

Calculation of the mobility and the Hall factor for doped *p*-type silicon

Frank Szmulowicz

Research Institute, The University of Dayton, 300 College Park Avenue, Dayton, Ohio 45469-0001

(Received 27 March 1986)

Conductivity and Hall mobilities have been calculated for nondegenerately doped *p*-type silicon. In addition to the two-phonon scattering mechanisms already incorporated by the author earlier [Phys. Rev. B **28**, 5973 (1983)], the ionized- and neutral-impurity scattering are also included. Care has been taken to include many contributions arising from the anisotropic-nonparabolic nature of the bands in modeling the transport. Results of the calculation are compared with data on boron-, aluminum-, gallium-, and indium-doped silicon. The comparison indicates that the Brooks-Herring model, modified to incorporate nonparabolic and most of the anisotropic effects, agrees well with experiments. On the other hand, it is shown that current models of neutral-impurity scattering are not sufficiently accurate as yet. A theory due to Ralph can, with one adjustable parameter, be made to give good agreement with experiments. The calculated ratio of the Hall mobility to the conductivity mobility, the so-called *r* factor, has been used to obtain the hole concentrations from the Hall-coefficient data. Least-squares fitting of the charge-balance equation to the hole-concentration data as a function of temperature produces superior χ^2 and acceptor activation energies compared to those resulting from the use of the *r*=1 approximation.

I. INTRODUCTION

Transport modeling in semiconductors is complicated by the simultaneous presence of several scattering mechanisms which limit carriers' mobility in the same temperature range.^{1,2} Quantitative success of transport models depends critically on the proper treatment of the scattering mechanisms, band structure, and on careful solution of the Boltzmann equation. In general, semiconductors have very complicated valence-band structures and offer a particularly difficult challenge in this regard.³ In a series of papers the conductivity and Hall mobilities of *p*-type silicon have been calculated as limited by the acoustic- and optical-phonon scattering.⁴⁻⁸ This has been achieved without the relaxation-time approximation, while retaining the nonparabolic-anisotropic nature of the valence bands, and in essentially a first-principles spirit. Given that only the intrinsic scattering mechanisms have been included so far, the results hold for relatively pure samples only.

The present paper incorporates another two scattering mechanisms: the ionized- and neutral-impurity scattering. Barring the presence of dislocations, charge clusters, or other less common scattering centers, the four scattering mechanisms determine the mobility of *p*-type silicon. (The hole-hole scattering will be operative at higher dopant concentrations than considered here.) The aim of this paper will be to introduce as few approximations into the treatment as possible. Coupled with the rigorous calculation of the phonon-limited mobility perhaps the limits of validity of ionized- and neutral-impurity scattering models can be assessed. At each step it will, therefore, be important to indicate some of the approximations involved. It will be the task of the following sections to show that good understanding has already been achieved.

Results of the present calculation will find their im-

mediate practical application in the interpretation of the Hall data. As is well known, the Hall coefficient must be multiplied by the *r* factor, the ratio of the Hall-to-conductivity mobility, to give the carrier concentration properly. The common practice of assuming *r*=1 results in wrong dopant concentrations and activation energies as calculated from fits of Hall data to the charge-balance equation.^{9,10}

A comprehensive review of the state-of-the-art in the ionized-impurity scattering is given in the review article by Chattopadhyay and Queisser.¹¹ The neutral-impurity scattering is one of the least understood mechanisms. The impurities of interest here are the neutral acceptors which by analogy with hydrogen can bind an extra carrier.¹²⁻¹⁷ The present day understanding of this scattering mechanism in semiconductors draws from works by Erginsoy,¹⁸ Anselm,¹⁹ Sclar,¹⁰ and McGill and Baron.²¹ Erginsoy, and later McGill and Baron, employ phase shifts calculated for the corresponding electron-neutral hydrogen problem.^{22,23} In practice a scaling factor is necessary to make Erginsoy's theory agree with experiment.²⁴⁻²⁷

There have been several comprehensive calculations of mobilities in *p*-type silicon.^{24,28-33} The way in which these authors approach the problem of phonon-limited mobilities is described in Ref. 8. For the treatment of the ionized-impurity scattering all the authors use the Brooks-Herring (BH) approach. Lin *et al.*³⁰ and Braggins²⁴ introduce various effective masses into the BH equation. Lin *et al.*³⁰ and Li²⁸ use a mixture of conductivity and density-of-states (DOS) effective masses, while Braggins uses curvature effective masses. The present paper will provide a more systematic approach to the problem of including the nonparabolic-anisotropic effects without using the relaxation-time approximation.

Lin *et al.*³⁰ and Li²⁸ employ neutral-impurity relaxation times derived by Sclar²⁰ using separate times for the

light- and heavy-hole bands by inserting DOS masses for these bands into Sclar's expression. This procedure neglects the fact that the bands are coupled even if the interband scattering is negligible. A better procedure is given by Ralph³³ and will be used in this paper. Braggins²⁴ does essentially what Lin *et al.* and Li have done but with the Erginsoy formulation. From Braggins's work it is apparent that Erginsoy's theory gives the correct order of magnitude result but an adjustable parameter is necessary for a good fit to experiment. Given that the phonon-limited mobilities have already been calculated with good precision,⁸ the quantitative predictions about the ionized- and neutral-limited mobilities can be made with more certainty.

In Sec. II the theory of ionized-impurity scattering will be presented together with a more detailed discussion of the state-of-the-art in this field. Section III discusses the neutral-impurity scattering in analogous manner. Section IV will be devoted to the presentation of results for the mobilities and to the use of the calculated r factor in the fit of the Hall carrier concentration data to the charge-balance equation. The quality of the fit will be used as one of the goodness criteria for the calculated r factors. Discussion will be given in Sec. V.

II. IONIZED-IMPURITY SCATTERING FORMALISM

The hole-ionized-impurity scattering rates for p -type silicon have been calculated using the first Born approximation. The formalism is summarized below and its limitations are pointed out. It should be remembered that the formalism adopted here was chosen because it retains most of the solid-state effects relevant to the scattering process. The phase-shift analysis is essentially an atomic-like description of the process, as practiced nowadays, which neglects such solid-state effects as valence-band coupling, band nonparabolicities, and anisotropies (see, for example, Refs. 34–36 and the review in Ref. 11).

A. Comparison of theories

The Brooks-Herring^{37,38} description of the carrier-ionized-impurity scattering has been adopted for the treatment of the transition rates. A detailed derivation of the scattering rates was provided by Falicov and Cuevas³⁹ whose treatment will be extended here to nonparabolic-anisotropic bands. Note that no mention is being made of the scattering potential in these developments. The derived transition rates can be obtained in the first Born approximation from a screened Coulomb potential of the Yukawa type. Yet the Falicov-Cuevas³⁹ treatment gives configuration-averaged transition rates from statistical analysis of an ionized center distribution. Therefore, there may be several potentials leading to the same scattering rates; the Yukawa potential is a convenient representation of just one such potential.

There is a large body of literature dealing with the Yukawa potential and its refinements.¹¹ The phase-shift analysis of the scattering rates is standardly employed with these potentials. Several authors^{35,36,40,41} self-consistently adjust these potentials in order to satisfy the Friedel sum rule.⁴² In general, it would be better to

dispense with potentials altogether since only the average scattering property of the medium is relevant. On the other hand, when screening by free carriers is the dominant screening mechanism each scattering center may scatter independently of the others. Multiple-ion encounters are thus avoided and a single scattering potential for an ion is a meaningful concept. In these cases the phase-shift analysis, the Yukawa potential, and the Friedel sum rules are valid concepts for treatment of the scattering problem. For p -type dopants in silicon with a reasonable amount of compensation the regime of dominant free-carrier screening occurs at temperatures greater than 30–50 K depending on the binding energy of the acceptors. That is precisely the temperature region where the phase-shift analysis and the Born-approximation yield similar results. It has been pointed out by Stern,⁴¹ and also by Krieger and Strauss,⁴⁰ that when the Born approximation is valid it also satisfies the Friedel sum rule.

In the low-temperature region where the Born-approximation becomes invalid the phase-shift analysis fails as well. Here it is the multiple-ion encounters that invalidate the single-ion analysis used in phase-shift calculations.^{43–46} As shown by Meyer and Bartoli⁴⁴ the multi-ion scattering has a negligible effect on mobilities when the screening length is much longer than the electron wavelength. On the other hand, the Born approximation coupled with the statistical analysis of Falicov and Cuevas does have the effect of averaging over the ensemble of scatterers. It may be added that it is not easy to envision a procedure to account for randomness of impurity distribution within the partial wave scheme.³⁴

Refinements to the Yukawa potential involve the use of spatially variable dielectric constants in order to account for a possibly different dielectric response around the impurity than in the rest of the bulk.^{47–60} Various calculations agree that the modifications of the dielectric constant from its long-wavelength limit in the bulk extend no farther than one interatomic distance.^{55,56,59,60} The models for the valence dielectric screening are isotropic and homogeneous rather than anisotropic and nonlocal.^{61,62} Nevertheless, the effect of the variable valence electron dielectric screening on the mobilities has been shown by Resta and Resca⁵⁶ and by Scarfone and Richardson⁵⁷ to be negligible for carrier densities below 10^{18} cm^{-3} . In general, the valence dielectric screening leads to negligible corrections to the BH results for nondegenerately doped silicon.^{56,57}

It is known that the Falicov-Cuevas scattering rates should fail at some low temperature where the free-electron screening is diminished. There have been several proposals to alleviate this shortcoming of the theory. By and large these proposals are untested. The temperature region in question is also the region where much bigger problems arise in the theory.

There is a body of literature concerned with the limits to the validity of the Born approximation.^{11,34–36} The Born approximation is known to be good for high-carrier energies which is the case at higher temperatures.³⁴ In general, one wants the perturbation to the wave function by the scattering potential to be small. Meyer and Bartoli³⁶ use an equation which shows the deviation of the wave

function from the plane-wave form⁶³

$$\psi'(\mathbf{r}) = -\frac{m^*}{2\pi\hbar^2} \int d\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} e^{ikz'} V(\mathbf{r}'), \quad (2.1)$$

which with a change of variables becomes

$$\psi'(\mathbf{r}) = -\frac{m^*}{2\pi\hbar^2} e^{ikz} \int \frac{e^{ik(R+Z)}}{R} V(\mathbf{r}+\mathbf{R}) d\mathbf{R}, \quad (2.2)$$

where $\mathbf{R}=\mathbf{r}-\mathbf{r}'$. It is usual at this point to let $\mathbf{r}=0$ and take the validity of the Born approximation to mean

$$|\psi'(\mathbf{r}=0)| \ll 1. \quad (2.3)$$

This is a very pessimistic outlook^{64,65} since it examines the wave function at the point $\mathbf{r}=0$ where the potential is actually infinite. Anyway, evaluating Eq. (2.3) for a screened Coulomb potential one gets the condition that

$$L, a_0 \gg \lambda, \quad (2.4)$$

where L is the screening length, a_0 is the hydrogenic radius, and λ is the wavelength of the incident charge carrier.

Equation (2.4) indicates that the Born approximation is bound to fail in the long-wavelength limit. In fact, examining the Born and the phase-shift studies of Meyer and Bartoli³⁶ shows that for moderate doping levels, less than 10^{17} cm^{-3} or so, and at temperatures above carrier freeze-out both ways of calculating mobilities agree very closely. The Meyer and Bartoli calculations do not include compensation and activation of carriers from the parent impurities and scattering mechanisms other than the ionized-impurity scattering. Perhaps rather fortuitously their own calculation for n -type GaAs (Ref. 36) and n -type Si (Ref. 65) shows better agreement with experiment for the Born approximation than for the phase-shift analysis.

There is a list of other potentially important effects that can arise in carrier-ionized-impurity scattering. When the field of the ionized center is perturbed by the oncoming charge carrier, i.e., the polarization effects, the treatment of Takimoto⁶⁷ and McIrvine⁶⁸ becomes relevant. Takimoto showed though that these complications arise only in the high carrier-concentration regime, greater than 10^{17} cm^{-3} . The samples used in the present work for comparison with theory have carrier concentrations $p < 10^{17} \text{ cm}^{-3}$. It is also known that the field of an ionized acceptor deviates from the simple screened Yukawa potential in the vicinity of the ion core. This contribution to the carrier scattering goes by the name of central-cell scattering. Ralph *et al.*⁶⁹ have calculated mobilities for n -type dopants in silicon and germanium including the central-cell scattering. Their model involves a short-range potential added to a screened Coulomb potential. The short-range potential is determined from the donor binding energy. Even here the central-cell contribution to the scattering is not important at room temperature for charge densities less than 10^{17} cm^{-3} . For higher densities mobilities are reduced from the purely Coulomb screened mobilities yet a sizable discrepancy with experiments remains. The authors suspect electron-electron and charge-cluster scattering has an effect on the results.

In summary of various approaches to the problem it appears that the standard Brooks-Herring theory should be tried first since it offers the best realization of a theory which employs both band nonparabolicities and anisotropies. These effects were shown to be crucially important in the treatment of phonon-limited mobilities. The phase-shift analysis obviously has its place as a valid scheme whose region of validity significantly overlaps the region of validity for the Born approximation. In the case of hole transport the proper phase-shift calculation ought to reflect the coupled valence-band nature of the resulting Hamiltonian as in the work of Ralph.³³

B. Transition rates—present theory

The transition rate for hole scattering from band n with wave vector $\mathbf{k}$ to band m with wave vector $\mathbf{k}'$ is given in the first Born approximation as

$$P(n, \mathbf{k}, m, \mathbf{k}') = \frac{2\pi}{\hbar} |\langle \psi_{n\mathbf{k}} | V | \psi_{m\mathbf{k}'} \rangle|^2 \delta(E_{n\mathbf{k}} - E_{m\mathbf{k}'}), \quad (2.5)$$

where $\langle \mathbf{r} | \psi_{n\mathbf{k}} \rangle$ is the Bloch wave function for the hole, V is the scattering potential, $E_{n\mathbf{k}}$ are the band energies, and the delta function ensures energy conservation for ionized-impurity scattering. Expanding the Bloch waves for $\mathbf{k}$ in the vicinity of the Brillouin-zone center in terms of Kohn-Luttinger⁷⁰ wave functions results in

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_m a_{nm}(\mathbf{k}) \chi_{m\mathbf{k}}(\mathbf{r}), \quad (2.6)$$

where the Kohn-Luttinger wave function is given by

$$\chi_{n\mathbf{k}} = u_{n\mathbf{k}_0}(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (2.7)$$

$u_{n\mathbf{k}_0}(\mathbf{r})$ is the periodic part of the Bloch wave at $\mathbf{k}_0$. The wave vector $\mathbf{k}_0$ is the expansion point for the series (2.6) and for p -type silicon is the Brillouin-zone center, point Γ .

Using the expansion (2.6) in the equation for the transition rates, Eq. (2.5), gives for the matrix element

$$\langle \psi_{n\mathbf{k}} | V | \psi_{m\mathbf{k}'} \rangle = \sum_{p,q} a_{np}^*(\mathbf{k}) a_{mq}(\mathbf{k}') \langle \chi_{p\mathbf{k}} | V | \chi_{q\mathbf{k}'} \rangle. \quad (2.8)$$

For sufficiently "gentle" potentials⁷⁰

$$\langle \chi_{p\mathbf{k}} | V | \chi_{q\mathbf{k}'} \rangle = \delta_{pq} V(\mathbf{k} - \mathbf{k}'), \quad (2.9)$$

where $V(\mathbf{k} - \mathbf{k}')$ is the Fourier transform of $V(\mathbf{r})$.

The transition rate for scattering from band manifold N with wave vector $\mathbf{k}$ to band manifold M with wave vector $\mathbf{k}'$ will now be given by

$$P(N, \mathbf{k}; M, \mathbf{k}') = \frac{2\pi}{\hbar} \left[\frac{1}{2} \sum_{n(N)} \sum_{m(M)} \left| \sum_p a_{np}^*(\mathbf{k}) a_{mp}(\mathbf{k}') \right|^2 \right] \times |V(\mathbf{k} - \mathbf{k}')|^2 \delta(E_{N\mathbf{k}} - E_{M\mathbf{k}'}), \quad (2.10)$$

where the sum within the large parentheses proceeds over members of the Kramers degenerate band manifolds N and M .

The approach of Brooks and Herring^{37,38} as given by Falicov and Cuevas³⁹ is to calculate the Fourier transform $V(\mathbf{k}-\mathbf{k}')$ from the Poisson equation. The method lets the carriers and ionized impurities redistribute themselves in a manner which produces mutual screening. Ensemble average of the resulting scattering rate is performed by assuming complete randomness of both the ionized majority and minority ion distributions and by assuming the two distributions to be uncorrelated. The configuration-averaged result is

$$\langle |V(\mathbf{k}-\mathbf{k}')|^2 \rangle = \frac{1}{\Omega} \left[\frac{4\pi e^2}{\epsilon} \right]^2 \frac{2N_D + p}{\left[|\mathbf{k}-\mathbf{k}'|^2 + \frac{1}{L_D^2} \right]^2}, \quad (2.11)$$

where Ω is the volume of the crystal, e is the electronic charge, ϵ is the static dielectric constant (11.7 for silicon), N_D is the density of donors, p is the hole concentration in the valence band, and $(2N_D + p)$ is the total ionized-impurity concentration. The Debye length is given by

$$1/L_D^2 = 4\pi e^2 / \epsilon k_B T p_s, \quad (2.12)$$

where k_B is the Boltzmann constant, T is the absolute temperature, and the screening charge density p_s is²

$$p_s = p + \sum_i N_i^0 \frac{N_i^-}{N_i}, \quad (2.13)$$

where the sum is over all acceptors of density N_i , with N_i^- acceptors ionized, and N_i^0 acceptors neutral. Equations (2.11) and (2.12) assume that all the donors are ionized so that $N_D^+ = N_D$ and the present treatment will exclude the temperature region where there is a significant amount of intrinsic electrons, $T > 350$ K for lightly doped p -type silicon.

Owing to the linearization of the Poisson equation in deriving Eq. (2.11) the treatment should break down at some low temperature for samples with compensating donors.³⁹ As discussed above the Born approximation breaks down at low temperatures anyway. Several schemes for calculating the screening¹¹ at low temperatures have been proposed but are not widely used. In practice, the expression for p_s , Eq. (2.13), appears to work rather well and its limitations are best assessed by comparison of experiment and theory. The reader may consult the review article by Chattopadhyay and Queisser¹¹ for further discussion of many other effects which may complicate the hole-ionized-impurity scattering. Their inclusion in the present work is not warranted before the major effects explored here are evaluated.

The unusual form of Eq. (2.13) for p_s should make it clear that we are not dealing with an ordinary scattering potential but with some average scattering property of the medium. Only when $p_s \approx p$ are the ions sufficiently screened to be regarded as isolated scattering centers.

C. Overlap factor

The overlap factor in Eq. (2.10) is given by

$$O_{NM}(\mathbf{k}, \mathbf{k}') \equiv \frac{1}{2} \sum_{n(N)} \sum_{m(M)} \left| \sum_p a_{np}^*(\mathbf{k}) a_{mp}(\mathbf{k}') \right|^2. \quad (2.14)$$

In a previous work by the author the transition rates for acoustic- and optical-phonon scattering process were expanded in a complete set of cubic harmonics.⁸ This approach is not practical here since the Fourier transform squared, Eq. (2.11), is very sharply peaked in the forward direction $\mathbf{k} = \mathbf{k}'$. Instead, I retain Eq. (2.11) as is and fit the angular variations for the overlap factor O_{NM} , Eq. (2.14).

One notes several symmetry properties on examining Eqs. (2.10) and (2.14). First, the microscopic reversibility holds

$$P_{NM}(\mathbf{k}, \mathbf{k}') = P_{MN}(\mathbf{k}', \mathbf{k}), \quad (2.15)$$

since the collision is elastic. Similarly,

$$O_{NM}(\mathbf{k}, \mathbf{k}') = O_{MN}(\mathbf{k}', \mathbf{k}), \quad (2.16)$$

on general grounds. Additionally though,

$$O_{NM}(-\mathbf{k}, -\mathbf{k}') = O_{NM}(\mathbf{k}, -\mathbf{k}') = O_{NM}(-\mathbf{k}, \mathbf{k}') = O_{NM}(\mathbf{k}, \mathbf{k}'), \quad (2.17)$$

which does not hold for the total rate $P_{NM}(\mathbf{k}, \mathbf{k}')$ [but $P_{NM}(-\mathbf{k}, -\mathbf{k}') = P_{NM}(\mathbf{k}, \mathbf{k}')$]. These properties are the consequences of the inversion and time-reversal symmetries.

Since the overlap factor is even in both $\mathbf{k}$ and $\mathbf{k}'$, only even representations can occur in its expansion in cubic harmonics. From the point-group symmetry one also has

$$O_{NM}(R\mathbf{k}, R\mathbf{k}') = O_{NM}(\mathbf{k}, \mathbf{k}'), \quad (2.18)$$

which requires that the cubic-harmonic expansion have the form

$$O_{NM}(\mathbf{k}, \mathbf{k}') = \sum_i \sum_{L, L'} \mathcal{O}_{NL, ML}^i(\mathcal{E}) \times \left[\sum_{\mu=1}^{l_i} K_L^{i\mu}(\hat{\mathbf{k}}) K_{L'}^{i\mu}(\hat{\mathbf{k}}') \right] \quad (2.19)$$

in standard notation, where $\mathcal{O}$ is the energy-dependent coefficient and the sum proceeds over the five even irreducible representations of the cubic point group O_h ($R \in O_h$). Unfortunately, this degree of rigor is very expensive to implement on a computer since it requires four-dimensional integrals to be performed in setting up solutions to the Boltzmann equation, Sec. II D.

Instead of the expansion (2.19) I have used a complete set of functions in relative angle between the incident $\hat{\mathbf{k}}$ and scattered $\hat{\mathbf{k}}'$ directions, i.e., the Legendre polynomials, so that

$$O_{NM}(\mathbf{k}, \mathbf{k}') = \sum_L \mathcal{O}_L^{NM}(\mathcal{E}) P_L(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}'). \quad (2.20)$$

One notes the following identity:

$$P_L(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') = \frac{4\pi}{2L+1} \sum_{i, \mu} K_L^{i\mu}(\hat{\mathbf{k}}) K_L^{i\mu}(\hat{\mathbf{k}}'), \quad (2.21)$$

which arises from unitarity of transformation between spherical and cubic harmonics. With (2.21) inserted in (2.20) one can see by comparison with Eq. (2.19) that the latter expansion is more general than Eq. (2.20). Nevertheless, the $\mathcal{O}$ coefficients in (2.20) will be obtained via a least-squares fit to the calculated O_{NM} 's and should be accurate in the least-squares sense. From Eq. (2.17) only even L values occur in (2.20).

A large number of O_{NM} 's were calculated. The $(\hat{\mathbf{k}}, \hat{\mathbf{k}}')$ mesh consisted of 378 direction pairs unrelated by the symmetry condition, Eq. (2.18). O_{NM} 's were obtained on a linear energy mesh of 52 energies from 0.0045 to 0.234 eV and a logarithmic energy mesh of 16 points given by

$$E = \exp[-13.305 + (j-1)0.6]$$

(measured in eV) with $1 \leq j \leq 16$. The least-squares fit procedure is similar to that developed in Ref. 5. I solve the following linear simultaneous equation for $\mathcal{O}_L^{NM}(\mathcal{E})$:

$$C_L = \sum_{L'} B_{LL'} \mathcal{O}_L^{NM}(\mathcal{E}), \quad (2.22)$$

where

$$C_L = \sum_{\hat{\mathbf{k}}, \hat{\mathbf{k}}'} O_{NM}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E}) P_L(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}'), \quad (2.23)$$

$$B_{LL'} = \sum_{\hat{\mathbf{k}}, \hat{\mathbf{k}}'} P_L(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') P_{L'}(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}'), \quad (2.24)$$

so that $\mathcal{O} = B^{-1}C$. The band energies and $a_{nm}(\mathbf{k})$ in Eq. (2.2) are obtained from solutions of the 6×6 Kane's $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian.

Using the same band parameters as in previous work,⁸ I obtain the expansion coefficients $\mathcal{O}$ as a function of energy displayed in Fig. 1. The fit was done with $L=0, 2, 4, 6$, and 8 where the higher angular coefficients, although small, are not negligible. It is apparent that the nonparabolic-anisotropic structure of the valence bands results in strongly energy-dependent overlap expansion coefficients, especially below the spin-orbit energy of 44 meV.

It is interesting to compare the results in Fig. 1 with those obtained for the isotropic-parabolic valence-band model. Here the valence-band Hamiltonian is given by (infinite spin-orbit coupling limit)⁷¹

$$H = \frac{\gamma_1 p^2}{2m} - \frac{\mu}{9} \frac{\gamma_1}{2m} (P^{(2)} \cdot J^{(2)}), \quad (2.25)$$

where $P^{(2)}$ and $J^{(2)}$ are irreducible second-rank tensors for pseudospin $J = \frac{3}{2}$, with eigenenergies,

$$E = \frac{\gamma_1 \hbar^2 k^2}{2m} (1 - \mu) \quad (2.26)$$

for $J_z = \pm \frac{3}{2}$ states (heavy-hole band H), and

$$E = \frac{\gamma_1 \hbar^2 k^2}{2m} (1 + \mu) \quad (2.27)$$

for $J_z = \pm \frac{1}{2}$ (light-hole band L). The split-off band S corresponds to $J = \frac{1}{2}$ and $J_z = \pm \frac{1}{2}$. The wave functions for the three top valence bands, in absence of the scattering potential, are for the heavy-hole band

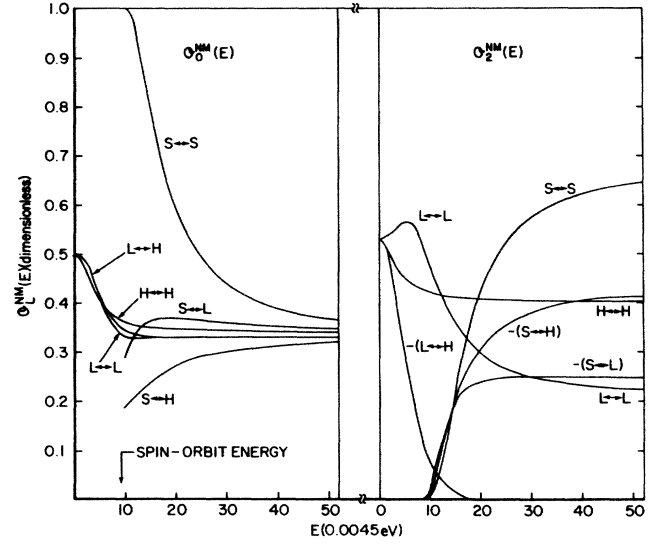


FIG. 1. The zeroth- and second-order Legendre polynomial expansion coefficients of the band overlaps as a function of energy. S , L , and H stand for the spin-orbit, light-, and heavy-hole bands, respectively.

$$e^{i\mathbf{k} \cdot \mathbf{r}} \left| \frac{3}{2} \pm \frac{3}{2} \right\rangle, \quad (2.28a)$$

for the light-hole band

$$e^{i\mathbf{k} \cdot \mathbf{r}} \left| \frac{3}{2} \pm \frac{1}{2} \right\rangle, \quad (2.28b)$$

and for the spin-orbit band

$$e^{i\mathbf{k} \cdot \mathbf{r}} \left| \frac{1}{2} \pm \frac{1}{2} \right\rangle. \quad (2.28c)$$

Far from the scattering potential $\mathbf{k}$, J , and J_z continue to be good quantum numbers and since $J = \frac{3}{2}$ for the L, H bands and $J = \frac{1}{2}$ for the S band, there is no interband scattering with the S band. Figure 1 shows that there indeed is a nonvanishing overlap between S and L, H bands which is not reflected by the spherical-parabolic model owing to the finite spin-orbit coupling.

Let us then consider scattering among the two top valence bands with $J = \frac{3}{2}$. Let the hole be incident along the $k = k_z$ direction, with J_z quantum number along the z direction being n . The scattered hole propagates along direction $\mathbf{k}'$. The scattered hole must belong to a definite band so that the projection of $\mathbf{J}$ onto the $\mathbf{k}'$ direction is m . In terms of eigenstates quantized along the k_z axis, $|jm_j\rangle$, the scattered state is³³

$$R(0, \theta_{\mathbf{k}'}, \phi_{\mathbf{k}}) \left| \frac{3}{2} m \right\rangle, \quad (2.29)$$

where R is the rotation operator and $\theta_{\mathbf{k}'}, \phi_{\mathbf{k}}$ are angles defining the direction of $\mathbf{k}'$. The overlap between the initial and final states is given by

$$\langle \frac{3}{2} n | R(0, \theta_{\mathbf{k}'}, \phi_{\mathbf{k}}) | \frac{3}{2} m \rangle = D_{nm}^{3/2}(\mathbf{k} - \mathbf{k}'), \quad (2.30)$$

where the D are the rotation matrices. For calculations of the scattering rates we square the overlap, sum it over initial and final states, and divide by two. Therefore,

$$\begin{aligned}
O_{LL} &= \frac{1}{2} \sum_{n=\pm\frac{1}{2}} \sum_{m=\pm\frac{1}{2}} |D_{nm}^{3/2}(\mathbf{k}-\mathbf{k}')|^2 = 1 - \frac{3}{4} \sin^2 \theta, \\
O_{HH} &= \frac{1}{2} \sum_{n=\pm\frac{3}{2}} \sum_{m=\pm\frac{3}{2}} |D_{nm}^{3/2}(\mathbf{k}-\mathbf{k}')|^2 = 1 - \frac{3}{4} \sin^2 \theta, \\
O_{LH} &= \frac{1}{2} \sum_{n=\pm\frac{1}{2}} \sum_{m=\pm\frac{3}{2}} |D_{nm}^{3/2}(\mathbf{k}-\mathbf{k}')|^2 = \frac{3}{4} \sin^2 \theta, \quad (2.31)
\end{aligned}$$

where $\cos \theta = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}'$. Note that in terms of Legendre polynomials

$$O_{LL} = O_{HH} = 1 - \frac{3}{4} \sin^2 \theta = \frac{1}{2} [P_0 + P_2], \quad (2.32)$$

$$O_{HL} = O_{LH} = \frac{3}{4} \sin^2 \theta = \frac{1}{2} [P_0 - P_2]. \quad (2.33)$$

In Fig. 1 the $\mathcal{O}_0^{NM} \simeq \frac{1}{2}$ for HH , HL , and LL scattering as $\mathcal{E} \rightarrow 0$ while $\mathcal{O}_2^{LL}$, $\mathcal{O}_2^{HH}$ are about $\frac{1}{2}$ and $\mathcal{O}_2^{LH} \simeq -\frac{1}{2}$ which we expect from Eqs. (2.32) and (2.33). For infinite spin-orbit coupling $O_{SS} = 1 = P_0$. In Fig. 1, though, the first two coefficients in the Legendre expansion are seen to vary significantly with energy. The coefficients appear to level off for energies higher than the spin-orbit energy. Clearly, the difference between results of Eqs. (2.32) and (2.33) and Fig. 1 are due to significant hybridizations of the three bands so that neither is a pure eigenstate of J or J_z .

D. The Boltzmann equation

The Boltzmann equation is given by⁸

$$\begin{aligned}
&\left[\frac{e}{\hbar} \mathbf{F} \cdot \nabla_{\mathbf{k}} - \frac{e}{\hbar^2} \mathbf{B} \cdot [\nabla_{\mathbf{k}} E_N(\mathbf{k}) \times \nabla_{\mathbf{k}}] \right] f_N(\mathbf{k}) \\
&= \sum_M \frac{\Omega}{(2\pi)^3} \int d\mathbf{k}' P^{\text{ion}}(M, \mathbf{k}'; N, \mathbf{k}) [f_M(\mathbf{k}') - f_N(\mathbf{k})] \\
&+ \dots, \quad (2.34)
\end{aligned}$$

where the ellipsis here represents phonon scattering terms and $\mathbf{F}$ and $\mathbf{B}$ are the electric and magnetic fields, respectively. Using the expansions

$$f_N(\mathbf{k}) = f_N^0(\mathbf{k}) [1 + \mathbf{F} \cdot \Phi_N(\mathbf{k}) + (\mathbf{F} \times \mathbf{B}) \cdot \mathbf{X}_N(\mathbf{k}) + \dots], \quad (2.35)$$

$$\Phi_N(\mathcal{E}, \hat{\mathbf{k}}) = \sum_{\lambda} e \left[\frac{4\pi}{k_B T} \right]^2 \mathbf{K}_{\lambda}^{\delta}(\hat{\mathbf{k}}) \theta_{N\lambda}^{\delta}(\mathcal{E}, T), \quad (2.36)$$

$$\mathbf{X}_N(\mathcal{E}, \hat{\mathbf{k}}) = \sum_{\lambda} \left[\left[\frac{4\pi}{k_B T} \right]^3 \frac{4\pi e^2}{\hbar} \right] \mathbf{K}_{\lambda}^{\delta}(\hat{\mathbf{k}}) \xi_{N\lambda}^{\delta}(\mathcal{E}, T), \quad (2.37)$$

as in Ref. 8, results in a system of linear equations for θ and ξ . The scattering matrix for the ionized-impurity scattering is given by integrals over constant energy $\mathcal{E}$ surfaces

$$\begin{aligned}
[S_{N\lambda'', M\lambda'}^{\delta}(\mathcal{E})]_{\text{ion}} &= \left[\frac{1}{3} \left[\frac{8\pi e^2}{\epsilon} \right]^2 \frac{2N_D + p}{k_B T} \right] \\
&\times \int d\hat{\mathbf{k}} \int d\hat{\mathbf{k}}' \left[\frac{\Gamma_N(\mathcal{E}, \hat{\mathbf{k}}) \Gamma_M(\mathcal{E}, \hat{\mathbf{k}}') O_{NM}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E}) \sum_{\mu} K_{\lambda''}^{\delta\mu}(\hat{\mathbf{k}}) K_{\lambda'}^{\delta\mu}(\hat{\mathbf{k}}')}{\left[|\mathbf{k}_N - \mathbf{k}_M|^2 + \frac{1}{L_D^2} \right]^2} \right. \\
&\quad \left. - \delta_{NM} \sum_P \frac{\Gamma_N(\mathcal{E}, \hat{\mathbf{k}}) \Gamma_P(\mathcal{E}, \hat{\mathbf{k}}') O_{NP}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E}) \sum_{\mu} K_{\lambda''}^{\delta\mu}(\hat{\mathbf{k}}) K_{\lambda'}^{\delta\mu}(\hat{\mathbf{k}}')}{\left[|\mathbf{k}_N - \mathbf{k}_P|^2 + \frac{1}{L_D^2} \right]^2} \right]. \quad (2.38)
\end{aligned}$$

In Eq. (2.38) $\mathbf{k}_N$ is the wave vector for band N and energy $\mathcal{E}$ in direction $\hat{\mathbf{k}}$. Other terms are as defined in Ref. 8. Note that the matrix S is symmetric in the combined $M\lambda$ indices.

Equation (2.38) is a general matrix element which in particular has terms for intraband $N=M$ and interband scattering. The angular momentum indices used in the present work range through $\lambda', \lambda'' = 1, 3, 5, 7$, which indicates that the formalism goes beyond the relaxation approximation where $\lambda', \lambda'' = 1$. The relaxation-time approximation is equivalent to retaining only the $\lambda=1$ term in the cubic-harmonic expansions of the distribution function, Eqs. (2.35)–(2.37).

There is a limited number of simplifications one may

accomplish with Eq. (2.38). It is possible to reduce the double angular integrations to within one irreducible wedge of the Brillouin zone yet the four-dimensional integration still has to be performed. The Debye length L_D depends on temperature, concentrations, and identities of dopants, so that there are a large number of these integrals to be performed. This would be inordinately time consuming.

Instead, I simplify certain of the quantities appearing in Eq. (2.38). $\Gamma_N(\mathcal{E}, \hat{\mathbf{k}})$ and calipers $k_N(\hat{\mathbf{k}}, \mathcal{E})$ both have the cubic symmetry and their cubic-harmonic expansion have large spherical terms $L=0$ and smaller $L=4, 6, 8, \dots$ contributions. I neglect those higher-order terms and compute spherically averaged quantities by integrating

over the solid angle:

$$k_N(\mathcal{E}) \equiv \frac{1}{4\pi} \int d\hat{\mathbf{k}} k_N(\hat{\mathbf{k}}, \mathcal{E}), \quad (2.39)$$

$$\Gamma_N(\mathcal{E}) \equiv \frac{1}{4\pi} \int d\hat{\mathbf{k}} \Gamma_N(\hat{\mathbf{k}}, \mathcal{E}). \quad (2.40)$$

In the spherical-parabolic band model one would have

$$k_N(\mathcal{E}) = (2m_0/\hbar^2)^{1/2} (m^* \mathcal{E}/m_0)^{1/2}, \quad (2.41)$$

$$\Gamma_N(\mathcal{E}) = (m^*/m_0)^{3/2} \mathcal{E}^{1/2}. \quad (2.42)$$

The spherically averaged quantities (2.39) and (2.40) are used only in evaluating the scattering matrix Eq. (2.38) and nowhere else. Those quantities still retain the band nonparabolicity through the energy dependence. Further analytic progress with Eq. (2.38) would have been impossible without these simplifications.

In order to obtain an explicit expression for the scattering matrix, Eq. (2.38), I use expansion (2.20) for the overlaps and replace Γ_N, k_N by their isotropic averages, Eqs. (2.39) and (2.40). The four-dimensional integral can be performed analytically to yield

$$S_{N\lambda'', M\lambda'}^{\delta}(\mathcal{E})_{\text{ion}} = \delta_{\lambda', \lambda''} \left[-\pi \left[\frac{8\pi e^2}{\epsilon} \right]^2 \frac{2N_D + p}{k_B T} \right] \left[\frac{\Gamma_N(\mathcal{E}) \Gamma_M(\mathcal{E})}{k_N^2(\mathcal{E}) k_M^2(\mathcal{E})} \sum_L \mathcal{O}_L^{NM}(\mathcal{E}) \frac{d}{dz} [P_{L<}(z) Q_{L>}(z)]_{z_{NM}} \right. \\ \left. - \delta_{NM} \sum_P \frac{\Gamma_N(\mathcal{E}) \Gamma_P(\mathcal{E})}{k_N^2(\mathcal{E}) k_P^2(\mathcal{E})} \sum_L \mathcal{O}_L^{NP}(\mathcal{E}) \left[\frac{dQ_L(z)}{dz} \right]_{z_{NP}} \right], \quad (2.43)$$

where $L_{<} = \min\{L, \lambda'\}$ and $L_{>} = \max\{L, \lambda'\}$, and where

$$z_{NM} = \frac{k_N^2(\mathcal{E}) + k_M^2(\mathcal{E}) + L_D^2}{2k_N(\mathcal{E})k_M(\mathcal{E})},$$

and $Q_L(z)$ is the Legendre polynomial of the second kind defined everywhere except along the cut on the real axis from $-\infty$ to $+1$.

One may want to note several familiar properties of the scattering matrix, which is related to the relaxation time. First, for intraband scattering, $N=M$, the following can be factored out from Eq. (2.38):

$$\sum_{\mu} K_{\lambda}^{\delta\mu}(\hat{\mathbf{k}}) [K_{\lambda'}^{\delta\mu}(\hat{\mathbf{k}}) - K_{\lambda}^{\delta\mu}(\hat{\mathbf{k}})], \quad (2.44)$$

which vanishes for $\hat{\mathbf{k}} = \hat{\mathbf{k}}'$, i.e., for forward scattering. If we keep $\lambda=1$ only in Eq. (2.44) we obtain $(1 - \cos\theta)$ which is the correct weighting factor for the momentum relaxation time in the relaxation-time approximation. Second, using Eq. (2.38) with $\lambda' = \lambda'' = 1$ and the overlaps for the two-band spherical-parabolic model, Eqs. (2.32) and (2.33), one will recover the interband and intraband relaxation times of Bir, Normantas, and Pikus.⁷² Third, retaining the intraband terms only with $\lambda' = \lambda'' = 1$ and

$$\mathcal{O}_L^{NM} = \delta_{Nm} \delta_{L0} \quad (2.45)$$

results in the Brooks-Herring relaxation times.

It is incorrect to take Eq. (2.43) and make the following replacements:

$$k_N(\mathcal{E}) \rightarrow k_N(\hat{\mathbf{k}}, \mathcal{E}), \quad (2.46a)$$

$$\Gamma_N(\mathcal{E}) \rightarrow \Gamma_N(\hat{\mathbf{k}}, \mathcal{E}), \quad (2.46b)$$

and then average S over the solid angle. The angular integrals have already been done in going from (2.38) to (2.43) and the reintroduction of angular dependence is artificial rigor. This has been done by Takeda *et al.*²⁹ in their Eqs. (2.17) and (2.19), and by Nakagawa and Zukotynski³² in their Eq. (14).

For low-carrier energies, or in the limit of parabolic bands, $\Gamma_N \sim \mathcal{E}$, $k_N^2 \sim \mathcal{E}$, so that

$$S \sim \begin{cases} N_I / \mathcal{E} T, & k^2 L_D^2 \gg 1 \\ \mathcal{E} T / N_I, & k^2 L_D^2 \ll 1. \end{cases} \quad (2.47a)$$

$$(2.47b)$$

Using Eqs. (2.67) or (2.68) of Ref. 8 for the mobilities one obtains

$$\mu \sim \begin{cases} T^{3/2} / N_I \\ N_I T^{-5/2} \end{cases} \quad (2.48a)$$

$$(2.48b)$$

for the cases of Eqs. (2.47a) and (2.47b), respectively.

The matrix equations for θ and ξ in Eqs. (2.36) and (2.37) were set up as in Eqs. (2.56) and (2.57) of Ref. 8 and then mobilities μ_c and μ_H were calculated from Eqs. (2.67) and (2.68) of that reference. The treatment of the acoustic- and optical-phonon scattering is identical to that in the earlier work. For future reference, I define partial band mobilities and Hall factors for each band in the following manner:

$$\mu_c = \sum_N p_N \mu_c^N / p, \quad (2.49)$$

$$\mu_c \mu_H = \sum_N p_N \mu_c^N \mu_H^N / p, \quad (2.50)$$

$$p = \sum_N p_N, \quad (2.51)$$

$$r_N = \mu_H^N / \mu_c^N, \quad (2.52)$$

which defines μ_c^N, μ_H^N , and r_N for each band N . The total r factor

$$r = \mu_H / \mu_c = \sum_N r_N p_N (\mu_c^N / \mu_c)^2 / p \quad (2.53)$$

has contributions from all bands weighted by both p_N and μ_c^N . Therefore, even if all r_N were equal, r itself may not equal any of the r_N owing to the band multiplicity. More familiar conductivity tensors are given by⁷³

$$\sigma_{11} = ep\mu_c = \sum \sigma_{11}^N, \quad (2.54)$$

$$\sigma_{123} = ep\mu_c\mu_H = \sum \sigma_{123}^N, \quad (2.55)$$

with

$$\sigma_{11}^N = ep_N\mu_c^N, \quad (2.56)$$

$$\sigma_{123}^N = ep_N\mu_c^N\mu_H^N. \quad (2.57)$$

III. NEUTRAL-IMPURITY SCATTERING

A. Erginsoy formulation

Neutral impurities have short-range potentials in comparison to the range of the Coulomb potential. One type of neutral impurities, similar to carbon and oxygen in silicon, have scattering cross sections on the order of their geometrical size. This number being small allows us to neglect this type of neutral-impurity scattering.²

There is another type of neutral impurity which can scatter strongly. At low temperatures acceptors are not ionized and there is a large number of neutral hydrogen-like acceptor centers. It is known that neutral hydrogen in empty space can bind a second electron with the binding energy of 0.75 eV, as compared with the binding energy of 13.6 eV for the first electron. It is this ability to bind an extra carrier which points to the possibility of strong scattering, especially for slow carriers at lower temperatures. The s -wave phase shift for electron-hydrogen scattering was worked out by Massey and Moiseiwitsch²² and applied to semiconductors by Erginsoy.¹⁸ The result for the differential scattering cross section, for the assumed elastic scattering, is

$$\sigma(\theta, \phi) = \frac{20a_0}{4\pi k}, \quad (3.1)$$

where a_0 is the Bohr radius and $k = (2m^*\mathcal{E}/\hbar^2)^{1/2}$ is the wave vector. For hydrogenic acceptors

$$a_0 = \frac{\epsilon\hbar^2}{e^2m^*}, \quad (3.2)$$

and the binding energy of the hole is

$$E_B = -\frac{m^*e^4}{2\epsilon^2\hbar^2}. \quad (3.3)$$

It is clear that Erginsoy's approach employs several approximations. Only $l=0$ phase shift is used and the bands are taken to be isotropic and uncoupled. If one wishes the dependence on energy of the scattered hole can be taken into account in the effective mass m^* . It is unclear which of the many effective masses one should use. Li²⁸ and Lin *et al.*³⁰ employ density-of-states effective masses. Braggins²⁴ uses the density-of-states effective mass of each band to obtain the scaled Bohr radius a_0 and a curvature related mass in the rest of the mobility calculation.

There still remains the problem that the acceptors are not hydrogenic. Baron *et al.*¹² attempt to account for this fact by noting the identity

$$E_B a_0 = \frac{e^2}{2\epsilon}, \quad (3.4)$$

which holds for Coulomb centers only, and then use

$$a_0 = \frac{e^2}{2\epsilon E_B}, \quad (3.5)$$

with experimentally measured E_B in Eqs. (3.1). In the same context, I would like to observe that another identity, which holds for Coulomb centers,

$$E_B a_0^2 = \frac{\hbar^2}{2m^*} \quad (3.6)$$

gives a_0 quite a bit different (80% difference for In) from the former estimate, Eq. (3.5). Nevertheless, Eq. (3.5) will be used in the present work since an adjustable parameter for the strength of Erginsoy's neutral impurity scattering will be needed anyway. Equation (3.5) alone does not correct for the difference in neutral impurity scattering by the various acceptors in silicon.

In order to implement Erginsoy's theory within the present formalism a few intermediate steps are noted. The transition rate for hole scattering between bands n and m with the change of wave vector from $\mathbf{k}$ to $\mathbf{k}'$ is

$$P_{nm}(\mathbf{k}, \mathbf{k}') = 2\pi/\hbar |T_{nm}(\mathbf{k}, \mathbf{k}')|^2 \delta(E_n(\mathbf{k}) - E_m(\mathbf{k}')) \quad (3.7)$$

from Fermi's golden rule. The square of the T matrix is related to the differential scattering cross section. For the isotropic scattering considered here

$$|T_{nm}(\mathbf{k}, \mathbf{k}')|^2 = \left[\frac{4\pi\hbar^2}{2m_n\Omega} \right]^2 \sigma_{nm}(\theta) \delta_{nm}, \quad (3.8)$$

where only intraband scattering is considered and θ is the relative angle between $\mathbf{k}$ and $\mathbf{k}'$. Using Eq. (3.1) results in

$$P_{nm}(\mathbf{k}, \mathbf{k}') = \frac{40\pi^2\hbar^3 a_0}{\Omega^2 m_n^2 k_n} O_{nm}(\mathbf{k}, \mathbf{k}') \times \delta(E_n(\mathbf{k}) - E_m(\mathbf{k}')) \delta_{nm}, \quad (3.9)$$

where the band overlap $O_{nm}(\mathbf{k}, \mathbf{k}')$ has been added, and [see Eq. (2.39)]

$$k_n(\mathcal{E}) = \langle k_n(\hat{\mathbf{k}}, \mathcal{E}) \rangle \quad (3.10)$$

is the spherically averaged constant $\mathcal{E}$ surface caliper. The effective mass m_n arises from velocity considerations, or band curvature, so that

$$\mathcal{E} = \frac{\hbar^2 k_n^2(\mathcal{E})}{2m_n(\mathcal{E})} \quad (3.11)$$

defines $m_n(\mathcal{E})$. Since the transition rate is nonzero for intraband scattering only

$$P_{NM} \equiv \frac{1}{2} \sum_{n(N)} \sum_{m(M)} P_{nm} = P_{nn} \delta_{NM}. \quad (3.12)$$

Manipulating with the collision term of the Boltzmann equation

$$\frac{\Omega}{(2\pi)^3} \int d\mathbf{k}' P_{NN}(\mathbf{k}, \mathbf{k}') [f_N(\mathbf{k}') - f_N(\mathbf{k})], \quad (3.13)$$

using the standard substitutions for f_N , Eqs. (2.35), gives the system of linear equations as before. The scattering matrix is diagonal in the band and angular momentum indices

$$S_{N\lambda'', M\lambda'}(\mathcal{E}, T)_{\text{neut}} = - \frac{320\pi^2 \hbar^4 \Gamma_N(\mathcal{E})^2 a_0 N_N}{m_N(\mathcal{E})^2 k_N(k_B T)} \mathcal{O}_0^{NM}(\mathcal{E}) \delta_{NM} \delta_{\lambda''\lambda'}, \quad (3.14)$$

where N_N is the density of neutral impurities. Note that this formalism decouples all three bands so that only intraband scattering is considered. The bands are taken to be isotropic yet nonparabolic. Also, the present treatment extends beyond the relaxation-time approximation since the angular momenta $\lambda=1, 3, 5, 7, 7'$ are used in this part of the calculation as well. Using the conversion from m_N to k_N , Eq. (3.11), and a_0 to E_B , Eq. (3.5), results in

$$S_{N\lambda'', M\lambda'}(\mathcal{E}, T)_{\text{neut}} = - \frac{640\pi^2 \Gamma_N(\mathcal{E})^2 e^2 N_N \mathcal{E}^2}{k_N^5 \epsilon E_B(k_B T)} \mathcal{O}_0^{NM}(\mathcal{E}) \delta_{NM} \delta_{\lambda''\lambda'}. \quad (3.15)$$

B. Ralph's formulation of the neutral-impurity scattering

An approach to neutral-impurity scattering based on scattering by a spherically symmetric square-well potential was originated by Anselm¹⁹ and Sclar.²⁰ The emphasis of this formulation is the role of the weakly bound hydrogenic-impurity ion for a single spherical-parabolic band. A numerical comparison between Sclar's, Erginsoy's, and the improved phase-shift calculation by McGill and Baron shows that the three calculations are very close except at very low temperatures, $T < 20$ K, or so.²¹ For completeness sake, I have also tried the parameterized expressions of Meyer and Bartoli⁶⁶ and have not found much difference between their results and those using Erginsoy's formulation.

In the case of scattering from a localized scattering center there is formulation which takes into account the multiband picture. In a largely unnoticed piece of work Ralph³³ has done a very comprehensive job in elucidating the physics and mathematics of scattering between coupled-band manifolds. The model exhibits the essential coupling between bands in the scattering process. It also avoids the arbitrariness with which effective masses of all sorts have been introduced by previous authors.

Following Ralph's formalism I get for the scattering rates

$$P_{MN}(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} \frac{N_N}{\Omega} |T_0|^2 \delta(E_N(\mathbf{k}) - E_M(\mathbf{k}')) \times \begin{cases} 1 - \frac{3}{4} \sin^2 \theta, & N=M \\ \frac{3}{4} \sin^2 \theta, & N \neq M \end{cases} \quad (3.16)$$

for scattering among the light- and heavy-hole manifolds in the spherical-parabolic band approximation. At temperatures where the spin-orbit band is appreciably occupied the number of neutral centers is very small so that this band may be neglected. Also,

$$|T_0|^2 = \frac{(4\pi)^2}{[E_N + \mathcal{E}]} \frac{4\mathcal{E}^3}{[k_L^3 + k_H^3]^2}, \quad (3.17)$$

where E_N is the binding energy of the second hole, and

$$\mathcal{E} = \frac{\hbar^2 k_L^2}{2m_L} = \frac{\hbar^2 k_H^2}{2m_H} \quad (3.18)$$

define the radii of the spherical constant energy surfaces for the light-, k_L , and heavy-hole, k_H , bands. The denominator in Eq. (3.17), $[k_L^3 + k_H^3]^{-2}$, exhibits the essential coupled-band nature of the problem. Apparently, Eq. (3.17) indicates that the proper mass to use is

$$(m^*)^3 = m_L^3 + m_H^3. \quad (3.19)$$

It also should be noted that the angular factors in Eq. (3.16) give rise to an angular dependence in P_{MN} . These factors arise from the overlap between the initial- and final-hole wave functions. At this stage I propose to replace Eq. (3.16) by its solid state analog, that is,

$$\begin{aligned} k_L &\rightarrow k_L(\mathcal{E}), \\ k_H &\rightarrow k_H(\mathcal{E}), \\ 1 - \frac{3}{4} \sin^2 \theta &\rightarrow O_{NM}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E}), \quad N=M \\ \frac{3}{4} \sin^2 \theta &\rightarrow O_{NM}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E}), \quad N \neq M. \end{aligned} \quad (3.20)$$

Also, E_N is on the order of 3–4 meV which is much less than hole energies of interest for $T > 20$ K. Thus to a good approximation $E_N \simeq 0$. With these substitutions the transition rate becomes

$$P_{MN}(\mathbf{k}, \mathbf{k}') = \frac{N_N}{\Omega} \frac{2\pi}{\hbar} \frac{4(4\pi)^2 \mathcal{E}^2}{[k_L(\mathcal{E})^3 + k_H(\mathcal{E})^3]^2} \times O_{NM}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E}) \delta(E_N(\mathbf{k}) - E_M(\mathbf{k}')). \quad (3.21)$$

This form contains the main consequences of nonparabolicity but neglects several important aspects of the band anisotropy. Admittedly, the substitutions, Eq. (3.20) are heuristic in nature. It is not apparent at this time how the full nonparabolic anisotropic nature of the valence bands can be included in the scattering problem.

Following the procedure outlined previously the scattering matrix becomes

$$S_{N\lambda'', M\lambda'} = -\delta_{\lambda''\lambda'} \delta_{NM} \frac{(16\pi)^2}{k_B T} \frac{N_N \mathcal{E}^2 \Gamma_N(\mathcal{E})}{[k_L(\mathcal{E})^3 + k_H(\mathcal{E})^3]^2} \times \left[\sum_{M=L,H} \Gamma_M(\mathcal{E}) \mathcal{O}_0^{NM}(\mathcal{E}) \right], \quad N, M = L, H \quad (3.22)$$

where the sum over M proceeds only over the light- and heavy-hole bands and only the $L=0$ component of

$O_{NM}(\hat{\mathbf{k}}, \hat{\mathbf{k}}', \mathcal{E})$ expansion is projected out.

It will be shown shortly that the theory outlined above produces Hall mobilities whose shape as a function of temperature agrees with experiment, unlike the Erginsoy theory. For quantitative agreement it has been found necessary to employ an adjustable parameter multiplying Eq. (3.22). There are several reasons why the theory, as advanced as it is, may still be lacking here. There are of course various simplifications regarding the band shapes and anisotropic effects. The solid-state analog argument is obviously an *ad hoc* procedure. Yet, I regard these as minor approximations compared to the assumption of a δ function for the scattering potential³³

$$V(\mathbf{r}, \mathbf{r}') = V_0 \delta(\mathbf{r}) \delta(\mathbf{r}') . \quad (3.23)$$

This assumption neglects the structure of the scattering center. Here, there is a real need for a calculation using phase-shift theories encompassing the multiband nature of the valence bands. The adjustable parameter will be found to be different for each acceptor.

In anticipation of numerical results in Sec. IV I develop for comparison the behavior of the Erginsoy and Ralph theories in the limit of a two spherical-parabolic band. Using Eqs. (2.41) and (2.42) with Erginsoy's theory gives

$$S_{NN} \sim N_N (m_N \mathcal{E})^{1/2} / T , \quad (3.24)$$

$$\mu_c, \mu_H \sim 1 / N_N , \quad (3.25)$$

and for a single noninteracting band

$$r_N = 1 . \quad (3.26)$$

Similarly, for Ralph's theory for $\mathcal{E} \gg E_N$

$$S_{NN} \sim (N_N / T) \frac{m_N^{3/2}}{(m_L^{3/2} + m_H^{3/2})} , \quad (3.27)$$

$$\mu_c, \mu_H \sim (T^{1/2} / N_N) , \quad (3.28)$$

and for a single noninteracting band

$$r_N = 45\pi / 128 = 1.1044 . \quad (3.29)$$

Apparently, in Erginsoy's theory the ratio of the light- to heavy-hole scattering efficiency is

$$S_{LL} / S_{HH} \sim (m_L / m_H)^{1/2} \sim 0.53 , \quad (3.30)$$

while in Ralph's theory it is

$$S_{LL} / S_{HH} \sim (m_L / m_H)^{3/2} \sim 0.15 , \quad (3.31)$$

so that in the latter theory light holes are scattered 3.5 times less than in the former theory. These relations will serve merely as guides for further discussion. The true picture is complicated by nonparabolicity, anisotropy, band coupling, and the inapplicability of the Matthiessen's rule.

In ending this review of the neutral-impurity scattering theories it should be observed that ultimately in semiconductor applications the effective-mass approximation (EMA) is invoked. As noted earlier, EMA works best for sufficiently gentle potentials. It may be that the short-range potentials of the neutral impurities are too abrupt for the validity of EMA to hold.

IV. RESULTS

Before exhibiting results for real samples a short discussion of model systems will be presented. This will have the advantage of tracking the influence of the ionized- and neutral-impurity scattering separately.

A. Model system for the study of the ionized- and the neutral-impurity scattering

The theory of Brooks-Herring as modified in Sec. II was applied to the ideal system where there is only one fully ionized acceptor level and there is no compensation, i.e.,

$$N_A = p , \quad (4.1)$$

$$L_D^{-2} = (4\pi e^2 / \epsilon k_B T) N_A . \quad (4.2)$$

I have calculated the Hall and conductivity mobilities as a function of temperature for different values of N_A , including both phonon scattering mechanisms. When $N_A = 10^{16}$ the characteristic turnover in the mobility is seen since the ionized-impurity limited mobility increases with temperature. Figure 2 shows the calculated r factors for several acceptor concentrations. The lowest concentration shown on the graph is representative of a sample with almost no ionized-impurity scattering and is indicative of the Hall factors for the two-phonon scattering mechanisms. As the ionized-impurity concentration increases the r factor for the light-hole band generally decreases and for the heavy-hole band increases. The effect of the ionized-impurity scattering is strongest, of course, at the lowest temperatures with the phonon scattering dominating the r -factor behavior at higher temperatures. Apparently the light-hole band is affected more than the heavy-hole band by the presence of the ionized-impurity scattering. With increasing N_A , the r factors for both bands in Fig. 2 tend to a limit determined by the ionized-

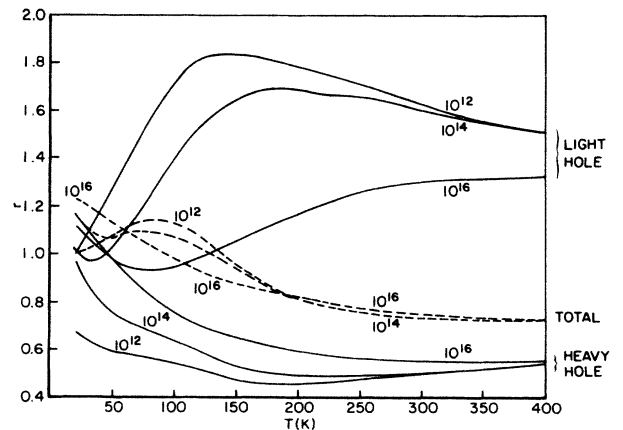


FIG. 2. Calculated Hall factors from calculated mobilities including the acoustic- and optical-phonon scattering with the Brooks-Herring ionized-impurity formalism. The r factor for the heavy- and light-hole bands are shown as well as the total r factor. The numbers next to each curve are the total number of ionized centers (cm^{-3}).

impurity scattering alone. Owing to the complicated energy dependence in the scattering matrix, Eq. (2.43), that limit is not $315\pi/512 = 1.9328$ as it is in simpler theories. As shown by Blatt³⁴ for both the partial wave and Born calculation, the r factor for a single spherical-parabolic band rises from about 1.1–1.2 at 20 K and saturates in value at about 1.7 above 100 K. The same conclusion was reached independently by Mansfield.⁷⁴ The present calculation confirms this trend which in large part is determined by the energy dependence of the scattering rates. Given the r factor behavior for ionized scattering alone, it is natural then that the r_H , being on the order of 0.5 for $N_A = 10^{12} \text{ cm}^{-3}$, will rise toward those values and r_L , being generally high for $N_A = 10^{12} \text{ cm}^{-3}$, will decrease toward the same limit. The mixing of the phonon- and ionized-impurity scattering in the r factor may not be a monotonic function of the degree of each type of scattering. As shown by Mansfield⁷⁴ it is possible for the r factor for the combined scattering to be outside the values for either scattering mechanism alone.

The total r factor in Fig. 2 is influenced more by r_H than by r_L since the weighting for r_N in Eq. (2.52) is p_N , the partial band occupancy. It is known that the heavy-hole band's carrier-concentration effective mass increases by a factor of 2 from low to high temperatures and its capacity to hold the carriers, that is p_N , also increases accordingly.⁷⁵ At these temperatures the weighting by p_N offsets the advantage that the light-hole band has in the weighting by virtue of its higher conductivity mobility.

The next model system incorporated both phonon scattering mechanisms and the neutral-impurity scattering using Ralph's formalism as developed in Sec. III. The model system this time has a constant density of neutral centers N_N . The neutral-impurity scattering appears important for N_N greater than 10^{15} cm^{-3} . The low-temperature mobility becomes proportional to N_N^{-1} for neutral-impurity concentrations above mid 10^{17} cm^{-3} . The neutral-impurity scattering using Ralph's formalism produces a turnover in μ_H at low temperatures for heavier doping. Flatness, if any, in μ_H versus T is produced only as a result of the competition between the neutral- and the acoustic-phonon scattering mechanisms.

Figure 3 shows that the effect of the neutral-impurity scattering, when combined with the acoustic- and optical-phonon scattering, is to raise the total r factor at low temperatures and lower it at intermediate temperatures. The behavior observed for r , r_L , and r_H in Fig. 3 is remarkably similar to that found in Fig. 2 for the effect of the ionized-impurity scattering. The explanation here is similar to that offered above. With increasing N_N , the partial r factors r_N for each band attempt to reach the limit indicative of pure neutral-impurity scattering for single noninteracting bands. That limit is on the order of unity, Eq. (3.25) and (3.29), for most models of the neutral-impurity scattering. This requires r_H to rise and r_L to fall in value, as N_N increases, from values r_L and r_H had for $N_N = 0$. This effect is strongest at low temperatures where the neutral-impurity scattering is strong. The total r factor is again dominated by r_H owing to the higher heavy-hole band concentration. Lastly, it is expected that the r factors for Erginsoy's theory should be lower than

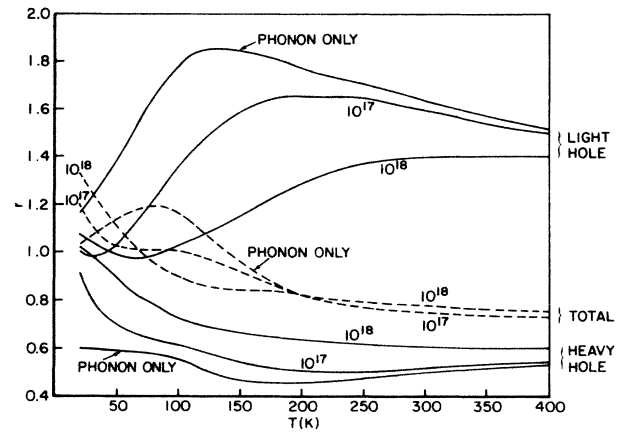


FIG. 3. Calculated Hall factors from calculated mobilities including the acoustic- and optical-phonon scattering with the Ralph neutral-impurity scattering formalism. The r factor for the heavy- and light-hole bands are shown as well as the total r factor. The numbers next to each curve are the total number of neutral centers (cm^{-3}).

for Ralph's since the light holes in Erginsoy's theory are scattered 3.5 times stronger in that theory than in Ralph's. Explicit calculations presented next confirm that prediction.

B. Comparison of theory with experiment

The theory of Sec. III has been applied to the calculation of mobilities of boron-, aluminum-, gallium-, and indium-doped silicon samples for which data could be gathered. The initial estimates of dopant densities were performed with an empirical r factor and a χ^2 fitting routine which fits the charge-balance equation to the carrier concentration versus temperature data.⁷⁶ The experimental setup and procedure are explained by Mitchel and Hemenger,⁷⁷ and Hemenger.⁷⁸

Figures 4(a)–4(d) display the calculated hall mobility for a representative set of boron-, aluminum-, indium-, and gallium-doped silicon samples together with the experimental data. In each graph three cases are considered: (1) phonon and the Brooks-Herring scattering; (2) phonon, BH, and the neutral-impurity scattering employing Ralph's theory; (3) phonon, BH, and Erginsoy's neutral-impurity scattering. An adjustable parameter multiplies the strength of Erginsoy's and Ralph's neutral-impurity scattering to bring the theory and experiment into agreement at low temperatures. More will be said about this later.

For lower doped samples $N_A < 10^{15} \text{ cm}^{-3}$, examined in the course of research, the ionized-impurity scattering, together with the phonon scattering, account very well for the magnitude and shape of the mobility curve. For samples in Figs. 4(a)–4(d) the acceptor concentrations are above 10^{15} cm^{-3} and the neutral-impurity scattering starts to rival the strength of the other scattering mechanisms as evidenced by the separation between the curves with and without the neutral-impurity scattering. For

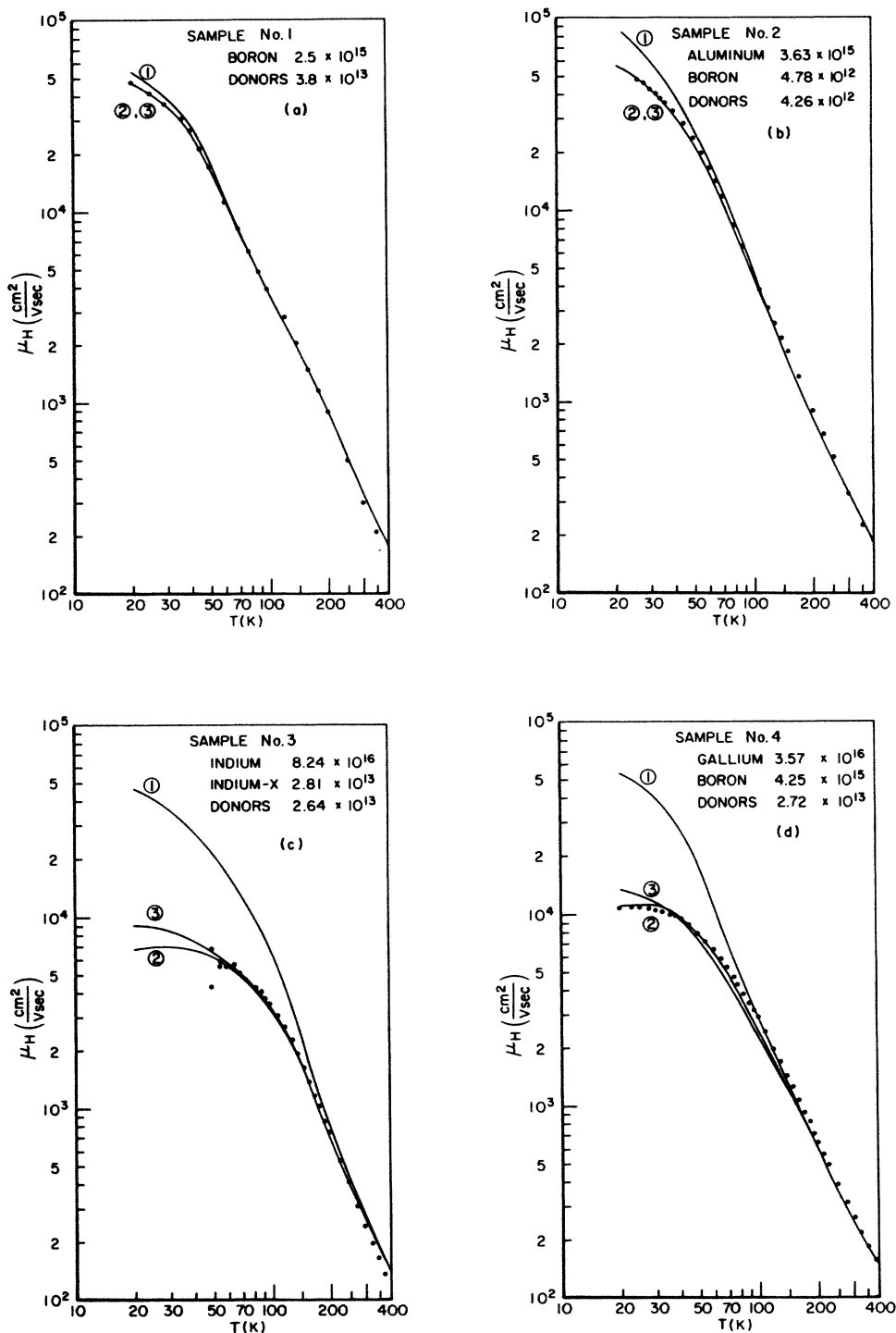


FIG. 4. Comparison of the calculated (solid curve) Hall mobility and data (circles) for (a) boron-, (b) aluminum-, (c) indium-, and (d) gallium-doped silicon samples. Curve 1 is calculated using both the acoustic- and optical-phonon scattering and the BH theory of ionized-impurity scattering. Curve 2 in addition uses Ralph's neutral-impurity scattering formalism while curve 3 uses Erginsoy's formalism.

four boron samples examined in the course of research the factor multiplying the strength of the neutral-impurity scattering was found to be between 2 to 2.8 for Ralph's formulation and 0.3 to 0.5 for Erginsoy's theory. The adjustable parameters for the strength of the neutral-impurity scattering were 3.5–4.5 for Ralph's formula-

tion, 1–1.5 for Erginsoy's formulation, for two aluminum samples. For two indium-doped samples the adjustable factors were between 3.5–5.5 for Ralph's theory and 1.9–2.9 for Erginsoy's. The adjustable parameters for the strength of the neutral-impurity scattering for four gallium-doped samples were between 5.5–6.0 for Ralph's

theory and between 1.8–2.0 for Erginsoy's theory. These numbers are based on data for a limited number of samples and are functions of the quality of data and initial χ^2 fits as these are inputs into the present calculation. The limited range of data for the indium-doped samples, i.e., lack of data in the freeze-out and exhaustion regions, makes the theoretical-experimental comparison less certain. In none of the samples examined did the variation of E_N in Eq. (3.17) make much difference in the value of calculated mobilities in the temperature range considered. Variations of E_N change mobilities by few percent and not by factors of 2 to 6 as is required. It is rather remarkable that the adjustable factors turn out to be rather constant for the given type of impurity implying that there is some physics underlying the adjustment.

The data displayed in Figs. 4(a)–4(d) show clearly that very good agreement with experiment has been reached when the BH theory and the phonon scattering are combined *without* any adjustable parameters. This is true as far as the magnitude and shape of the mobility curves are concerned. The fine undulations in Fig. 4(a) are due to the varying degree of ionization of boron as a function of temperature. At the lowest of temperatures $p = 0$, so that

$$L_D^{-2} \sim N_D(1 - N_D/N_A)/T, \quad (4.3)$$

and L_D increases with temperature. As temperature rises, p increases as holes are promoted to the valence band with p rising exponentially. As the consequence the free-carrier screening becomes effective and L_D drops in value. On further increase in temperature the exhaustion region is reached and

$$p_s = p = N_A - N_D \quad (4.4)$$

and

$$L_D^{-2} \sim T^{-1} \quad (4.5)$$

again, so that L_D increases with temperature. The three temperature regimes find their reflections in the mobility. Given the adjustment of the strength of the neutral-impurity scattering there is very little that differentiates between the Ralph and Erginsoy theory in Fig. 4(a). The neutral-impurity scattering using Sclar's theory,²⁰ and also the theory of Meyer and Bartoli,⁶⁶ were tried as well and found to be practically indistinguishable from the theory of Erginsoy, as was also found to be the case by McGill and Baron.²¹ The Matthiessen's rule was found to be totally inappropriate.

Figures 4(b) and 4(c) show the comparison of theory with experiment for aluminum- and an indium-doped samples with doping exceeding 10^{15} and 10^{16} cm⁻³, respectively. In the higher temperature regions the combination of the ionized-impurity and optical-phonon scattering account well for both the shape and magnitude of the data. The influence of the neutral-impurity scattering is more visible here since aluminum and especially indium have large ionization energies so that these impurities remain neutral up to higher temperatures. The fine undulations in the mobilities seen in Fig. 4(a) are not seen here since by the time the exhaustion region is reached the ionized-impurity scattering is not as important as it is at lower temperatures. There is an important difference be-

tween the Ralph and Erginsoy theories which is apparent in Fig. 4(c) for indium-doped silicon. Unfortunately, that difference arises in the freeze-out region, $T < 50$ K, where no accurate data could be taken. It is quite possible that the indium samples contain boron which is overcompensated so that the χ^2 fit would not reveal its presence.

Figure 4(d), on the other hand, is representative of a large selection of data available to the author for gallium-doped silicon (mostly mid- 10^{16} cm⁻³ region, though). The need for the neutral-impurity scattering is self-evident in Fig. 4(d). The agreement between experiment and theory for lower-doped gallium samples indicated that the phonon and BH scattering are well understood. On the basis of these results, it can now be said that scaling of the scattering cross section using the experimental E_B , Eq. (3.4), is not enough to account for the strength of Erginsoy's neutral-impurity scattering.

Most serious is the discrepancy between the Erginsoy theory and experiment at low temperatures in Fig. 4(d) and other similarly doped gallium samples. The calculated mobility does not bend over as required by the experiment. Given that Erginsoy's theory yields temperature-independent mobility this result may seem surprising, at first. A large amount of numerical "experimentation" has shown that Matthiessen's rule is not obeyed and the mobility is the result of competition between the four scattering mechanisms considered here. The knowledge of the temperature dependence of each mobility mechanism alone is but a rough guide of how the total mobility will behave with all the mechanisms present all at once. Higher doped, mid- 10^{17} cm⁻³ and up, silicon gallium data⁷⁹ in fact shows the mobility bending over at low temperatures and not merely flattening out. (At still lower temperatures there is of course a precipitous drop in the mobility as the hopping mobility regime is reached. The bend over discussed above is not related to this large drop in the mobility.)

It has been shown by McGill and Baron²¹ that there is little that distinguishes the theories of Sclar, Erginsoy, and McGill and Baron for neutral-impurity scattering involving a single band. Sclar's formulation is the form that Ralph's theory assumes in the limit of noninteracting bands. When I applied Sclar's theory treating the light- and heavy-hole bands as decoupled indeed the results were close to the predictions of Erginsoy's theory in Figs. 4(a)–4(d). Therefore, it is the coupled nature of the valence bands in Ralph's theory that accounts for the correct temperature dependence of the calculated mobility. The need for adjustment of the strength of neutral scattering is common to all calculations to date. In the case of Ralph's theory the discrepancy may be connected with the zero-potential range assumption and possible breakdown of the effective-mass formulation for rapidly varying potentials.

C. Use of the calculated r factors in the Hall data fits

The ultimate utility of the calculated r factors is in fitting the carrier concentration data to the charge-balance equation in order to extract the dopant densities and activation energies.¹⁰ The hole concentration is given by

$$p = r/qR_H, \quad (4.6)$$

where the Hall coefficient R_H is measured experimentally. The carrier concentration is then used in the charge-balance equation

$$p + N_D = \sum_i N_{A_i}^-, \quad (4.7)$$

where the sum proceeds over all acceptors. The density-of-states effective masses of Madarasz, Lang, and Hemenger^{75,80} were used to calculate p in Eq. (4.6).

The calculated r factors for samples in Figs. 4(a)–4(d) are shown in Figs. 5(a)–5(d). In each graph four cases are exhibited: (1) phonon scattering only, (2) phonon and BH scattering, (3) phonon, BH, and Ralph's neutral scattering, (4) phonon, BH, and Erginsoy's neutral scattering. The main effect of the ionized-impurity scattering is to raise the r factor in the freeze-out region, $p \approx 0$, and lower it as the temperature is increased. The inclusion of the neutral-impurity scattering further lowers the r factor for Erginsoy's theory for the whole temperature range while in Ralph's theory at the lowest of temperatures the r factor increases over the value for phonon and BH scattering only. These trends were explained in Sec. IV A. Here it ought to be added that owing to the stronger scattering of

light holes in Erginsoy's theory, see Eqs. (3.30) and (3.31), the r factor is diminished from that in Ralph's theory since, as discussed in Sec. IV A, light holes increase the r factor while the heavy holes have the opposite effect. For more heavily doped samples, the differences between the two neutral-impurity scattering theories in Fig. 5(a)–5(d) are large enough to make a difference in the quality of the subsequent χ^2 fits.

The carrier-concentration data was multiplied by the calculated r factors and then used in the fit to the charge-balance equation, Eq. (4.7). Table I provides a quantitative measure for the goodness of fit using the calculated r factors and including $r = 1$ for comparison. It is assumed that R_H is determined with experimental accuracy of 3%. For boron-doped samples the χ^2 for both theories are comparable since the calculated r factors differ little for samples used in this study. These fits are generally better than the fit with $r = 1$ as far as the calculated χ^2 and when one considers the calculated activation energies. The accepted values of activation energies of acceptors in silicon are 44.39, 69.03, 72.73, and 155.58 meV for boron, aluminum, gallium, and indium.⁸¹ Lipari *et al.*⁸² gives activation energies about 1.5 meV larger.

The comparison for aluminum-doped samples also gives good χ^2 and activation energies using the calculated

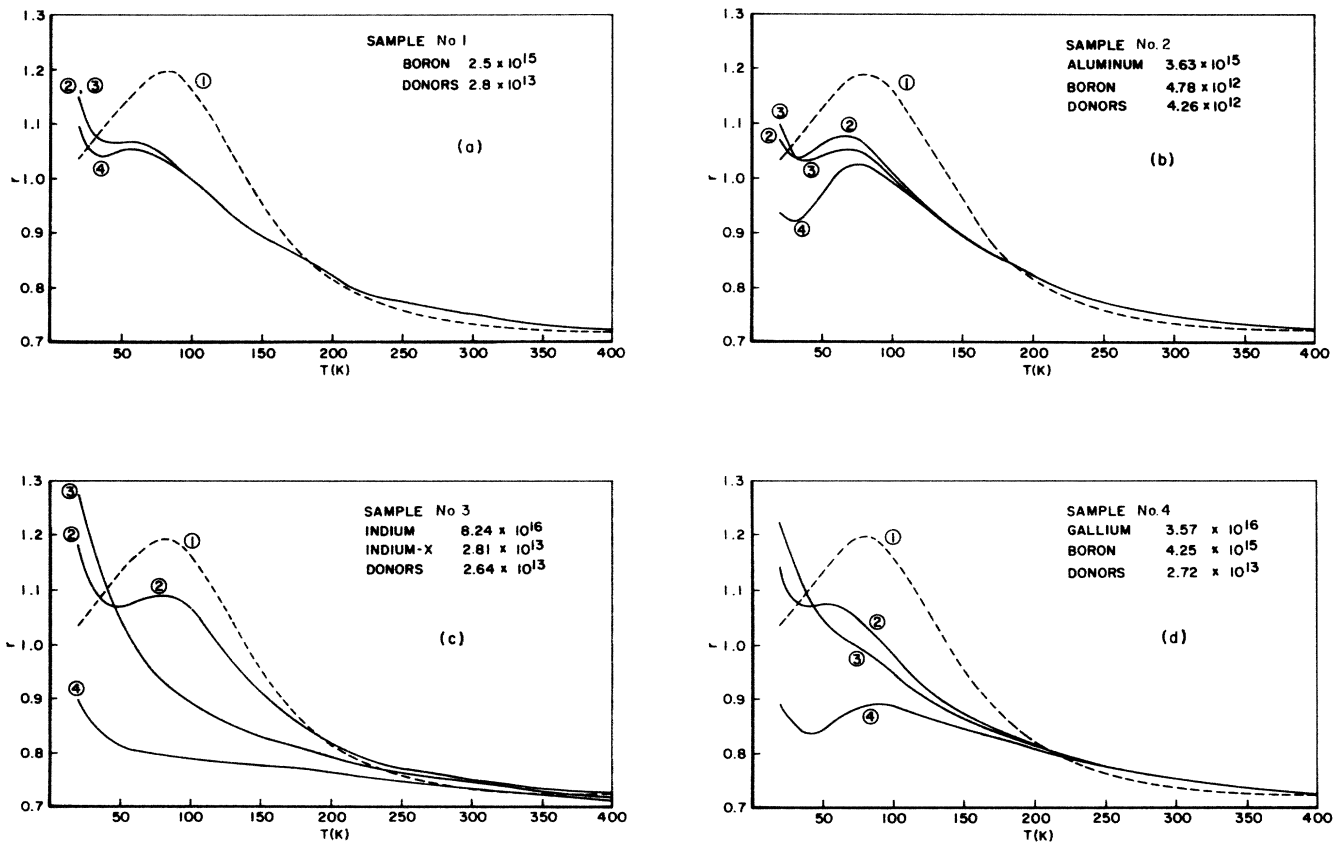


FIG. 5. Calculated Hall factors for the samples in Fig. 4. Curves 1–4 are as follows: (1) optical and acoustic phonon only, (2) phonon + BH ionized-impurity scattering, (3) phonon, BH, and Ralph's neutral-impurity scattering, (4) phonon, BH, and Erginsoy's neutral-impurity scattering.

TABLE I. Comparison of χ^2 fits to the data for doped silicon samples using calculated r factors and $r = 1$. The concentrations of acceptors and donors are in cm^{-3} and the activation energies are in meV.

Sample		$r = 1$	Erginsoy	Ralph
No. 1 (119-757-1037)	N_B	2.79×10^{15}	2.42×10^{15}	2.42×10^{15}
	N_D	2.11×10^{13}	2.88×10^{13}	2.88×10^{13}
	E_B	45.13	43.89	43.89
	χ^2	5.11	1.37	1.37
No. 2 (188-458-1087)	N_{Al}	4.61×10^{15}	3.67×10^{15}	3.65×10^{15}
	N_B	3.13×10^{12}	5.50×10^{12}	5.15×10^{12}
	N_D	2.81×10^{12}	5.20×10^{12}	4.78×10^{12}
	E_{Al}	72.96	71.40	70.99
	E_B	44.08	43.11	43.36
	χ^2	1.15	0.30	0.17
No. 3 (109-285-0957)	N_{In}	1.33×10^{17}	7.99×10^{16}	7.73×10^{16}
	N_{In-X}	2.42×10^{13}	1.89×10^{13}	2.05×10^{13}
	N_D	2.38×10^{13}	1.84×10^{13}	1.99×10^{13}
	E_{In}	168.60	168.54	166.39
	E_{In-X}	112.67	113.96	113.46
	χ^2	0.881	0.902	1.061
No. 4 (0164-432-0724)	N_{Ga}	5.28×10^{16}	3.80×10^{16}	3.75×10^{16}
	N_B	3.01×10^{15}	2.11×10^{15}	3.00×10^{15}
	N_D	2.58×10^{13}	2.36×10^{13}	2.54×10^{13}
	E_{Ga}	74.15	72.31	72.79
	E_B	44.85	44.58	44.49
	χ^2	0.453	0.882	0.373

r factors. The $r = 1$ assumption leads to activation energies higher than experimentally accepted. There is a long-standing puzzle regarding the degeneracy factor for indium. The theoretically accepted value for the ground-state degeneracy for column III acceptors in silicon is 4. It is also true that the degeneracy factor in the charge-balance equation takes into account all the excited states of the impurity. The position of odd-parity excited states is well known from optical experiments.^{81,82} Their inclusion in the “effective” degeneracy factor does not influence the fit much, except for very heavily doped samples, owing to the proximity of these states to the valence-band edge.¹⁰ Less is known about the even-parity states.^{82,83,84} Experiment and theory place Γ_8^+ states close to the valence-band edge again,⁸² but there is evidence of a Raman-active twofold degenerate Γ_7^+ level in Si:B 23.8 meV above the ground state. To date there has been no observation of a similar level for other column III acceptors in silicon. Nevertheless, Hall-data fits for indium-doped silicon give activation energies closer to the optically measured values for the degeneracy of 6 rather than 4.⁸⁵

To complicate matters further there is evidence that there is an excited level of neutral indium acceptor about 4 meV above the ground state.^{86,87,88} The consensus appears to attribute the level to the dynamic Jahn-Teller effect⁸⁹ caused by the strong vibronic coupling with the highly localized hole wave function on indium in silicon. The X level in indium^{90–93} has been identified as an indium-carbon complex with activation energy of 112.8

meV. Like a similar Al- X level⁹⁴ in aluminum-doped silicon, it has trigonal symmetry and its ground-state degeneracy is therefore 2.

Table I gives the fit for the gallium-doped sample. From the χ^2 one sees that the Ralph theory is better than Erginsoy's theory and produces activation energies closer to the experimental values. The choice of $r = 1$ consistently overestimates the activation energies. It is true that parameters for residual acceptors, such as boron in Si:Ga, will be less well determined from a fit than parameters for the major dopant. The use of $r = 1$ leads to major dopant concentrations as much as 50–80 % off from densities obtained using the theoretically determined r factors.

V. SUMMARY AND CONCLUSIONS

The conductivity and Hall mobilities for doped p -type silicon have been calculated with a new theory which combines the acoustic-phonon, optical-phonon, ionized-impurity, and neutral-impurity scattering. The effect of band nonparabolicity has been taken into account exactly in the treatment of the scattering mechanisms and in the solution of the Boltzmann equation. The band anisotropy was incorporated in the calculation of the phonon scattering rates and the major effects of band anisotropy were incorporated in the treatment of the ionized- and neutral-impurity scattering. The present treatment obviates the need for the relaxation-time approximation.

The comparison between experiment for four dopants in silicon and theory leads to several conclusions. The Brooks-Herring theory of ionized-impurity scattering, as modified in Sec. II, yields good agreement with experiment, without any adjustable parameters, for the doping concentrations and temperatures considered.

The state of neutral-impurity scattering theories is a less well-settled matter. The comparison between Ralph's and Erginsoy's formulations, using mobilities and carrier-concentration data fits as criteria, shows that Ralph's approach is better. The advantage is traced to the inclusion of the coupled valence-band nature of the problem in Ralph's formulation. The need for adjusting the strength of the neutral-impurity scattering in either theory

is a glaring deficiency in our state of knowledge of scattering mechanisms.

ACKNOWLEDGMENTS

This work was performed under U.S. Air Force Contracts No. F33615-81-C-5095 and No. F33615-85-C-5062. A debt is owed to Dr. P. M. Hemenger and Mr. John A. Detrio for their support and encouragement. I would also like to thank Dr. Frank L. Madarasz, Dr. J. E. Lang, Dr. W. Mitchel, Mr. Steven R. Smith, and Mr. Timothy Peterson for many helpful discussions. Mr. Kenneth Bloch, Mr. Timothy Peterson, and Mr. Kenneth Beasley performed the χ^2 fits.

- ¹B. R. Nag, *Electron Transport in Semiconductors*, Vol. 11 of *Springer Series in Solid State Sciences* (Springer-Verlag, New York, 1980).
- ²B. K. Ridley, *Quantum Processes in Semiconductors* (Clarendon, Oxford, 1982).
- ³J. D. Wiley, in *Transport Phenomena*, Vol. 10 of *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1975).
- ⁴F. L. Madarasz and F. Szmulowicz, *Phys. Rev. B* **24**, 4611 (1981).
- ⁵F. Szmulowicz and F. L. Madarasz, *Phys. Rev. B* **26**, 2101 (1982).
- ⁶F. Szmulowicz and F. L. Madarasz, *Phys. Rev. B* **27**, 2605 (1982).
- ⁷F. Szmulowicz and F. L. Madarasz, *Phys. Rev. B* **27**, 6279 (1983).
- ⁸F. Szmulowicz, *Phys. Rev. B* **28**, 5943 (1983); *Appl. Phys. Lett.* **43**, 485 (1983).
- ⁹W. C. Mitchel and P. M. Hemenger, *J. Appl. Phys.* **53**, 6880 (1982).
- ¹⁰J. S. Blakemore, *Semiconductor Statistics* (Pergamon, New York, 1962).
- ¹¹D. Chattopadhyay and H. J. Queisser, *Rev. Mod. Phys.* **53**, 745 (1981).
- ¹²E. E. Haller, Robert E. McMurray, Jr., L. M. Falicov, N. M. Haegel, and W. L. Hansen, *Phys. Rev. Lett.* **51**, 1089 (1983).
- ¹³W. Burger and K. Lassman, *Phys. Rev. Lett.* **53**, 2035 (1984).
- ¹⁴N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Clarendon, Oxford, 1965).
- ¹⁵G. F. Drukarev, *The Theory of Electron-Atom Collisions* (Academic, New York, 1965).
- ¹⁶Leonard S. Rodberg and R. M. Thaler, *Introduction to the Quantum Theory of Scattering* (Academic, New York, 1967).
- ¹⁷Sydney Geltman, *Topics in Atomic Collision Theory* (Academic, New York, 1969).
- ¹⁸C. Erginsoy, *Phys. Rev.* **79**, 1013 (1950).
- ¹⁹A. I. Anselm, *Zh. Eksp. Teor. Fiz.* **24**, 85 (1953).
- ²⁰N. Sclar, *Phys. Rev.* **104**, 1559 (1956).
- ²¹T. C. McGill and R. Baron, *Phys. Rev. B* **11**, 5208 (1975).
- ²²H. S. W. Massey and B. L. Moiseiwitsch, *Proc. R. Soc., London, Ser. A* **135**, 483 (1951); *Phys. Rev.* **78**, 180 (1950).
- ²³A. Temkin and J. C. Lamkin, *Phys. Rev.* **121**, 788 (1961).
- ²⁴T. T. Braggins, Ph.D. thesis, Syracuse University, 1975 (unpublished) [available from University Microfilms, Ann Arbor, MI (1975)].
- ²⁵P. Norton, T. Braggins, and H. Levinstein, *Phys. Rev. B* **8**, 5632 (1973).
- ²⁶Eiji Ohta and Makoto Sakata, *Jpn. J. Appl. Phys.* **17**, 1795 (1978).
- ²⁷R. Baron, M. H. Young, and T. C. McGill, *Physics of Semiconductors*, Proceedings of the 13th International Conference on the Physics of Semiconductors, Rome, 1976, edited by F. G. Fumi (North-Holland, New York, 1976), p. 1158.
- ²⁸Sheng S. Li, *Solid-State Electron.* **21**, 1109 (1978).
- ²⁹K. Takeda, K. Sakui, A. Taguchi, and M. Sakata, *J. Phys. C* **16**, 729 (1983).
- ³⁰J. F. Lin, S. S. Li, L. C. Linares, and K. W. Teng, *Solid-State Electron.* **24**, 827 (1981).
- ³¹D. A. Tjapkin, T. I. Tošić, and M. M. Jevtić, *Solid-State Electron.* **27**, 667 (1984). This calculation was performed at 300 K only.
- ³²H. Nakagawa and S. Zukotynski, *Can. J. Phys.* **56**, 364 (1977).
- ³³H. I. Ralph, *Philips Res. Rep.* **32**, 160 (1977).
- ³⁴F. J. Blatt, *J. Phys. Chem. Solids* **1**, 262 (1957).
- ³⁵A. D. Boardmann and D. W. Henry, *Phys. Status Solidi B* **60**, 633 (1974).
- ³⁶J. R. Meyer and F. J. Bartoli, *Phys. Rev. B* **23**, 5413 (1981).
- ³⁷H. Brooks, *Phys. Rev.* **83**, 879 (1951); *Advances in Electronics and Electron Physics*, edited by L. Morton (Academic, New York, 1955), Vol. 7, p. 85.
- ³⁸C. Herring (unpublished).
- ³⁹L. M. Falicov and M. Cuevas, *Phys. Rev.* **169**, 1025 (1967).
- ⁴⁰J. B. Krieger and S. Strauss, *Phys. Rev.* **164**, 674, (1968).
- ⁴¹F. Stern, *Phys. Rev.* **158**, 697 (1967).
- ⁴²J. Friedel, *Adv. Phys.* **3**, 446 (1954).
- ⁴³D. L. Rode and S. Knight, *Phys. Rev. B* **3**, 2534 (1971).
- ⁴⁴J. R. Meyer and F. J. Bartoli, *J. Phys. C* **15**, 1987 (1982).
- ⁴⁵J. R. Meyer and F. J. Bartoli, *Phys. Rev. B* **31**, 2353 (1985).
- ⁴⁶J. R. Meyer and F. J. Bartoli, *Phys. Rev. B* **30**, 1026 (1984).
- ⁴⁷P. Csavinszky, *Phys. Rev. B* **14**, 1649 (1976).
- ⁴⁸P. Csavinszky, *Int. J. Quantum Chem.* **13**, 221 (1978).
- ⁴⁹M. A. Paesler, *Phys. Rev. B* **17**, 2059 (1978).
- ⁵⁰L. M. Richardson and L. M. Scarfone, *Phys. Rev. B* **18**, 5892 (1978).
- ⁵¹L. M. Richardson and L. M. Scarfone, *Phys. Rev. B* **19**, 925 (1979).
- ⁵²D. E. Theodorou and H. J. Queisser, *Phys. Rev. B* **19**, 2092 (1979).
- ⁵³P. Csavinszky and R. A. Morrow, *Int. J. Quantum Chem.* **19**, 957 (1981).

- ⁵⁴P. Csavinszky, L. Beauprethuy, and R. A. Morrow, Phys. Rev. B **24**, 3602 (1981).
- ⁵⁵R. Resta, Phys. Rev. B **19**, 3022 (1979).
- ⁵⁶R. Resta and R. Resca, Phys. Rev. B **20**, 3254 (1979).
- ⁵⁷L. M. Scarfone and L. M. Richardson, Phys. Rev. B **22**, 982 (1980).
- ⁵⁸D. Chandramohan and S. Balusubramanian, Phys. Rev. B **31**, 3899 (1985).
- ⁵⁹P. Csavinszky and K. R. Brownstein, Phys. Rev. B **25**, 1362 (1982).
- ⁶⁰K. A. Chao, Phys. Rev. B **26**, 3463 (1982).
- ⁶¹H. J. Mattausch and W. Hanke, Phys. Rev. B **27**, 3735 (1983).
- ⁶²R. Car and Annabella Selloni, Phys. Rev. Lett. **42**, 1365 (1979).
- ⁶³Albert Messiah, *Quantum Mechanics*, 6th ed. (Wiley, New York, 1966), Vol. II, p. 817.
- ⁶⁴Leonard S. Rodberg and R. M. Thaler, *Introduction to The Quantum Theory of Scattering* (Academic, New York, 1967), p. 145.
- ⁶⁵Eugen Merzbacher, *Quantum Mechanics*, 2nd Ed. (Wiley, New York, 1970), p. 224.
- ⁶⁶J. R. Meyer and F. J. Bartoli, Phys. Rev. B **24**, 2089 (1981).
- ⁶⁷N. Takimoto, J. Phys. Soc. Jpn. **14**, 1142 (1959).
- ⁶⁸E. C. McIrvine, J. Phys. Soc. Jpn. **15**, 928 (1960).
- ⁶⁹H. I. Ralph, G. Simpson, and R. J. Elliot, Phys. Rev. **11**, 2948 (1975).
- ⁷⁰J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
- ⁷¹A. Baldereschi and Nunzio O. Lipari, Phys. Rev. B **8**, 2697 (1973).
- ⁷²G. I. Bir, E. Normantas, and G. E. Pikus, Fiz. Tverd. Tela (Leningrad) **4**, 1180 (1962) [Sov. Phys.—Solid State **4**, 867 (1962)].
- ⁷³The definitions of μ_H^N and σ_{123} differ from those in Ref. 8. The present way conforms to the definition used by other authors (Refs. 30 and 32). The advantage of the new definition is that the r_N approximate r factors for isolated bands whereas the total r is complicated by the weighting factors in Eq. (2.53).
- ⁷⁴R. Mansfield, Proc. Phys. Soc., London, Ser. B **69**, 862 (1956).
- ⁷⁵Frank L. Madarasz, Joseph E. Lang, and Patrick M. Hemenger, J. Appl. Phys. **52**, 4646 (1981).
- ⁷⁶Joseph E. Lang (unpublished).
- ⁷⁷W. C. Mitchel and P. M. Hemenger, J. Appl. Phys. **53**, 6880 (1982).
- ⁷⁸P. M. Hemenger, Rev. Sci. Instrum. **44**, 698 (1973).
- ⁷⁹R. Baron (unpublished); T. Peterson (unpublished).
- ⁸⁰Joseph E. Lang, Frank L. Madarasz, and Patrick M. Hemenger, J. Appl. Phys. **54**, 3612 (1982).
- ⁸¹David W. Fischer and John. J. Rome, Phys. Rev. B **27**, 4826 (1983).
- ⁸²N. O. Lipari, A. Baldereschi, and M. L. Thewalt, Solid State Commun. **33**, 277 (1980).
- ⁸³N. O. Lipari and A. Baldereschi, Solid State Commun. **25**, 665 (1978).
- ⁸⁴M. L. W. Thewalt, Solid State Commun. **23**, 733 (1977).
- ⁸⁵Joseph E. Lang (private communication).
- ⁸⁶Hp. Schad and K. Lassman, Phys. Lett. **56A**, 409 (1976).
- ⁸⁷R. Sauer, W. Schmid, and J. Weber, Solid State Commun. **27**, 705 (1978).
- ⁸⁸K. R. Elliot, S. A. Lyon, D. L. Smith, and T. C. McGill, Phys. Lett. **70A**, 52 (1979).
- ⁸⁹T. N. Morgan, Phys. Rev. Lett. **24**, 887 (1970).
- ⁹⁰R. Baron, M. H. Young, J. K. Neeland, and O. J. Marsh, Appl. Phys. Lett. **30**, 594 (1977).
- ⁹¹Walter Scott, Appl. Phys. Lett. **32**, 540 (1978).
- ⁹²C. E. Jones, D. Schafer, W. Scott, and R. J. Hager, J. Appl. Phys. **52**, 5148 (1981).
- ⁹³C. W. Searle, P. M. Hemenger, and M. C. Ohmer, Solid State Commun. **48**, 995 (1983).
- ⁹⁴H. R. Chandrasekhar and A. R. Ramdas, Phys. Rev. B **33**, 1067 (1986).