiation will then vary as the nuclei oscillate, thus leading to a broadening of the absorption line. The mean square width of the line will then be proportional to the mean square amplitude of nuclear oscillations. Thus, at absolute zero there will be a zero-point width associated with the zero-point oscillations of the lattice. The squared width will increase gradually with temperature. At temperatures at which the classical equipartition law becomes valid,  $kT \gtrsim \hbar\omega$  (where  $\omega$  is a frequency of a typical mode of importance in the

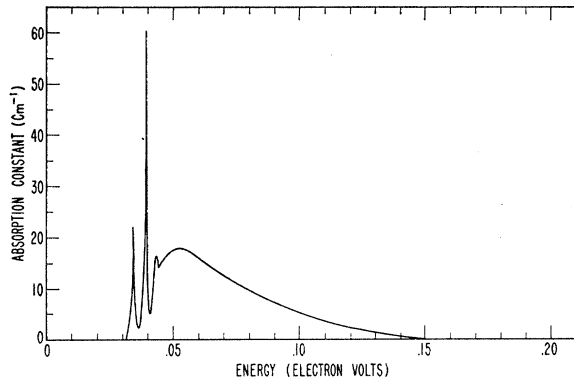


FIG. 1. Infrared absorption of a single crystal of silicon doped with boron [Burstein, Bell, and Hennis (unpublished)]. Further discussion of this data may be found in Burstein, Bell, Davisson, and Lax, *J. Phys. Chem.* **57**, 849 (1953).

<sup>7</sup> Burstein, Bell, Davisson, and Lax, *J. Phys. Chem.* **57**, 849 (1953). This paper establishes the existence of excited impurity states which hitherto had only been conjectured. See this paper for a summary of the experimental procedure. Qualitatively similar results have also been obtained by R. Newman for In-doped silicon, *Phys. Rev.* **99**, 465 (1955).

<sup>8</sup> C. Kittel and A. H. Mitchell, *Phys. Rev.* **96**, 1488 (1954); M. A. Lampert, *Phys. Rev.* **97**, 352 (1955); W. Kohn and J. Luttinger, *Phys. Rev.* **97**, 1721 (1955); **98**, 915 (1955). Since this work was completed, wave functions for the acceptor states have also been obtained by W. Kohn (private communication).

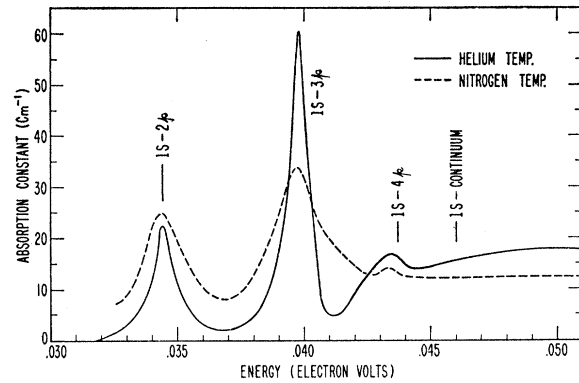


FIG. 2. Infrared absorption of a single crystal of silicon doped with boron [Burstein, Bell, and Hennis (unpublished)]. Further discussion of this data may be found in Burstein, Bell, Davisson, and Lax, *J. Phys. Chem.* **57**, 849 (1953).

broadening) the squared width becomes linear in the temperature (Fig. 3).

The extent to which a given mode of vibration  $q(\tau)$  of propagation constant  $\tau$  participates in the broadening of a level depends on the extent to which a perturbation potential  $q(\tau) \exp(i\tau \cdot r)$  shifts the electronic energy. To a first approximation, this shift may be computed by taking the mean of  $\exp(i\tau \cdot r)$  over the state in question. It is clear that phase cancellations will occur unless the wavelength of the phonon is comparable to, or larger than the Bohr radius of the state. More precisely, the phonons of importance will be those which obey

$$\tau \lesssim 2/a \approx 1.48 \times 10^7 \text{ cm}^{-1}, \quad (1.1)$$

when the Bohr radius  $a$  is given by:

$$a = (m/m_h) K a_H = (m/m_h) K (\hbar^2/m_e^2) \approx 13.5 \text{ \AA}. \quad (1.2)$$

In (1.2) we have used a dielectric constant  $K \approx 11.5$ . The effective mass for holes  $m_h/m \approx 0.45$  is determined by requiring that the hydrogenic formula yield agreement with the observed ground-state energy<sup>7</sup>  $E_g \approx 0.046 \pm 0.001$  eV. This "empirical" effective mass is intermediate between the two observed cyclotron resonance masses of the  $v_1$  and  $v_2$  bands (see Dexter and Lax).<sup>6</sup> It is hoped that the use of such an intermediate mass, and hence intermediate Bohr radius will partly cancel the error of neglecting the degeneracy in the bands.

The condition (1.1) applies to broadening of the  $1s$  ground state. For broadening of the  $2p$ ,  $3p$ , or  $4p$  states, we should insert a Bohr radius two, three, or four times as large. Since the number of phonons with propagation constant  $< 2/a$  is proportional to  $(2/a)^3$ , the broadening of the  $2p$ ,  $3p$ , and  $4p$  states should be related to the broadening in the  $1s$  state roughly by a factor  $1/8$ ,  $1/27$ , and  $1/64$  respectively. Thus, the broadening associated with a transition  $1s - np$  is produced almost entirely by the broadening of the  $1s$  level. For this reason, we expect that the  $1s - 2p$ ,  $1s - 3p$  and  $1s - 4p$  transitions show about the same width. This conclusion is in agreement with experiment (Fig. 2).

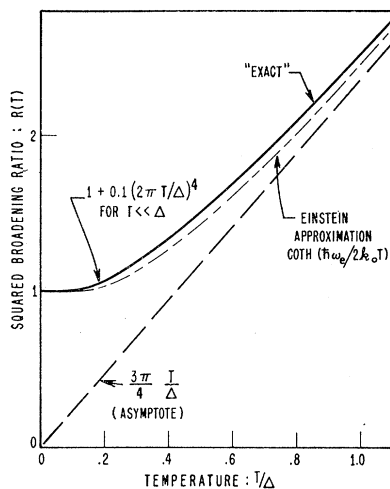


FIG. 3. The theoretical ratio,  $R(T)$ , of the squared broadening at the temperature  $T$  to its zero-point value, is plotted as a function of the dimensionless variable  $T/\Delta$ . The characteristic temperature  $\Delta = 2\hbar v/(k_0 a) \approx 94^\circ\text{K}$  in silicon, where  $v$  is the velocity of sound,  $a$  is the Bohr radius in the crystal and  $k_0$  is Boltzmann's constant.

The modes of importance in the broadening can according to (1.1) be assigned to a characteristic temperature  $\Delta$  according to

$$k_0 \Delta \approx \hbar \omega = \hbar v \tau \approx \hbar v (2/a), \quad (1.3)$$

where  $k_0$  is Boltzmann's constant, and  $v$  is the velocity of sound. For Silicon, we may use<sup>9</sup>  $v = (C_{11}/\rho)^{1/2} \approx 0.83 \times 10^6$  cm/sec. This leads to a characteristic temperature  $\Delta \approx 94^\circ\text{K}$ . The ratio  $R(T)$  of the mean squared broadening at temperature  $T$  to the zero-point broadening can then be approximated roughly<sup>10</sup> by the corresponding ratio of mean-squared oscillations of a harmonic oscillator of frequency  $\omega = k_0 \Delta/\hbar$ :

$$R(T) \approx \coth(\hbar \omega/2k_0 T) = \coth(\Delta/2T). \quad (1.4)$$

Thus, a noticeable increase in broadening should occur at a temperature  $T = \Delta$ . In fact,  $R(\Delta)$  is roughly 2, so that a 40 percent increase in the line width should be observed from absolute zero to  $90^\circ\text{K}$ . This is in good agreement with the actual increase in width between helium and nitrogen temperatures. The observed increase at low temperatures is strong evidence for the importance of long wavelength acoustic modes in the broadening; these temperatures are much smaller than the Debye temperature<sup>11</sup>  $\theta \approx 658^\circ\text{K}$ .

The good qualitative agreement between the experimental and theoretical temperature dependence of the broadening lends support to our proposed mechanism of broadening by the acoustical modes through the electron lattice interaction. A more stringent test would be a

comparison between the predicted zero-point broadening and the experimental broadening at liquid helium temperatures. Our detailed theory, to be presented later, shows that the rms width of the line at  $0^\circ\text{K}$  can be written in the form [see Eq. (5.15)]:

$$\begin{aligned} \langle (h\nu - h\bar{\nu})^2 \rangle^{1/2} &\approx (\hbar/3\pi^2 v \rho a^4)^{1/2} E_1 \\ &\approx 2.3 \times 10^{-4} E_1, \end{aligned} \quad (1.5)$$

where  $\rho \approx 2.33$  g/cm<sup>3</sup> is the (x-ray) density of silicon,  $v = 0.83 \times 10^6$  is the sound velocity, and  $a \approx 13.5 \times 10^{-8}$  cm is the Bohr radius.

The deformation potential constant<sup>12</sup>  $E_1$  measures the strength of the electron-lattice interaction and was chosen to be 15 eV to fit the experimental mobility at  $300^\circ\text{K}$  (see Sec. 4). Thus, the predicted rms width is about 3.6 millivolts, whereas the experimental full width at half-power is about 1 millivolt. (Experimental widths of donor states have a similar value.) For a Gaussian line shape, the full width at half-power is  $(8 \ln 2)^{1/2} \approx 2.35$  times bigger than the rms width. This would increase the predicted width to about 8.4 mV. However, an investigation of the theoretical line shape shows that the line shape is far from Gaussian due to the presence of "motional narrowing." When this is taken into account roughly, [see Eq. (6.2)] the predicted line width becomes 3.0 mV or about three times the experimental width.

The discrepancy between theory and experiment is sufficiently small, considering the difficulties of making an absolute calculation, that the proposed broadening mechanism seems quite reasonable. It nevertheless may be of some interest to examine the possible reasons for the discrepancy. Perhaps the first assumption to be re-examined is the use of the Born-Oppenheimer<sup>13</sup> approximation. We have assumed that the electronic wave function and energy adjusts to the instantaneous positions of the nuclei. If the nuclei move too fast, this will not be the case: the energy levels will be less sensitive to the nuclear motion and a smaller broadening will result. We know, therefore, that our predicted broadening is an overestimate.

Usually the validity of the Born-Oppenheimer approximation is guaranteed by the large mass ratio<sup>13</sup>: in a molecule the electrostatic restoring forces that act on an electron and on a nucleus are comparable in size. The frequencies of oscillation of the electron and nucleus will then be proportional to the inverse square root of the respective masses. However, in a semiconductor, the impurity center electron is not responsible for the cohesion of the lattice and there is no necessary relation between the forces acting on this electron and the nuclei. The nuclear vibrations are determined by the usual cohesive forces which are practically independent of the presence of the impurity electron. The electron, on the

<sup>9</sup> The velocity of sound varies with orientation in silicon. The longitudinal velocity in the (1,0,0) direction is taken as representative.

<sup>10</sup> More precise estimates of the broadening ratio given by Eqs. (5.22), (5.25) and Fig. 3 agree closely with (1.4).

<sup>11</sup> N. Pearlman and P. H. Keesom, Phys. Rev. **88**, 398 (1952).

<sup>12</sup> J. Bardeen and W. Shockley, Phys. Rev. **80**, 72 (1950).

<sup>13</sup> M. Born and J. R. Oppenheimer, Ann. Physik **84**, 457 (1927); M. Born and K. Huang, *Dynamics of Crystal Lattices* (Oxford University Press, Oxford, 1954).

other hand, is subject to forces weaker than usual because it moves in a large orbit, subject primarily to the electrostatic attraction of the center reduced by the dielectric constant. The validity of the Born-Oppenheimer approximation must therefore be reexamined in detail for the case of boron doped silicon:

The root mean square velocity of a hole trapped in the ground state is  $(2E_g/m_h)^{1/2} \simeq 1.9 \times 10^7$  cm/sec where the experimental ionization energy<sup>7</sup> is  $E_g = 0.046 \pm 0.001$  eV, and the effective mass  $m_h$  is chosen to be  $m_h = 0.45m$  [see Eq. (1.2), ff]. This may be compared with the velocity of sound in the material  $v \simeq 0.83 \times 10^6$  cm/sec.<sup>9</sup> The rms velocity of a typical nucleus is of the order of  $(k\theta/M_{\text{Si}})^{1/2} \simeq 10^5$  cm/sec, where  $\theta$  is the Debye temperature<sup>11</sup> 650°K and the mass of a silicon atom is  $M_{\text{Si}} = 4.65 \times 10^{-23}$  grams. Thus the electron moves considerably faster than the nuclei and will surely adjust somewhat to their motion.

However, the full validity of the Born-Oppenheimer approximation presumably requires the electron to traverse its orbit several times during one vibration of a nucleus. Thus, we should compare the frequencies of the two motions. The frequency of the electronic motion is of the order  $E_g/\hbar \simeq 7 \times 10^{13}$  cycles per second. The nuclear motion has a spectrum of frequencies whose maximum is of the order  $k\theta/\hbar \simeq 10^{14}$  cycles per second. If the broadening were predominantly produced by such high frequencies, the Born-Oppenheimer approximation would fail. However, as we remarked in Eq. (1.1) the phonons of importance have a propagation constant  $\tau \leq 2/a \simeq 1.48 \times 10^7$  cm<sup>-1</sup>. This result is only 0.13 of the maximum  $\tau$ :

$$\tau_{\text{max}} = (6\pi^2 n_0)^{1/3} \simeq 1.14 \times 10^8 \text{ cm}^{-1} \quad (1.6)$$

obtained by converting one cell of the reciprocal lattice space to a sphere of equal volume with radius  $\tau_{\text{max}}$ , (where  $n_0 = 2.5 \times 10^{22}$  primitive cells per cm<sup>3</sup>). The important vibrational frequencies will then be less than  $(0.13) \times 10^{14}$  cycles per second, so that the electron (whose frequency is  $0.7 \times 10^{14}$ ) will traverse its orbit five times before the important normal coordinates traverse theirs once. From this, we may conclude that the Born-Oppenheimer approximation will yield a broadening that is of the right order of magnitude.

The major source of error that remains is the use of a one-band picture that neglects the degeneracy of the valence bands at the center of the Brillouin zone. (A further, related assumption is that the band energy near  $k=0$  is approximately quadratic in  $k$  for energies of the order of  $k_0 T$ .) This error could be remedied by recalculating  $V_{aa}$  using (3.1) with a more accurate wavefunction  $\psi_a(r)$  for the impurity state. Such a change is not likely to reduce the theoretical answer by a factor as large as three. It seems much more likely that an error of a factor of three can be produced by attempting to determine the constant  $E_1$  using the standard one-band formula for the mobility Eq. (4.6). These points will be discussed further in Secs. 4 and 6.

## 2. MATHEMATICAL FORMULATION

If the Condon approximation is made,<sup>14</sup> the cross section for absorption of light by an impurity center can be written:

$$\sigma(\nu) = \left[ \frac{n \left( \frac{E_e}{E} \right)^2}{\epsilon \left( \frac{E_e}{E} \right)} \right] \frac{8\pi^3 \nu}{3c} |M_{ba}|^2 G(\nu), \quad (2.1)$$

where

$$G(\nu) = A v_m \sum_n \left| \int X_{bn}^*(x) X_{am}(x) dx \right|^2 \times \delta(E_{bn} - E_{am} - h\nu) \quad (2.2)$$

is the normalized absorption spectrum (15.2). Here  $X_{am}(x)$  is the  $m$ th vibrational state when the electron is in state  $a$  and  $X_{bn}(x)$  is the  $n$ th vibrational state when the electron is in state  $b$ , where  $x$  is an abbreviation for the  $10^{23}$  nuclear coordinates or normal coordinates of the lattice, and  $m, n$  are abbreviations for the quantum numbers associated with the  $10^{23}$  corresponding harmonic oscillators.

In the Born-Oppenheimer approximation, the effective Hamiltonian for nuclear vibration depends on the electronic state and will be denoted  $H_a$  or  $H_b$  for the ground or excited electronic states, respectively. Thus the eigenfunctions  $X_{am}(x)$  and  $X_{bn}(x)$  belong to two different Hamiltonians and are not orthogonal with respect to each other. As a consequence (2.2) permits multi-phonon transitions in which the individual normal modes of oscillation change their quantum numbers by zero or plus or minus one.

The difficulty in evaluating (2.2) lies in performing the sum over those final states  $n$  that satisfy the conservation of energy requirement. Huang and Rhys<sup>3</sup> were able to perform this sum by special devices that are applicable only when all the modes have the same frequency (e.g., an idealized optical branch). In I, we showed that the sum over final states could be performed by closure techniques after the energy conservation delta function was replaced by an integral representation and energy eigenvalues are replaced by the appropriate Hamiltonians  $H_a$  and  $H_b$ . Regardless of the nature of the vibration spectrum, we found

$$G(\nu) = h^{-1} \int_{-\infty}^{+\infty} \exp(-i2\pi\nu t) dt G(t), \quad (2.3)$$

where

$$G(t) = av [\exp(iH_b t/\hbar) \exp(-iH_a t/\hbar)] \quad (2.4)$$

still contains the average to be performed over initial states. The Boltzmann averaging symbol  $av$  acting on any operator  $O$  is an abbreviation for:

$$av(O) = \text{trace}[O \exp(-H_a/kT)] / \text{trace}[\exp(-H_a/kT)], \quad (2.5)$$

<sup>14</sup> See reference 1 for terminology and notation. For further discussion of (2.1) with reference to absolute absorption cross sections and broadening, see M. Lax, "The influence of lattice vibrations on electronic transitions in solids" in *Proceedings of the November 1954 ONR Conference on Photoconductivity* (John Wiley and Sons, Inc., New York, to be published).

so that (2.4) has been expressed in an invariant way in terms of the many-body Hamiltonians  $H_a$  and  $H_b$ .

A reduction of this many-body problem to a set of one-body problems can be performed if the normal coordinates  $q_j$  associated with  $H_a$  and  $H_b$  are the same, i.e., if  $H_a$  and  $H_b$  can be decomposed into terms depending on the same normal coordinates with no cross terms:

$$\begin{aligned} H_a &= \sum_j H_{aj}(q_j, p_j), \\ H_b &= \sum_j H_{bj}(q_j, p_j). \end{aligned} \quad (2.6)$$

In this case, the complete average (2.4) can be factored into a product of averages each pertaining to an individual normal mode:

$$G(t) = \prod_j G_j(t), \quad (2.7)$$

$$G_j(t) = \text{av}[\exp(iH_{bj}t/\hbar) \exp(-iH_{aj}t/\hbar)]. \quad (2.8)$$

This factorization was an essential simplification accomplished in I. The resulting averages (2.8) for individual harmonic oscillators are easy to perform by a wide variety of methods. An early derivation<sup>15</sup> was based on methods similar to Huang and Rhys. In I, ordered operator methods were used because of their elegance and simplicity. O'Rourke<sup>5</sup> pointed out the usefulness of Dirac density matrix methods, and included a quadratic term in the energy difference. Quadratic differences in  $H_b - H_a$  were not discussed in I because in general they will cause a mixing of the normal coordinates destroying the separability exhibited in (2.7). For the special case considered by O'Rourke<sup>4</sup> in which the quadratic energy difference contains only diagonal terms we can write:

$$\Delta E(q_j) = H_{bj} - H_{aj} = N^{-\frac{1}{2}} A_j q_j + N^{-1} B_j q_j^2. \quad (2.9)$$

Equations I(7.7), (7.10), and (8.6) then yield:

$$G_j(t) = \text{av} \exp_{[-]} \left[ i \int_0^t \Delta E(q_j(s)) ds / \hbar \right] \quad (2.10)$$

$$\begin{aligned} &= \exp \{ N^{-1} C_j^2 [i(\sin \omega_j t - \omega_j t) - (2\bar{n}_j + 1) \\ &\quad \times (1 - \cos \omega_j t)] + i N^{-1} (B_j / M \omega_j) (\bar{n}_j + \frac{1}{2}) t \}, \end{aligned} \quad (2.11)$$

with

$$C_j^2 = A_j^2 / (2\hbar M \omega_j^3). \quad (2.12)$$

To obtain  $G(t)$  from (2.11) one need only sum over  $j$  in the exponent. The resulting exponent then agrees (if  $B_j = 0$ ) with the previous calculation for a linear energy shift I(8.12). The effect of the quadratic energy difference is to add a term in the exponent linear in  $t$ . Such a linear term merely shifts the origin of  $\nu$  in (2.3), thus yielding the same frequency spectrum displaced by a constant amount.<sup>16</sup> The amount of the shift in energy units according to (2.3) and (2.11) is

$$\Delta(h\nu) = N^{-1} \sum_j (B_j \hbar / M \omega_j) (\bar{n}_j + \frac{1}{2}) \quad (2.13)$$

<sup>15</sup> M. Lax, Naval Research Laboratory Report 3973, June 17, 1952 (unpublished).

<sup>16</sup> If off-diagonal quadratic terms had been included, the shape of the spectrum would have changed as well. Further discussion of this point is given by Lax.<sup>14</sup>

or

$$\Delta(h\nu) = \text{av} [N^{-1} \sum_j B_j q_j^2], \quad (2.14)$$

i.e., the amount of the energy shift may be obtained by replacing the quadratic part of the energy difference by its thermal average value. The change of this shift with temperature was used by O'Rourke (following a suggestion of Huang and Rhys) to explain the shift in the  $F$ -center absorption peak with temperature.<sup>17</sup> The constants  $B_j$  were not calculated, but  $\sum B_j / N$  was chosen to give agreement with the experimental shift<sup>17</sup> (neglecting the fact that part of the shift may be ascribed simply to expansion of the lattice with increasing temperature).

The calculation made in (2.10) is easily extended to cubic and higher terms in the energy difference. However, a term of order  $N^{-\frac{3}{2}} q_j^3$  will yield no contribution of order  $N^{-1}$ , and thus will not contribute to the result. Does this mean that anharmonic terms have no influence on the absorption spectrum? Surely a cubic term  $x^3$  in the displacement of an impurity atom or a neighboring atom will affect the spectrum. Yet when  $x^3$  is expressed in terms of normal modes the terms are of order  $N^{-\frac{3}{2}} i q_j q_k$  that would be omitted in (2.9). This suggests that it may not be permissible to neglect cross terms among the quadratic terms. The diagonal quadratic terms will undoubtedly continue to yield an energy shift (2.14), but the cross terms will cause a mixing of modes. This mixing of modes may not change the peak or the width of the absorption spectrum appreciably, but it may scramble up or destroy fine structure that would otherwise be present.

In our calculation, we shall take account only of linear terms because there is no easy way to make a reliable estimate of the higher-order terms. The linear terms will be estimated from the electron-lattice interaction using the deformation potential approximation.<sup>12,18</sup> The electron-lattice interaction is weak in silicon and germanium—as shown by the high mobility. Thus the quadratic terms should be small. (As corroborating experimental evidence we note that the absorption line does not shift with temperature over a range in which appreciable broadening is observed.) We assume, therefore, that the influence of the quadratic terms on the first few moments of the spectrum are small—though they may have a significant effect on the higher moments, i.e., on details of the shape. We shall therefore use our calculation only to estimate the first few moments of the spectrum and not to make a detailed evaluation of the line shape.

### 3. EFFECTIVE VIBRATIONAL HAMILTONIANS

Let us suppose that the change in electronic potential associated with an arbitrary (small) displacement of the nuclei is  $V(\mathbf{r}, \dots)$ . The dots represent the dependence of

<sup>17</sup> Further discussion of the  $F$ -center problem is given by Lax.<sup>14</sup>

<sup>18</sup> An analysis of the validity of the deformation potential approximation is being made by R. Rosenberg as part of a thesis at Syracuse University on absorption of light by free carriers.

this potential on the displacements of the nuclei, or the normal coordinates of the latter's motion. To our present accuracy, the dependence of  $V$  on these displacements is linear. If the trapped electronic wave function in a ground or excited state is  $\psi_a(\mathbf{r})$  or  $\psi_b(\mathbf{r})$  for the case of a perfect lattice, the change in energy due to the lattice vibration is  $V_{aa}$  or  $V_{bb}$  correct to terms linear in the displacements. And a typical matrix element important in thermal transition problems is  $V_{ba}$ , where

$$V_{ba} = [\psi_b(\mathbf{r}), V(\mathbf{r}, \dots) \psi_a(\mathbf{r})] \\ \equiv \int \psi_b^*(\mathbf{r}) V(\mathbf{r}, \dots) \psi_a(\mathbf{r}) d\mathbf{r}. \quad (3.1)$$

However, neither the potential  $V$  nor the wave functions are known accurately unless a detailed knowledge of the band structure is available. We shall show nevertheless that an evaluation of (3.1) can be made without such information (for the nondegenerate case).

A wave function such as  $\psi_a(\mathbf{r})$  can be expanded<sup>19-21</sup> in terms of Bloch orbitals  $\psi(\mathbf{k}, \mathbf{r})$  or Wannier orbitals  $\alpha(\mathbf{r}-Q)$

$$\psi_a(\mathbf{r}) = \sum b_a(\mathbf{k}) \psi(\mathbf{k}, \mathbf{r}) = \sum c_a(Q) \alpha(\mathbf{r}-Q). \quad (3.2)$$

If there is no degeneracy between bands, and if the trap states are sufficiently spread out, it is sufficient to expand in terms of Bloch waves or Wannier orbitals of one band<sup>22,23</sup> the conduction band for electrons, or the valence band for holes. (Actually the valence band in silicon contains three bands degenerate at  $k=0$ , one of which is split off by spin-orbit interaction.<sup>18-20</sup> Since a detailed analysis of the impurity states is not yet available<sup>8</sup> under such circumstances, we shall be content here to do the calculation as if only one band were present. Part of the error in such a procedure will be cancelled by using an effective mass determined from the ionization energy of boron impurities.)

Although we do not have detailed information concerning the Wannier orbitals in (3.2), we know that  $c_a(Q)$  is a hydrogenic function of its discrete argument  $Q$ . We may regard  $b_a(k)$  as the same wave function in the momentum representation:

$$b_a(k) = N^{-\frac{1}{2}} \sum c_a(Q) \exp(-ik \cdot Q). \quad (3.3)$$

If (3.2) and a corresponding equation for  $\psi_b(\mathbf{r})$  are inserted into (3.1) we find:

$$V_{ba} = \sum_{kk'} b_b^*(k') b_a(k) V_{k'k}. \quad (3.4)$$

In principle, (3.4) constitutes the evaluation of (3.1) since the matrix elements  $V_{k'k}$  are the usual matrix elements of the theory of electrical conductivity. In-

serting (3.3) into (3.4), we obtain a form

$$V_{ba} = \sum_{Q'Q} c_b^*(Q') c_a(Q) V_{Q'Q} \quad (3.5)$$

expressed directly in terms of the hydrogenic functions  $c(Q)$ , with

$$V_{Q'Q} = N^{-1} \sum V_{k'k} \exp(ik'Q' - ikQ). \quad (3.6)$$

The momentum wave functions  $b(k)$  fall off rapidly for  $k$  larger than a reciprocal Bohr radius. Furthermore, this reciprocal Bohr radius is small compared to the size of the Brillouin zone—compare (2.3) and (2.4). Thus the dominant contribution to  $V_{ba}$  in (3.4) comes from  $V_{k'k}$  with sufficiently small  $k$  and  $k'$  that the deformation potential<sup>12</sup> is likely to be a good approximation.<sup>24</sup> In the latter approximation  $V_{k'k}$  is proportional to  $(\mathbf{k}' - \mathbf{k})$ . To effect a simplification of (3.5) and (3.6), however, all we need is that  $V_{k'k}$  be a function only of  $\mathbf{k}' - \mathbf{k}$ , a result that remains valid even in a self-consistent theory of the electron-lattice interaction.<sup>25</sup> If we write  $V_{k'k} = V(\Delta k, \dots)$  where  $\Delta k = k' - k$  and the dots are as usual the displacement or normal coordinates, then (3.5) becomes:

$$V_{ba} = \sum_{\tau} V(\tau, \dots) N_{ba}(\tau), \quad (3.7)$$

where

$$N_{ba}(\tau) = \sum_Q c_b^*(Q) \exp[i\tau \cdot Q] c_a(Q). \quad (3.8)$$

Here  $\tau$  is used as a summation index, replacing  $\Delta k$ . From (4.4), it can be seen to have the significance of a phonon propagation vector. The coefficients  $c(Q)$  are normalized in the discrete sense:

$$\sum_Q |c_a(Q)|^2 = 1. \quad (3.9)$$

Thus,  $N_{aa}(\tau)$  is the discretized mean value of  $\exp[i\tau \cdot Q]$  in the hydrogenic ground state. Since the ground state  $c_a(Q)$  varies slowly as  $Q$  moves from one lattice point to another, we may replace the sum (3.8) by the corresponding normalized integral:

$$N_{aa}(\tau) = \int \exp(i\tau \cdot \mathbf{r}) (\pi a^3)^{-1} \exp(-2r/a) d\mathbf{r} \\ = [1 + (\frac{1}{2}a\tau)^2]^{-2}. \quad (3.10)$$

For the first excited state, the  $2p$  hydrogenic state, we have

$$N_{bb}(\tau) = \int \exp(i\tau \cdot \mathbf{r}) (32\pi a^5)^{-1} [r \cos(r, e)]^2 \exp(-r/a) d\mathbf{r} \\ = [1 + (a\tau)^2]^{-4} [1 - (a\tau)^2 - 4(a\tau)P_2(\cos(\tau, e))],$$

when we have chosen a state with  $l=1, m=0$  with our  $z$  axis in the direction of the polarization vector  $e$  of the light quantum. This is the state that will absorb light polarized in the  $z$ -direction.

<sup>19</sup> G. H. Wannier, Phys. Rev. **52**, 191 (1937).  
<sup>20</sup> J. C. Slater, Phys. Rev. **76**, 1592 (1949); G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954).

<sup>21</sup> H. M. James, Phys. Rev. **76**, 1602, 1611 (1949).

<sup>22</sup> E. N. Adams, Phys. Rev. **85**, 41 (1952), J. Chem. Phys. **21**, 2013 (1953).

<sup>23</sup> P. Feuer, Phys. Rev. **88**, 92 (1952).

<sup>24</sup> Strictly speaking, one should use a generalization of the deformation potential to include scattering by transverse modes. See, however, reference 26. For degenerate bands one should also include the effect of interband transitions.

<sup>25</sup> J. Bardeen, Phys. Rev. **52**, 688 (1937).

## 4. DEFORMATION POTENTIAL APPROXIMATION

A comparison of (3.7) and (3.10) shows that the evaluation of  $V_{aa}$  requires an accurate knowledge of  $V(\tau, \dots)$  only for  $\tau < (2/a)$ , i.e., we are concerned with phonons of sufficiently long wavelength compared to the lattice spacing [compare (1.1) and (1.6)] that the deformation potential approximation should be adequate.<sup>24</sup> We shall in this paper neglect entirely the contribution of the optical modes and assume that for acoustical modes the two particles in the primitive cell of the diamond structure move in unison. (An analysis of the errors involved in the deformation potential approximation is in progress.<sup>18</sup>) Thus we can simplify our notation by proceeding as if we were dealing with a lattice containing one atom per cell whose mass  $M$  is twice that of a silicon atom  $M_{Si}$ . An arbitrary deformation  $\delta\mathbf{R}(\mathbf{r}_j)$  can then be expressed in terms of the complex normal coordinates  $q_\alpha(\tau)$  by

$$\delta\mathbf{R}(\mathbf{r}_j) = N^{-\frac{1}{2}} \sum_{\alpha=1}^3 \mathbf{e}_\alpha(\boldsymbol{\tau}) q_\alpha(\boldsymbol{\tau}) e^{i\boldsymbol{\tau} \cdot \mathbf{r}_j}, \quad (4.1)$$

where  $\mathbf{e}_\alpha(\boldsymbol{\tau})$  are the unit polarization vectors, and for sufficiently small  $\tau$  we may adopt the convention that  $\mathbf{e}_3(\boldsymbol{\tau}) = \boldsymbol{\tau}/\tau$  is the longitudinal mode. The deformation potential is then given by<sup>26</sup>

$$V(\mathbf{r}, \dots) = E_1 \text{div} \delta\mathbf{R}(\mathbf{r}) = iE_1 N^{-\frac{1}{2}} \sum_{\alpha} \tau q_\alpha(\boldsymbol{\tau}) e^{i\boldsymbol{\tau} \cdot \mathbf{r}} \quad (4.2)$$

and the standard matrix element is given by

$$\begin{aligned} V(\Delta k, \dots) &= [\psi(\mathbf{k}', \mathbf{r}), V(\mathbf{r}, \dots) \psi(\mathbf{k}, \mathbf{r})] \\ &\simeq iE_1 N^{-\frac{1}{2}} \Delta k q_3(\Delta \mathbf{k}), \end{aligned} \quad (4.3)$$

which depends on the approximation

$$[\psi(\mathbf{k}', \mathbf{r}), e^{i\boldsymbol{\tau} \cdot \mathbf{r}} \psi(\mathbf{k}, \mathbf{r})] \simeq \delta_{\Delta k \boldsymbol{\tau}}, \quad (4.4)$$

whose validity follows from the smallness of  $\Delta k$ .

Since only longitudinal modes are involved hereafter, we shall drop the subscript 3 on  $q_3$ . Furthermore,  $q(\boldsymbol{\tau})$  can be written

$$q(\boldsymbol{\tau}) = [\hbar/2M\omega(\boldsymbol{\tau})]^\frac{1}{2} [a(\boldsymbol{\tau}) - a^\dagger(\boldsymbol{\tau})] \quad (4.5)$$

where  $a(\boldsymbol{\tau})$  and  $a^\dagger(\boldsymbol{\tau})$  are conventional destruction and creation operators for phonons of propagation constant  $\boldsymbol{\tau}$  with matrix elements  $[n(\boldsymbol{\tau})]^\frac{1}{2}$  and  $[n(\boldsymbol{\tau})+1]^\frac{1}{2}$  respectively.

The matrix elements (4.3) and (4.5) lead to the customary formula for the mobility,<sup>12</sup>

$$\mu = \left(\frac{8\pi}{9}\right)^\frac{1}{2} \frac{e\hbar^4 n_0 M v_s^2}{E_1^2 (kT)^\frac{3}{2} (m^*)^{5/2}}, \quad (4.6)$$

for the case of spherical energy surfaces.<sup>27</sup>

<sup>23</sup> See reference 12, Eq. (A30). The neglect of the transverse contribution is not serious here, as pointed out by Bardeen and Shockley. The contribution of transverse modes can be included roughly, providing we understand that  $E_1^2$  now represents the sum of the squares of three deformation constants. Since  $E_1^2$  will later be chosen by comparison with experiment, transverse contributions are automatically included. We shall, therefore, use the

An estimate of  $E_1$  for holes in silicon can be obtained from (4.6). Using  $\mu_p \simeq 500$  cm<sup>2</sup>/volt-sec<sup>28</sup> at 300°K, a velocity of sound  $v \simeq 0.83 \times 10^6$  cm/sec, an effective mass  $m^* \simeq 0.45m$ , and  $M = 2M_{Si}$ , we find that  $E_1 \simeq 15$  ev. Our estimate is based on a high-temperature mobility because the broadening mechanism automatically selects lattice scattering in preference to impurity or dislocation scattering: The trapped hole will only be scattered by vibrations or imperfections in its immediate vicinity. Imperfections sufficiently close to the impurity atom to be of importance will very likely shift the entire line by more than its width in a somewhat random way and thus will not appear with any appreciable intensity in the observed absorption spectrum.

Thus, it is necessary to determine separately the effect of lattice scattering, which is partially accomplished by using a room temperature mobility. However, the experimental mobility<sup>28</sup>  $\mu_p \simeq 2.4 \times 10^8 T^{-2.3}$  disagrees with the  $T^{-1.5}$  power predicted by the deformation theory. If, for example, we were to fit our constant  $E_1$  at 100°K instead of 300°K, we would find  $E_1 \simeq 10$  ev. This leads to a reduction of a factor of 2 in the width and leads to much better agreement with experiment. We shall prefer, however, for the sake of conservatism to use the room temperature fit, and regard the factor of 2 as an uncertainty in the predicted broadening because of the absence of an adequate mobility theory.

## 5. CHARACTERISTIC FUNCTION AND THE MOMENTS

Combining (3.7) and (4.3), the difference between the two Hamiltonians  $H_b - H_a$  (aside from a constant term) is given by

$$V_{bb} - V_{aa} = N^{-\frac{1}{2}} \sum A(\boldsymbol{\tau}) q(\boldsymbol{\tau}), \quad (5.1)$$

where

$$A(\boldsymbol{\tau}) = iE_1 \tau N(\boldsymbol{\tau}); \quad N(\boldsymbol{\tau}) = N_{bb}(\boldsymbol{\tau}) - N_{aa}(\boldsymbol{\tau}). \quad (5.2)$$

The characteristic function  $G(t)$  can now be obtained from (2.11) or (I8.12) providing we merely replace  $A_j^2$  by  $|A(\boldsymbol{\tau})|^2$  since the normal coordinates  $q(\boldsymbol{\tau})$  are complex (see Appendix 1).

$$G(t) = \exp[i2\pi\nu_0 t + f(t)], \quad (5.3)$$

$$\begin{aligned} f(t) &= \langle |C(\boldsymbol{\tau})|^2 [i \sin \omega_\tau t \\ &\quad - (2\bar{n}_\tau + 1)(1 - \cos \omega_\tau t)] \rangle, \end{aligned} \quad (5.4)$$

$$|C(\boldsymbol{\tau})|^2 = E_1^2 \tau^2 |N(\boldsymbol{\tau})|^2 / (2\hbar M \omega_\tau^3). \quad (5.5)$$

The average,  $\langle \rangle$  taken in (5.4) is defined as in (I5.21) to be an average over the modes  $\tau$ , i.e.,  $N^{-1}$  times the sum

original deformation potential thereby avoiding complicated notation. The most serious error made by this procedure is that we do not explicitly take account of interband transitions.

<sup>27</sup> For ellipsoidal surfaces centered at  $k=0$  with effective masses  $m_1, m_2, m_3$ , Eq. (4.6) is still valid providing  $m^*$  is defined by:

$$(m^*)^{5/2} = 3(m_1 m_2 m_3)^\frac{1}{2} / (m_1^{-1} + m_2^{-1} + m_3^{-1}).$$

For an eight-ellipsoid picture such as prevails in germanium,<sup>6</sup> (4.6) must be reduced by the presence of interellipsoid transitions. The latter effect can be shown to be small in germanium however.<sup>18</sup>

<sup>28</sup> M. B. Prince, Phys. Rev. **93**, 1204 (1954).

of the contribution of the  $N$  different modes  $\tau$ . The energy  $h\nu_0$  constitutes the thermal activation energy,<sup>1</sup> and will be defined more precisely in (5.40).

The moments of a distribution function can be obtained directly from the derivatives of the corresponding characteristic function at the origin, see I (4.2), I (8.14). Explicit formulas for the moments are identical to I (8.16)

$$h\bar{\nu} = h\nu_0 \pm \langle |C(\tau)|^2 \hbar\omega_\tau \rangle, \quad (5.6)$$

$$\langle (h\nu - h\bar{\nu})^2 \rangle = \langle |C(\tau)|^2 (2\bar{n}_\tau + 1) (\hbar\omega_\tau)^2 \rangle, \quad (5.7)$$

$$\langle (h\nu - h\bar{\nu})^3 \rangle = \pm \langle |C(\tau)|^2 (\hbar\omega_\tau)^3 \rangle, \quad (5.8)$$

$$\langle (h\nu - h\bar{\nu})^4 \rangle = 3 \langle (h\nu - h\bar{\nu})^2 \rangle^2 + \langle |C(\tau)|^2 (2\bar{n}_\tau + 1) (\hbar\omega_\tau)^4 \rangle. \quad (5.9)$$

The upper signs refer to absorption, the lower to emission. The first moment or mean absorption energy  $h\bar{\nu}$  will be discussed later in this section.

The second moment (5.7) is a measure of the breadth of the absorption line in the transition  $a \rightarrow b$ . The breadths of the individual levels may be obtained by using  $N_{aa}$  or  $N_{bb}$  in place of  $N(\tau)$  in (5.5). The squared breadth associated with the transition, however, may differ from the sum of breadths of the individual levels through a cross or correlation term:

$$\text{av}(V_{bb} - V_{aa})^2 = \text{av}V_{aa}^2 + \text{av}V_{bb}^2 - 2 \text{av}(V_{aa}V_{bb}). \quad (5.10)$$

The significance of (5.10) may be visualized as follows: during a lattice displacement the energy of level  $a$  changes by an amount  $V_{aa}$ . Thus,  $\text{av}V_{aa}^2$  is the mean-squared excursion of width of state  $a$ . Similar remarks apply to state  $b$ . However, the width of an absorption line is determined by  $V_{bb} - V_{aa}$ , the shift of line  $b$  relative to  $a$ , so that (5.10) automatically takes account any correlation in the shifts of the two lines.

We shall now calculate the broadening of the ground hydrogen-like state, by inserting (3.10) and (5.5) into (5.7):

$$\langle (h\nu - h\bar{\nu})^2 \rangle_{aa} = \frac{\hbar E_1^2}{Mv_s a} \left\langle \left[ \frac{x}{(1+x^2)^4} (1+2\bar{n}_x) \right] \right\rangle, \quad (5.11)$$

where

$$x = \frac{1}{2}a\tau, \quad \bar{x} = \frac{1}{2}a\tau_{\text{max}} \simeq 7.7, \quad (5.12)$$

$$\langle [ \ ] \rangle = (3/\bar{x}^3) \int_0^{\bar{x}} [ \ ] dx \quad (5.13)$$

and

$$\bar{n}_x = [\exp(x\Delta/T) - 1], \quad (5.14)$$

with

$$\Delta = (2\hbar\nu/k_0a) \simeq 94^\circ\text{K}.$$

The zero-point broadening, obtained by setting  $\bar{n}_x = 0$  in (5.11) may be calculated readily by noting that

$$\int_0^\infty x^3(1+x^2)^{-4} dx = (1/12); \quad [\hbar/(Mva)] \simeq 10^{-4}, \quad (5.4)$$

so that

$$\begin{aligned} \langle (h\nu - h\bar{\nu})^2 \rangle_{aa} &\simeq \left( \frac{\hbar}{4Mva\bar{x}^3} \right)^{\frac{1}{2}} E_1 \\ &\simeq 2.3 \times 10^{-4} E_1 \simeq 3.6 \times 10^{-3} \text{ ev}. \end{aligned} \quad (5.15)$$

The ratio of the broadening in the  $2p$  state to that in the  $1s$  state is given by

$$\frac{\langle V_{bb}^2 \rangle}{\langle V_{aa}^2 \rangle} = \frac{\langle x(1+4x^2)^{-8} [1-4x^2-8xP_2(\cos(x,e))]^2 \rangle}{\langle x(1+x^2)^{-4} \rangle} \quad (5.16)$$

$$= \frac{\langle x(1+4x^2)^{-8} [1+(24/5)x^2+16x^4] \rangle}{\langle x(1+x^2)^{-4} \rangle}. \quad (5.17)$$

$$\frac{\langle V_{bb}^2 \rangle}{\langle V_{aa}^2 \rangle} = 89/(5600) \simeq 0.016, \quad (5.18)$$

where the step from (5.15) to (5.16) involved an average over the orientation of the light polarization vector  $\mathbf{e}$  relative to the propagation direction.

We may conclude from (5.18) that the broadening of the  $2p$  state is only one-eighth that of the ground state. The physical reason for this is that the Bohr radius of a  $2p$  state is twice that of a  $1s$  state and the number of phonons with sufficiently large wavelength to produce broadening is correspondingly reduced. The broadening of the  $3p$  and  $4p$  states will be even less. The entire width of the  $1s-2p$ ,  $1s-3p$  and  $1s-4p$  transitions can therefore be ascribed to the ground state. (There is, of course, also broadening of the  $1s$ -continuum transition, but this can only be observed as a rounding of the edge).

The ratio of the second moment at temperature  $T$  to its zero-point value is given by

$$R(T) = 1 + 24 \int_0^\infty x^3(1+x^2)^{-4} [\exp(x\Delta/T) - 1]^{-1} dx \quad (5.19)$$

for the ground state. At sufficiently high temperatures, the second moment becomes linear in  $T$ :

$$R(T) \xrightarrow{T \rightarrow \infty} (3\pi/4)(T/\Delta), \quad (5.20)$$

whereas at low temperatures it approaches its zero-point value as  $T^4$ :

$$R(T) \xrightarrow{T \rightarrow 0} 1 + (8\pi^4/5)(T/\Delta)^4. \quad (5.21)$$

An approximate expression for  $R(T)$  may be obtained by replacing the distribution of modes by a single effective frequency. In this case

$$R(T) \approx \coth(\frac{1}{2}\hbar\omega_e/k_0T), \quad (5.22)$$

where the best choice for the effective frequency,<sup>14</sup> using Eq. I (5.22) and I (5.23), is

$$\hbar\omega_e = (8/3\pi)k_0\Delta. \quad (5.24)$$



The Einstein approximation (5.22) approaches unity exponentially at low temperatures in contrast to the power law behavior of (5.21) which follows from the more accurate Debye treatment.

The integral (5.19) can be carried out exactly using contour integration methods with the result:

$$R(T) = (1/z) \left[ (3/8) + (3/4)z^2\psi'(z) - (3/4)z^3\psi''(z) - (1/4)z^4\psi'''(z) \right], \quad (5.25)$$

where

$$z = \Delta / (2\pi T) \quad (5.26)$$

and

$$\psi(z) = d/dz \ln \Gamma(1+z); \quad \psi'(z) = \sum_{\nu=1}^{\infty} \frac{1}{(z+\nu)^2}. \quad (5.27)$$

The digamma, trigamma, tetragamma, and pentagamma functions  $\psi(z)$ ,  $\psi'(z)$ ,  $\psi''(z)$  and  $\psi'''(z)$  are tabulated by Davis.<sup>29</sup>

A comparison of the "exact" broadening (5.26) and the "Einstein" broadening *versus* temperature is shown in Fig. 3.

Equations (5.19)–(5.21) describing the change in broadening with temperature apply to trapped electrons as well as holes. However, it must be remembered that the temperature  $\Delta$  in (5.13) varies inversely with the Bohr radius  $a$ ;  $\Delta_e/\Delta_h = a_h/a_e = m_e/m_h$  where the last equality applies if both states are hydrogenlike.

The ratio of the zero-point broadening for electrons to that for holes according to (5.15) and (4.6) is given by:

$$\frac{\langle (h\nu - h\bar{\nu})^2 \rangle_e}{\langle (h\nu - h\bar{\nu})^2 \rangle_h} \xrightarrow{T \rightarrow 0} \left( \frac{E_e m_e^2}{E_h m_h^2} \right)^2 = \left( \frac{m_e}{m_h} \right)^2 \frac{\mu_h}{\mu_e}. \quad (5.28)$$

The third moment according to (5.8) is independent of temperature and can be written in the form:

$$\begin{aligned} \langle (h\nu - h\bar{\nu})^3 \rangle &\simeq \pm \frac{E_1^2 \hbar^2}{M a^2 \bar{x}^3} \int_0^{\infty} \frac{x^4}{(1+x^2)^4} dx \\ &\simeq \frac{3\pi}{16} \frac{E_1^2 \hbar^2}{M a^2 \bar{x}^3}. \end{aligned} \quad (5.29)$$

The contribution of the second term in (5.9) to the fourth moment can be written

$$\langle |C(\tau)|^2 (2\bar{n}_\tau + 1) (\hbar\omega_\tau)^4 \rangle \equiv \frac{2E_1^2 \hbar^3 v_s}{M a^3 \bar{x}^3} R_4(T), \quad (5.30)$$

where

$$R_4(T) \simeq 6 \int_0^{\infty} \frac{x^5}{(1+x^2)^4} \coth\left(\frac{x\Delta}{2T}\right) dx \quad (5.31)$$

reduces to unity at absolute zero. In fact at sufficiently

low temperatures ( $T \ll \Delta$ )

$$R_4(T) \rightarrow 1 + 1440\zeta(5)(T/\Delta)^6, \quad (5.32)$$

where  $\zeta(5) = 1.0369$ , and at sufficiently high temperatures ( $T \gg \Delta$ )

$$R_4(T) \rightarrow 3\pi T / (8\Delta) \quad (5.33)$$

whereas at intermediate temperatures  $T \sim \Delta$ , or higher, a sufficiently accurate expression can be obtained by a mean value estimate of (5.31):

$$R_4(T) \simeq \coth[8\Delta / (3\pi T)]. \quad (5.34)$$

A dimensionless description of the shape of the absorption curve can be obtained from the skewness and kurtosis:

$$\gamma_3(T) = \frac{\langle (h\nu - h\bar{\nu})^3 \rangle}{[\langle (h\nu - h\bar{\nu})^2 \rangle]^{3/2}} \quad (5.35)$$

$$= -\frac{3\pi}{2} (\bar{x})^3 \frac{M v_s^2}{E_1} \left( \frac{\hbar}{M v_s a} \right)^{1/2} [R(T)]^{-3/2}.$$

$$\gamma_4(T) = \frac{\langle (h\nu - h\bar{\nu})^4 \rangle}{[\langle (h\nu - h\bar{\nu})^2 \rangle]^2} - 3 \quad (5.36)$$

$$= 32 (\bar{x})^3 \left( \frac{M v_s^2}{E_1} \right)^2 \frac{\hbar}{M v_s a} \frac{R_4(T)}{[R(T)]^2}.$$

$$\gamma_3(0) \simeq 1.8; \quad \gamma_4(0) \simeq 8.9. \quad (5.37)$$

From (5.37) we see that the skewness and kurtosis are quite appreciable near absolute zero. However, they decrease at higher temperatures. Since  $\gamma_3(T)$  semiclassical = 0, and  $\gamma_4(T)$  semiclassical = 0, we see that the semiclassical approximation is not valid at low or even moderate temperatures.

It may be of interest to note that

$$\langle |C(\tau)|^2 \rangle = E_1^2 a / (8\hbar M v_s^3 \bar{x}^3) \quad (5.38)$$

$$\simeq E_1^2 / 590 \simeq 0.38, \quad (5.39)$$

where  $E_1$  has been set equal to 15 eV in (5.39). The smallness of this parameter indicates the feasibility of expanding  $\exp[f(t)]$  in powers of  $f(t)$  in (5.3), i.e., of expressing the results in terms of zero-phonon, one-phonon, two-phonon contributions, etc. (The parameter  $\langle |C(\tau)|^2 \rangle$  corresponds to the  $S$  of Huang and Rhys and indicates the average number of phonons to be expected in a transition.) This point will be discussed again in Sec. 6.

In I, we showed that because of the interaction between the electron and the lattice, the lattice relaxes to a new set of equilibrium positions one for state  $a$ , and one for state  $b$ . These positions are the minima of the parabolas that give the configurational energy of the system (see Fig. 1 of I). The energy difference between these minima denoted  $h\nu_0$ , corresponds to the thermal activation energy between the states  $a$  and  $b$ . We

<sup>29</sup>H. T. Davis, *Tables of the Higher Mathematical Functions* (Principia Press, Bloomington, 1933), Vols. I and II.

showed in (I 6.7) that

$$h\nu_0 = E_b - E_a + \langle [(A_j^a)^2 - (A_j^b)^2] / (2M\omega_j^2) \rangle, \quad (5.40)$$

where  $E_b - E_a$  is the electronic energy difference with the nuclei constrained to the perfect lattice positions. The constants  $A_j^a$  and  $A_j^b$  are the linear interaction constants appropriate to state  $a$  or  $b$  respectively [see (5.2)]. For our purposes we have shown that the interaction in the  $2p$  state is negligible compared to that in the  $1s$  state, so that we may set  $A_j^b = 0$ ,  $A_j^a = A$ . For the present case, then,

$$h\nu_0 = E_b - E_a + \langle |C(\tau)|^2 \hbar\omega \rangle \quad (5.41)$$

and

$$h\bar{\nu}_a = E_b - E_a + 2\langle |C(\tau)|^2 \hbar\omega \rangle \quad (5.42)$$

follows from (5.6). Thus the absorption energy  $h\bar{\nu}_a$  will shift from the constrained lattice value. An estimate of this shift may be obtained by noting that

$$\begin{aligned} \langle \hbar\omega |C(\tau)|^2 \rangle &= \frac{3\pi \langle (h\nu - h\bar{\nu})^2 \rangle_{T=0}}{8 k_0 \Delta} \\ &\simeq \frac{3\pi (3.6 \text{ mv})^2}{8 \cdot 8.1 \text{ mv}} \simeq 1.9 \text{ mv}. \end{aligned} \quad (5.43)$$

Thus a total shift of 3.8 mv could occur. Since our theoretical breadth is too large by a factor three, the shift is more likely about 1.3 mv. This shift cannot, of course, be observed directly because it is temperature independent. It could affect slightly, a comparison between theoretical and experimental estimates of impurity level energies.

## 6. COMPARISON WITH EXPERIMENT

There are two qualitative experimental results that bear emphasis and explanation: Only a finite number (three) of transitions has been observed, and the width of the absorption line is approximately the same in each case. The second result we have already explained by noting that the broadening in the three transitions  $1s-2p$ ,  $1s-3p$ ,  $1s-4p$  is primarily the broadening of the  $1s$  level. The  $2p-3p-4p$  levels have a much smaller broadening because their Bohr radii are larger, and the number of lattice modes with sufficiently long wavelength to produce broadening is correspondingly reduced.

It has been previously suggested that only a finite number of excited states exist because the wave functions associated with different impurity centers will eventually begin to overlap.<sup>30</sup> This is a limiting factor only at very high impurity concentrations, however. The crystal for which the measurements in Figs. 1 and 2 were taken (referred to as Si 155 in reference 7) is a fairly pure specimen having about  $5 \times 10^{15}$  free holes at room temperature. Thus the separation between boron centers is of the order of 1000 Å. This is enough larger

than the first Bohr radius 13.5 Å that the centers may be regarded as independent.

A more elementary explanation can be given for the fact that only a finite number of transitions is observed—namely that such transitions would not be resolved since (a) they are close together in energy: they must all occur between 0.043 eV and 0.046 eV; see Fig. 2; (b) they should have widths of the order of 0.001 eV equal to those for the  $1s-2p$ ,  $1s-3p$ , or  $1s-4p$  transition; (c) they should have low intensities.

Another striking experimental result is the appreciable increase in width (about 40 percent) between liquid helium and liquid nitrogen temperatures. This is evidence that the phonons of importance have low energies  $\hbar\omega \simeq k_0(80)$ . Our analysis agrees with this by showing that the phonons of importance have wavelengths about equal to  $\pi$  times a Bohr radius. These phonons have a characteristic temperature of  $\Delta = 2\hbar\nu/k_0 a \simeq 94^\circ\text{K}$  in silicon, see (5.14). The squared broadening ratio is plotted *versus*  $T/\Delta$  in Fig. 3. At  $T = \Delta$ , the squared broadening ratio is slightly over 2 in qualitative agreement with the experimental 40 percent increase.

Although the qualitative agreement with experiment suggests that our picture of the broadening as due to interaction with long wavelength acoustic modes is correct, our absolute estimate of the line width is slightly too large. The experimental full width at half power (F.W. at H.P.) is about one millivolt. The rms width [Eq. (5.15)] is about 3.6 millivolts. To convert from the rms width to the full width at half power, we need to know the shape of the line. For a Gaussian shape the

$$\begin{aligned} \text{F.W. at H.P.} &= (8 \ln 2)^{1/2} [\langle (h\nu - h\bar{\nu})^2 \rangle]^{1/2} \\ &\simeq (2.35)(3.6) \simeq 8.4 \text{ mv} \end{aligned} \quad (6.1)$$

at absolute zero (or liquid helium temperatures).

However, the line shape at absolute zero may be expected to be far from Gaussian as an examination of the kurtosis (5.36) shows. A large kurtosis implies a large fourth moment for the same second moment—so that more of the strength of the line is in the wings—with a correspondingly narrower, more peaked center. This “motional narrowing” of the line is a quantum mechanical effect that would not have appeared in a semiclassical treatment and results from the weakness of the electron-lattice interaction in silicon. A rough estimate of the importance of the line shape on the width may be obtained by comparing with the case of exchange narrowing in paramagnetic resonance treated by Anderson and Weiss.<sup>31</sup> Their estimate of the F.W. at H.P.  $\simeq 2\omega_p^2/\omega_e$  can be rewritten in the form

$$\begin{aligned} \text{F.W. at H.P.} &\simeq (2\pi/\gamma_4)^{1/2} [\langle (h\nu - h\bar{\nu})^2 \rangle]^{1/2} \\ &\simeq (0.84)(3.6) \simeq 3.0 \text{ mv} \end{aligned} \quad (6.2)$$

by using the fourth moment to estimate  $\omega_e$ . Because of

<sup>30</sup> K. S. Shifrin, J. Tech. Physics (U.S.S.R.) 14, 43 (1944).

<sup>31</sup> P. W. Anderson and P. R. Weiss, Revs. Modern Phys. 25, 269 (1953).

the large kurtosis, an appreciable narrowing results, and the estimated width of 3 millivolts is only about three times too large.

One of the possible reasons that the theoretical answer is too large is that the calculation was made using the Born-Oppenheimer approximation which inevitably overestimates somewhat the sensitivity of the electronic energy to the nuclear coordinates. This point was discussed in detail in Sec. 1. Another reason for error in our calculation of the electronic energies was the use of a simplified one band picture rather than the use of a degenerate band picture for the valence band in silicon. This error should be partly cancelled by using an effective mass determined from the experimental binding energy. Another approximation was the use of the deformation potential which neglects any contribution of transverse modes to the scattering matrix elements. This should be partly cancelled by the use of the experimental mobility and the deformation potential theory in the determination of the electron-lattice constant. Perhaps the worst error is that constant  $E_1$  is estimated neglecting the possibility of hole scatterings between the degenerate bands and quartic terms in the energy-momentum relationship. Because of this neglect, we have probably overestimated  $E_1$  and hence overestimated our theoretical width. (If we had used the 100°K mobility to determine  $E_1$ , the predicted broadening would have been reduced by a factor of  $3.8 \approx 2.35$ ).

As a concluding remark, we would like to explain why we used the Anderson-Weiss narrowing estimate (6.2) rather than working directly from the line shape which is presumably known from (2.3) and (5.3). Because of the smallness of the parameter  $E_1$  in silicon ( $|C(\tau)|^2 \approx (E_1/24)^2 \approx 0.4$  and it becomes legitimate to expand (5.3) in powers of  $f(t)$ . The term of order  $[f(t)]^n$  then corresponds to an  $n$ -phonon process. The zeroth-order term in this expansion corresponds to a delta function, i.e., an infinitely sharp line at the electronic energy difference. The one- and two-phonon processes, etc., yield a continuous background. Because the integrated intensity is independent of  $E_1$ , a large  $E_1$  (the semi-classical limit) would throw most of the intensity into the many phonon processes with a corresponding reduction in the intensity of the center (zero-phonon) line. It is then more legitimate to approximate the line shape by a continuous curve. For small  $E_1$ , as is the case in silicon, a fair fraction of the total intensity resides in the center line and no simple description of the line shape can be given. Of course, an infinitely sharp center line could never be observed because of lack of perfect resolution in the detection instruments. The fact that our line

width increases from helium to nitrogen temperatures, however, shows that the observed broadening cannot be entirely instrumental.

One possible theoretical explanation is that it is not possible to neglect the terms in the energy difference of second order in the nuclear displacements even when such terms are small. For such terms imply that the normal coordinates of the lattice are different in the ground and excited electronic states. There is then, strictly speaking, no such thing as a zero-phonon process. The presence of a small second-order term could then conceivably result in a continuous line shape without appreciably affecting the first few moments of the spectral distribution. Thus, we have assumed that our calculated values of the first few moments are more reliable than our calculated line shape and have used them to make a crude estimate of the line width.

#### APPENDIX: INTRODUCTION OF REAL COORDINATES

The complex linear energy difference (5.1) can be rewritten in the symmetrized form:

$$\Delta E(q) = N^{-1/2} \sum_{\tau > 0} [A(\tau)q(\tau) + A(-\tau)q(-\tau)], \quad (\text{A1})$$

where we ignore the term  $\tau=0$  since  $A(0)=0$ , and in any case, one term will make no difference after the limit  $N \rightarrow \infty$  is taken. Since  $\Delta E$  is a Hermitian operator and  $q(-\tau) = \pm q^+(\tau)$  we must have also  $A(-\tau) = \pm A^*(\tau)$ . If we introduce these relations into (A1) and define

$$\begin{aligned} A(\tau) &= [A_1(\tau) + iA_2(\tau)]/\sqrt{2}, \\ q(\tau) &= [q_1(\tau) + iq_2(\tau)]/\sqrt{2}, \end{aligned} \quad (\text{A2})$$

then

$$\Delta E(q) = \sum_{\tau > 0} [A_1(\tau)q_1(\tau) + A_2(\tau)q_2(\tau)]. \quad (\text{A3})$$

Since the two modes  $q_1(\tau)$  and  $q_2(\tau)$  have the same frequency  $\omega(\tau)$ , any sum of the form

$$\begin{aligned} f &= \sum_{\tau > 0, \alpha=1,2} A_\alpha^2(\tau)h(\omega(\tau)) \\ &= \sum_{\tau > 0} [A_1^2(\tau) + A_2^2(\tau)]h(\omega(\tau)) \end{aligned} \quad (\text{A4})$$

can be rewritten using the evenness of  $\omega(\tau)$  and  $|A(\tau)|^2$  in the form

$$f = \sum_{\text{all } \tau} |A(\tau)|^2 h(\omega(\tau)), \quad (\text{A5})$$

where  $h(\omega)$  is an arbitrary function of  $\omega$ . The function  $f(t)$  defined by (5.3) and (2.11) is a sum of this form. Thus, we may apply the formulas of reference 1 based on the real sum  $\sum A_j q_j$  by replacing  $A_j^2$  by  $|A(\tau)|^2$ .