[24]L. Finegold, J. K. Hulm, R. Mazelsky, N. E. Phillips, and B. B. Triplett, Ann. Acad. Sci. Fennicae A 6, 129 (1966).

[25]J. K. Hulm, in NRL Report No. 6972, p. 30 (unpublished), available from the Naval Research Laboratory, Washington, D. C. 20390.

[26]J. R. Burke, B. Houston, and H. T. Savage, Bull. Am. Phys. Soc. 14, 329 (1969).

[27]R. F. Bis and J. R. Dixon, J. Appl. Phys. 40, 1918 (1969).

[28]L. Muldawer, Bull. Am. Phys. Soc. 16, 84 (1971).

[29]G. S. Pawley, W. Cochran, R. A. Cowley, and G. Dolling, Phys. Rev. Letters 17, 753 (1966).

[30]See comment at end of Ref. 14.

[31]J. R. Burke, J. D. Jensen, and B. B. Houston, in Ref. 8, p. 393.

[32]J. R. Burke and J. D. Jensen (private communication).

[33]S. Rabii (private communication).

[34]R. F. Bis and J. R. Dixon, Phys. Rev. B 2, 1004 (1970).

Excitonic Effects in Landau Transitions at the E_1 Edges of InSb and GaSb[†]

S. O. Sari[*]

Joseph Henry Physics Laboratory, Princeton University, Princeton, New Jersey 08540

(Received 12 November 1971)

Measurements are presented of differential reflectivity in a strong magnetic field in InSb and GaSb at the E_1 and $E_1 + \Delta_1$ edges. A description of the circular-polarization modulation technique used to obtain the data is given. Optical absorption at a critical point along the [111] direction in a semiconductor lacking inversion symmetry is discussed. An exciton model in a magnetic field is developed. Numerical solutions for the first four bound-state energies having azimuthal quantum number $m = 0$ are tabulated for the two-dimensional Schrödinger equation of an attractive Coulomb potential in a magnetic field. Solutions to the three-dimensional problem are discussed using the Born–Oppenheimer approximation. The exciton model is compared to experimental spectra of Landau levels at the E_1 edge of InSb. A value of 2.8 ± 0.6 meV is obtained for the exciton rydberg. This agrees with the value obtained using the measured effective transverse mass $m/\mu_t = 19.7 \pm 1.3$ and the static dielectric constant of InSb. A lower limit on the longitudinal mass of $15 \lesssim |\mu_l/\mu_t|$ is estimated. The sign of the longitudinal mass has not been unambiguously obtained; however, the critical-point symmetry at the E_1 edge may be of type M_0 instead of M_1. The E_1 excitonic energy gap is determined to be 2.015 ± 0.001 eV. Spectra of GaSb are compared to InSb. A line splitting of 33 meV due to the inversion asymmetry in GaSb is proposed.

I. INTRODUCTION

Effects due to electromagnetic transitions between Landau levels were first observed in crystalline solids in cyclotron-resonance experiments in 1953[1] and in infrared magneto-optical experiments[2] a short time later. Extensive work has followed on optical magnetic effects,[3] yielding accurate measurements of band masses and a wealth of other band-structure parameters. Up to the present time quantized harmonic-oscillator levels have been seen only in the infrared. They have been studied at the lowest-energy-gap thresholds, and in transitions from the spin-orbit split-off valence bands[4,5] in a number of semiconductors. Attempts to look at higher bands have not resolved magnetic levels.[6]

In the present experiment, however, Landau levels are seen in the visible energy range at the E_1 edge in InSb.[7] The measurements have been obtained using an optical-polarization modulation technique, which is relatively noiseless. The method is not unlike other techniques used to make

similar measurements,[8] but it has the virtue of requiring no modulation to be applied directly to the sample. In the present paper we have computed differential dielectric constants ($\Delta \epsilon_1$ and $\Delta \epsilon_2$) from our data. The "density of states" calculated from a model including excitons in a magnetic field has been compared to $\Delta \epsilon_2$. This approach has been necessitated by previous evidence for strong Coulomb effects at the E_1 thresholds.[9] A discussion of inversion asymmetry along the Λ direction has been included in the analysis. This may be important in interpreting new observations of optical transitions in semiconductors lacking inversion symmetry.

The "edge" E_1 where Landau-level transitions are seen is shown in the energy-band-structure diagram of InSb in Fig. 1. It is away from the Brillouin-zone center in the [111] direction and has a valence-conduction gap of approximately 2 eV. The component split-off by the spin-orbit interaction is $E_1 + \Delta_1$ about 0.5 eV higher. No Landau levels have been seen there thus far. The same statement may be made for still higher energy
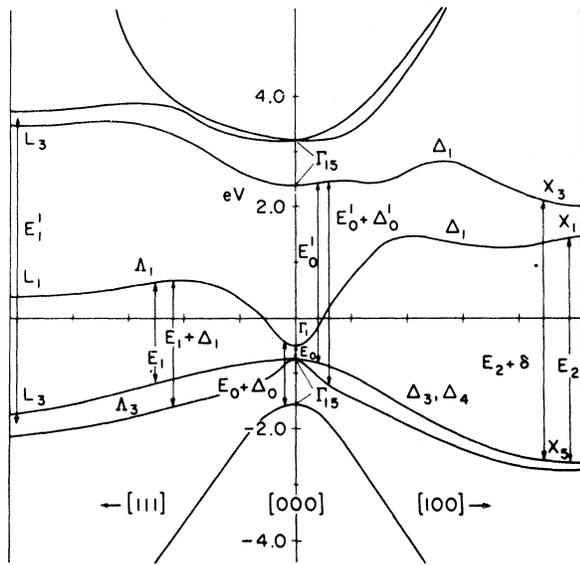
FIG. 1. Energy-band structure of InSb showing the location of various critical points where experimental structure is observed. (Reproduced from Ref. 8, p. 34, with permission.)

thresholds, also labeled in the diagram. The fundamental threshold E_0 at the zone center has an energy gap of 0. 2355 eV and optical magnetic effects have been extensively studied there.[3,5,10] Landau transitions at the spin-orbit split-off band $E_0 + \Delta_0$ ($\Delta_0 = 0. 803$ eV) have also been investigated.[11]

In Sec. II of this paper we describe the apparatus and technique for making our measurements in

some detail. The method of data analysis is considered in Sec. III. A simple model for magneto-optical absorption at critical points is developed in Sec. IV. A discussion of energy bands in the [111] direction in a zinc-blende semiconductor lacking inversion symmetry is given in Sec. V. The exciton model of Sec. VI is compared to experimental results in InSb in Sec. VII. We also compare spectra of GaSb to InSb (Sec. VIII). Landau levels are not resolved in GaSb in magnetic fields up to 62 kG, the highest available to us. The effective mass at the corresponding energy threshold appears to be larger than for InSb. A new peak in GaSb is seen, however, which may be due to inversion asymmetry. Finally, a summary of new InSb parameters and conclusions are presented in Sec. IX.

II. EXPERIMENT

The apparatus used to obtain the measurements reported here is shown in Fig. 2. The present setup is similar to that used for ellipsometry and magnetoreflection measurements described previously.[12-14] A 150-W Osram Xe arc lamp was used for a light source. Monochromatic light was produced by a $\frac{3}{4}$-m Spex Czerny-Turner spectrometer with a 102×102-mm 1200-line/mm grating blazed at 3000 Å. A cut-off filter was used behind the monochromator so only first-order diffracted light remained in the transmitted beam. Spectrosil quartz lenses and windows were used throughout the system. As a detector we used an EMI 9558QB photomultiplier (S-20 response) in a magnetically shielded can. A Magnion superconducting solenoid
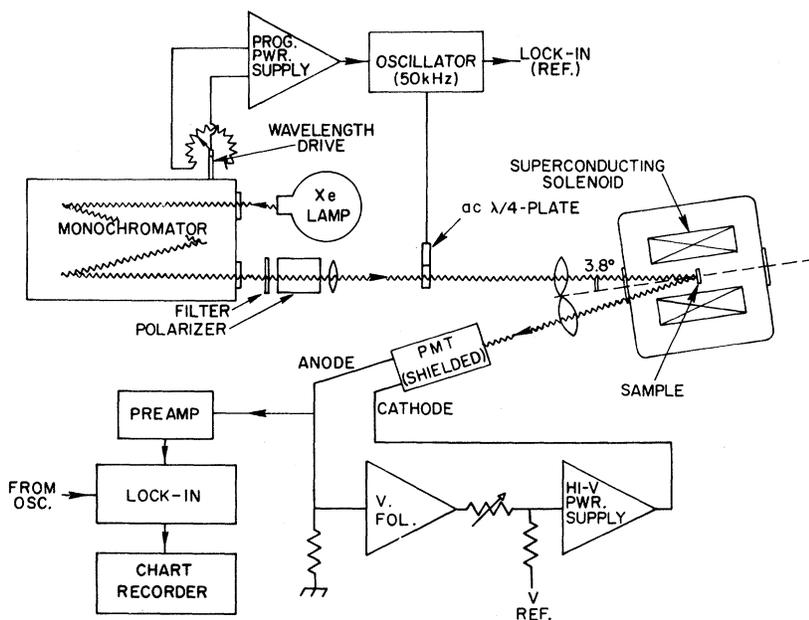
FIG. 2. Apparatus for the observation of interband magneto-optic effects.

with a 2-in. internal bore mounted horizontally in a Janis Dewar provided homogeneous magnetic fields up to 63.6 kG.

The light beam incident on the sample is first plane polarized. It is then passed through a modulated phase shifter, which is a uniaxially stressed fused-quartz rod.[12] The rod is resonantly coupled to a piezoelectrically driven quartz crystal, self-oscillating at 50 kHz. When plane-polarized light passes through the fused-quartz bar, a relative sinusoidal phase retardation $\delta = \delta_0 \sin\omega t$ is introduced between the axes parallel and perpendicular $(\hat{i}, \hat{j})$ to the stress. By orienting the linear polarizing axis at 45° to these directions, a modulated, elliptically polarized beam is produced with an electric field vector

$$\vec{E} = E_0(\hat{i} + e^{i\delta}\hat{j}) \ . \tag{1}$$

This may be decomposed into a basis of oppositely rotating circular vectors. The resultant intensity detected at the phototube after reflection from the sample is then obtained as

$$I \propto 1 + (\Delta R/2R)\sin\delta \ , \tag{2}$$

where $\Delta R/2R$ is the difference in reflectivity between left and right circularly polarized light normalized by twice the average reflectivity.

The oscillating part of the intensity ($\sin\delta$) will have a Fourier component at the fundamental frequency of the phase modulator at 50 kHz, as well as at odd harmonics. The 50-kHz signal is detected with a Princton Applied Research HR-8 lock-in amplifier and type-A preamplifier. The oscillator which drives the modulator is wavelength programed from the monochromator to ensure that the maximum amplitude of δ, namely δ_0, remains constant as the wavelength is scanned.[13]

Keeping the average photocurrent fixed normalizes the dc part of the intensity in Eq. (2). This is accomplished by comparing the photomultiplier anode voltage to a fixed reference voltage and employing any error signal to modify the high voltage on the cathode of the phototube. This feedback network averages out low-frequency intensity fluctuations. The bandwidth of the system is adjusted to be a few hundred cycles so that the 50-kHz modulation signal will not be averaged and can thus be detected.[13,14]

Measurements were made at near-normal incidence on samples anchored by a copper mount to a 4.2 °K liquid-helium bath. Vacuum grease held the samples to the mount. A check was made to see if the samples were under a static strain, which can shift the experimental spectra.[4] We loosely wrapped crystal specimens with soft 0.005-in. copper, allowing for optical access to the reflecting surface. The wrapper was then vacuum greased to the mount. The measured wavelength shift between this mounting scheme and direct mounting of the samples proved to be less than 1 Å, which is the approximate accuracy limit of our spectrometer. Hence strain effects due to mounting procedure can be ignored.

The samples were undoped, single crystals of InSb and GaSb cut to expose a [111] face.[15] Each specimen (0.8 mm thick) was polished with three grades of alumina and etched using standard recipes just prior to runs.[16]

Continuous wavelength scans were recorded on a Moseley 7100B strip-chart recorder. To normalize the measurements, a perfect circular polarizer was placed in the light beam to determine the unity signal. Markers could be set at 50 or 100-Å intervals for each scan. Wavelength calibration was periodically checked with low-pressure Hg and Tl lamp standards. Prior to analysis all spectra were converted to be functions of photon energy rather than wavelength. All measurements at high magnetic fields were taken at liquid-helium temperature in the Faraday configuration.

In Fig. 3, the measured differential reflectivity spectrum $\Delta R/R$ for InSb, as a function of incident photon energy, is shown in incremental magnetic field steps of 10 kG. The signal approaches zero at the low-energy side of the threshold and is taken as zero for $\hbar\omega < 1.85$ eV. The absolute sign of $\Delta R/R$ is not measured in this experiment. For any given sample and fixed magnetic field, the experimental peak heights were reproducible within 5%; from sample to sample, however, peak heights varied to about 20%. Variations between samples are observed in other modulation experiments[8] and in our case may be due to details of the surface preparation. On the other hand, the energies at which peaks occurred reproduced to within 1 Å for all runs.

The quiescent noise level of the apparatus, using a 1-sec lock-in integrating time, was approximately 5×10^{-5} so that the measured signal-to-noise ratio was usually better than 20 to 1.

III. ANALYSIS OF DATA

In analyzing the measurements we must consider the effect of circularly polarized light incident on a sample with its propagation vector parallel to the static magnetic field $\vec{B}$. For materials having at least a threefold rotational symmetry about the z axis, the dielectric response function $\overleftrightarrow{\epsilon}$ for $\vec{B} \parallel \hat{z}$ has the form[17]

$$\begin{pmatrix} \epsilon_0' & \epsilon_1 & 0 \\ -\epsilon_1 & \epsilon_0' & 0 \\ 0 & 0 & \epsilon_0 \end{pmatrix} \ . \tag{3}$$

In general, the off-diagonal component $\epsilon_{xy} = -\epsilon_{yx}$
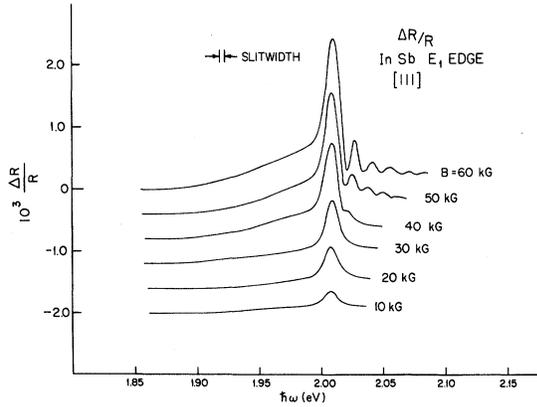
FIG. 3. Differential reflectivity spectrum for InSb taken at liquid-helium temperature. All signals approach zero for $\hbar\omega < 1.85$ eV. Curves for $B < 60$ kG have been displaced for clarity.

$= \epsilon_1$ must be an odd function of the applied B field. This must hold if the Faraday rotation is to reverse when the field is reversed. The diagonal component $\epsilon_{xx} = \epsilon_{yy} = \epsilon_0'$ is an even function of the field to zeroth (ϵ_0) and higher orders. These statements are consistent with the Onsager relations $\epsilon_{ij}(B) = \epsilon_{ji}(-B)$.[18]

In the presence of the dielectric tensor above, the eigenvectors for electromagnetic fields are states of left and right circular polarization. The propagation constant $k_{\pm} = (\omega/c)\epsilon_{\pm}^{1/2}$, where $\epsilon_{\pm} = \epsilon_0'$ $\pm i\epsilon_1$. According to standard convention, the upper (lower) sign is appropriate for left (right) circularly polarized waves.[17]

Using Maxwell's equations one may obtain the normal-incidence[19] amplitude from a plane interface as

$$r_{\pm} e^{-i\theta_{\pm}} = \frac{E_{r\mp}}{E_{i\pm}} = \frac{1 - N_{\pm}}{1 + N_{\pm}} , \qquad (4)$$

where $N_{\pm} = \epsilon_{\pm}^{1/2} = n_{\pm} + i k_{\pm}$ is the complex refractive index. If the reflectivity is defined as $R = r^2$, then Eq. (4) may be used to obtain[20]

$$i \Delta R/2R + \Delta\theta = i\Delta\epsilon/\epsilon_0^{1/2}(\epsilon_0 - 1) . \qquad (5)$$

The following definitions have been used:

$$\Delta\epsilon = \Delta\epsilon_1 + i\Delta\epsilon_2 = \epsilon_+ - \epsilon_- = 2i\epsilon_1 ,$$
$$\Delta R/2R = (R_+ - R_-)/(R_+ + R_-) , \qquad (6)$$
$$\Delta\theta = \theta_+ - \theta_- .$$

Equation (5) may be rewritten in terms of real and imaginary components as

$$\Delta\epsilon_1 = A_2 \Delta\theta + A_1 \Delta R/2R , \qquad (7)$$
$$\Delta\epsilon_2 = -A_1 \Delta\theta + A_2 \Delta R/2R ,$$

where $A_1 = n(\epsilon_1 - 1) - k\epsilon_2$ and $A_2 = k(\epsilon_1 - 1) + n\epsilon_2$.

The real part of ϵ_0 is ϵ_1 $(= n^2 - k^2)$ and the imaginary part is ϵ_2 $(= 2nk)$.

The coefficients A_1 and A_2 for InSb were determined from the room-temperature optical constants (n, k) of Phillipp.[21] The optical constants were shifted up in energy by 150 meV to correspond to liquid-helium temperature. The amount of energy shift was determined by comparing the location of the E_1 peak of Phillipp's normal-incidence reflectivity curve to the corresponding structure in 5 °K reflectivity (R) data.[22] Admittedly such a procedure cannot ensure that the optical constants used to calculate $\Delta\epsilon$ will be accurate. However, if we recalculate the low-temperature R from them, there is agreement to about 15%. Moreover, since the parameters A_1 and A_2 are to a large measure structureless, errors introduced in $\Delta\epsilon$ will be primarily in absolute magnitude and not in line shape. The coefficients A_1 and A_2 are shown in Fig. 4.

The differential phase shift $\Delta\theta$ (Fig. 5) was obtained by Kramers–Kronig transform of the differential reflectivity. We used the subtracted integral[20,14]

$$\Delta\theta(\omega) = \frac{\omega}{\pi} \int_0^\infty \frac{[\omega(\Delta R/R)(\omega') - \omega'(\Delta R/R)(\omega)] d\omega'}{\omega'[(\omega')^2 - \omega^2]} , \qquad (8)$$

in which the principal-value singularity does not appear. This integral is a transform between quantities of definite circular polarization and is not the usually encountered differential reflectivity phase relation for linearly polarized quantities. The differential phase-shift curves presented here were calculated numerically by applying Simpson's rule to Eq. (8). The E_1 and $E_1 + \Delta_1$ parts of $\Delta R/R$ from 1.75 eV to about 3 eV were included in the integration. To accomplish this, $\Delta R/R$ was smoothly continued to zero near 3 eV.
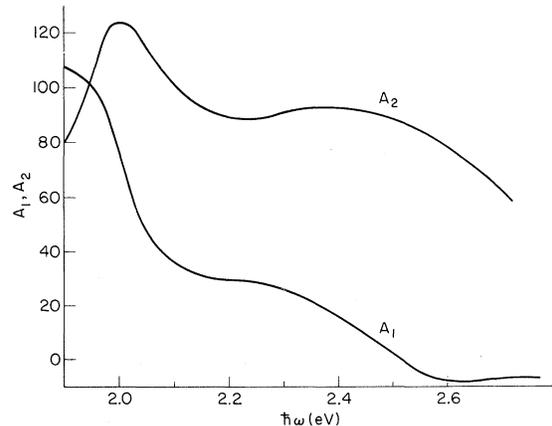
FIG. 4. Optical constants for InSb used for obtaining $\Delta\epsilon$.
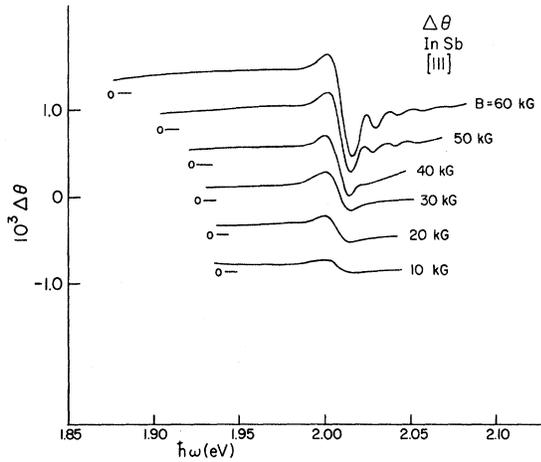
FIG. 5. Faraday rotation angle obtained from the data of Fig. 3 by Kramers–Kronig transformation. The zero level for each scan is shown.

The results were also calculated using the measured spectrum from 1.75 eV out to approximately 5 eV, where a smooth tail was again added. Near 2 eV, $\Delta\theta$ is shifted in magnitude by a few percent between these two cases. No correction for the infrared direct gap transitions at E_0 and $E_0 + \Delta_0$ was included. However, we expect that the only result would again be to shift $\Delta\theta$ slightly. Primarily affected is $\Delta\epsilon_1$, but only by a few percent.

The real part of the differential dielectric function $\Delta\epsilon_1$ is shown in Fig. 6. Because the constant $A_2 \gg A_1$, it is evident that over most of the photon energy range $\Delta\epsilon_1 \propto \Delta\theta$ and $\Delta\epsilon_2 \propto \Delta R/R$. Hence the $\Delta R/R$ measurements are primarily of differential absorption.

IV. SIMPLE MODEL FOR MAGNETIC DENSITY OF STATES AT CRITICAL POINTS

At a critical point $\vec{k}_0$ energy bands between valence and conduction states are, by definition, parallel. Hence in the vicinity of such a point the surfaces of constant energy can be approximated by

$$E(k) = \hbar^2(k_x^2 + k_y^2)/2\mu_t + \hbar^2 k_z^2/2\mu_l \,, \qquad (9)$$

where

$$E_c - E_v = E_g + E(k) \,. \qquad (10)$$

Here we have assumed at least threefold symmetry in the transverse plane to ensure that $\mu_x = \mu_y = \mu_t$. We have also assumed simple nondegenerate bands as occur along Λ (excluding spin). The wave vector $\vec{k}$ is measured from $\vec{k}_0$, and E_g is the energy gap. The reduced conduction-valence-band effective masses (μ_t, μ_l) measure the curvature of the bands near E_g. We have defined $1/\mu_i = 1/m_{ci} - 1/m_{vi}$, which is $1/m_{ci} + 1/$

$|m_{vi}|$ if m_{vi} is negative.

The sign of the effective reduced masses influences the line shape due to electronic transitions, which is seen for photon energies near E_g. If there are j negative masses, the critical point is said to be of type M_j. If all the masses are positive or negative the energy surface $E(k)$ is ellipsoidal, and the critical point is a maximum or a minimum. If the masses have different signs, $E(k)$ is hyperboloidal, and the critical point becomes a saddle point.

Since energy is conserved, photons of energy $\hbar\omega$ generate electron-hole pairs on the energy surface $E(k) = \hbar\omega - E_g$. Let us now suppose that a magnetic field B is applied. Then, as in other magnetic experiments, the quantum-mechanical effect of the magnetic field is to make only certain orbits allowed which satisfy an area quantization. The area A_t transverse to the field must satisfy

$$A_t = (2\pi e B/\hbar c)(n + \tfrac{1}{2}) \,. \qquad (11)$$

The area enclosed by each orbit (dA_t/dn) is independent of n. This means that in a magnetic field the same number of transverse plane wave states have coalesced to each magnetic level. The expression for the energy is given by

$$E_{nk_z} = \int \frac{\partial E(k)}{\partial A_t} \, dA_t + \frac{\hbar^2 k_z^2}{2\mu_l}$$

$$= \pm (n + \tfrac{1}{2}) \hbar\omega_c + \hbar^2 k_z^2/2\mu_l \,, \qquad (12)$$

where the integral encloses the area out to the nth orbit. The cyclotron frequency $\omega_c = eB/|\mu_t|c$, where $1/\mu_t = (2\pi/\hbar^2)\partial E(k)/\partial A_t$. The + or - sign in Eq. (12) allows μ_t to be either positive or negative.

If we examine which states are excited in an energy interval δE about $E(k)$, we find them to be
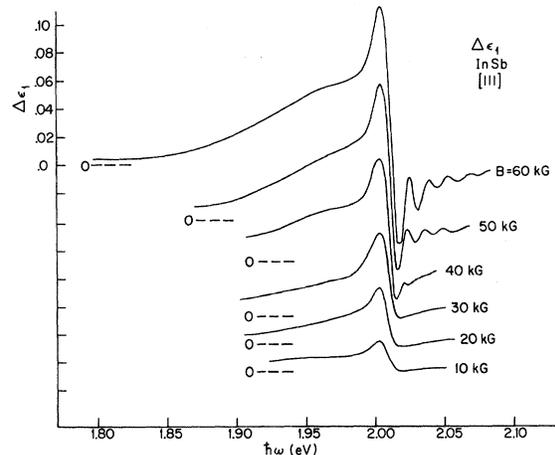
FIG. 6. Dispersive part of the optical dielectric function of InSb near the E_1 edge.
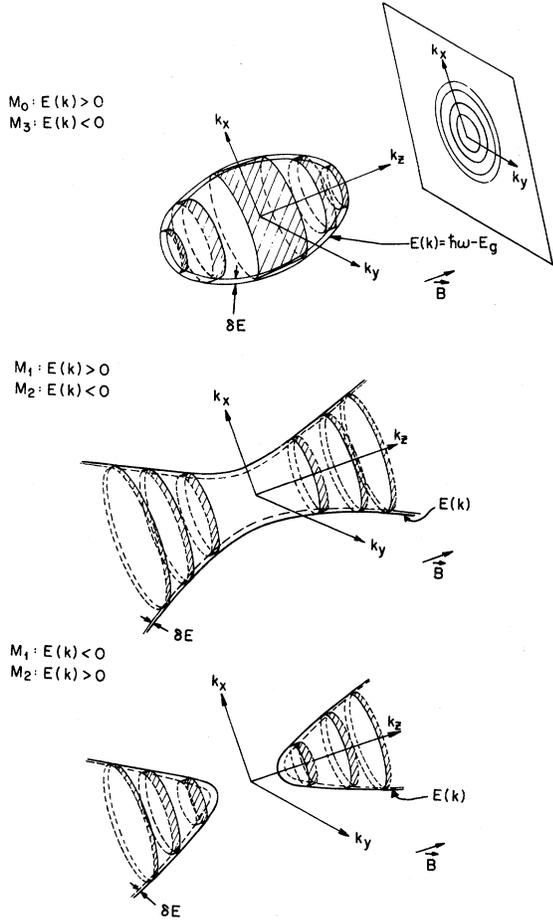
FIG. 7. Excited states in an energy interval δE in the presence of a magnetic field on an energy surface $E(k)$. The third orbit for the M_0 and M_3 case is extremal. The width of the ring is large indicating that the density of states is large.

given by those states in the shaded rings in Fig. 7. The number of states in the energy interval for each excited orbit will be proportional to the width of each ring. Hence that width will represent the density of states for each possible orbit. Since $\delta E = |\nabla_{k_z} E|\, \delta k_z$ we may write

$$\epsilon_2 \sim (\text{density of states}) \sim \sum_n \frac{dN}{dk_z}\frac{dk_z}{dE}$$

$$\sim \sum_n \frac{1}{|\nabla_{k_z} E|} \; . \qquad (13)$$

The sum is over all orbits which can be excited at energy $E(k)$ and the gradient is evaluated at $E(k) = E_{nk_z}$. This gives us an interpretation for extremal orbits for which $\nabla_{k_z} E = 0$ and for peaks in our experimental spectra which result from these orbits. $\nabla_{k_z} E = 0$ is a stationary phase con-

dition which means many states over a wide region of phase space will be coherently excited. As an example, the third orbit for M_0 and M_3 critical points is extremal. Hence the density of states has a singularity there which is proportional to $1/k_z$. The Coulomb interaction and other sources of broadening will prevent Eq. (13) from actually blowing up at the singularity.

We wish to examine some properties of this pictorial model. We have included the excited states (shaded rings) for critical points of all symmetries in Fig. 7. Changing the energy $\hbar\omega$ of the incident light means that a different energy surface will be sampled. Changing the magnetic field will change the area quantization instead. The density of states on each quantized orbit will, of course, change with either the field or the energy. A difference between ellipsoidal (M_0 and M_3 critical points) and hyperboloidal surfaces (M_1 and M_2 critical points) is that in the latter case one may have excitations and a nonzero density of states for both $\hbar\omega > E_g$ and $\hbar\omega < E_g$. For an M_0 critical point states are excited only when $\hbar\omega > E_g$, and for an M_3 critical point $\hbar\omega$ must be $< E_g$, in the absence of broadening. Hence only one figure is necessary for these cases. It appears as if an infinite number of states are excited simultaneously for each $\hbar\omega$ for the hyperbolic critical points whereas only a finite number are excited for ellipsoidal surfaces. This is, however, an error introduced by the effective-mass approximation, since we must restrict ourselves to wave vectors well within the Brillouin zone.

We have evaluated the optical density of states using Eqs. (12) and (13). The result is[7]

$$\epsilon_2(\omega) = \frac{e^2}{m^2\omega^2}\frac{eB}{\hbar c}\left(\frac{2|M|\mu_t}{\hbar^2}\right)^{1/2} |P|^2$$

$$\times \sum_n \frac{\theta(\pm[\hbar\omega - E_g \mp (n+\tfrac{1}{2})\hbar\omega_c])}{\{\pm[\hbar\omega - E_g \mp (n+\tfrac{1}{2})\hbar\omega_c]\}^{1/2}} \; . \qquad (14)$$

Here M is defined as μ_l/μ_t. The signs for critical points M_0 to M_3 in that order are $(+\,-)$, $(-\,-)$, $(+\,+)$, and $(-\,+)$. The line shape above is shown in Fig. 8. It has been Lorentzian broadened.

If the Coulomb interaction is included in the analysis, several important modifications to the free-particle model in a magnetic field will then occur. We summarize these briefly and give a discussion of details and of approximations to obtain them in Sec. V of this paper. Clearly the Landau-level spectrum will no longer be given by the transverse cyclotron energies $(n+\tfrac{1}{2})\hbar\omega_c$. For M_0 and M_1 critical points the levels will be shifted down in energy due to the attractive Coulomb field, the ground state shifting the most. The resonances in the spectrum will be determined by the eigen-
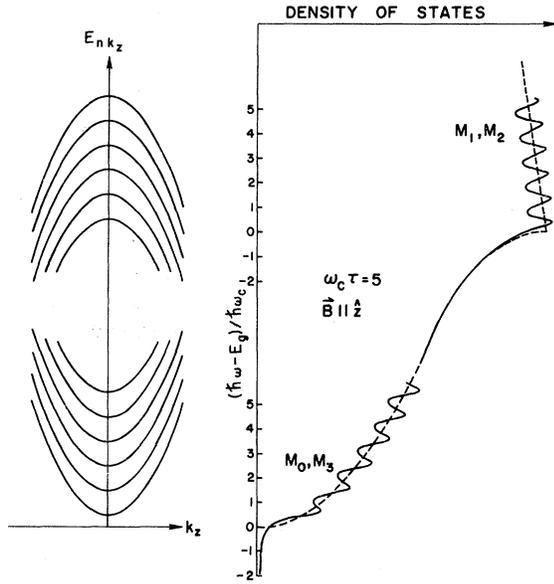
DENSITY OF STATES

FIG. 8. Density of states for a particle in a magnetic field for critical points of all symmetries. The energy scale must be reversed for the M_2 and M_3 critical points (from Ref. 7).

energies of states with no azimuthal angular momentum component, for direct allowed transitions. Only these states contribute, since the transition strength will be proportional to the probability that the electron and hole which are optically created lie in the same unit cell.[23]

In the limit that M is very large the absorptive part of the dielectric function $\epsilon_2(\omega)$ in Eq. (13) will be modified to be

$$\epsilon_2(\omega) \sim \sum_n \frac{1}{|\nabla_{k_z} E|} \times \left\{ \begin{array}{c} \text{amplitude for creating} \\ n\text{th exciton state} \end{array} \right\}^2 .$$

(15)

Here the excitonic envelope functions are solutions of the two-dimensional Schrödinger equation with magnetic and Coulomb fields. Only in the limit of very large magnetic fields do all amplitudes become equal so that Eq. (13) is recovered. When M becomes smaller but remains large in comparison to unity, the approximation (15) can under certain conditions still be used, as we will see, and we will examine its validity. For zero and small magnetic fields the $n = 0$ line will be strongly enhanced. On the basis of this model this is the line responsible for the strong "hyperbolic" exciton resonances observed in many semiconductors at anisotropic E_1 edges which are attributed to have M_1 symmetry.[8] It also accounts for the large oscillator strength of ellipsoidal excitons at M_0 critical points. We should add that resonances can also appear at both M_2 and M_3 critical points for finite magnetic fields. The ground-state amplitude

will be strongly depressed and the energy spectrum will be reversed. We will not discuss these negative-mass resonances in this paper.

V. ENERGY BANDS IN [111] DIRECTION IN ZINC-BLENDE MATERIAL

The shape of the energy bands and the associated wave functions near the Γ point and elsewhere in the the Brillouin zone have been extensively explored.[24-30] Dresselhaus[31] and Kane[32] have considered the $\vec{k} \cdot \vec{p}$ band structure for zinc-blende materials such as InSb which lack inversion symmetry. In the [111] Λ direction the results for the valence bands of highest energy are summarized in Fig. 9. This figure should be compared with Fig. 1. All splittings are exaggerated here for emphasis.

At the Brillouin-zone center the fourfold degenerate states having $p_{3/2}(\Gamma_8)$ symmetry lie at the top of the valence band. They are separated by the spin-orbit splitting Δ_0 from the twofold degenerate $p_{1/2}(\Gamma_7)$ states. Away from the zone center the fourfold symmetry of Γ_8 is lifted. The bands split into twofold degenerate light-hole and heavy-hole bands. The heavy-hole bands are further split by inversion asymmetry which introduces a term linear in the wave vector ($\pm \sqrt{2}\, Ck$ in Dresselhaus's notation) in the E-vs-k curves near $k = 0$. The energy maxima for these curves are hence moved away from Γ. Further out in the zone the splitting remains, and we have called it Δ_1^K. This is the splitting of $\Lambda_4 \Lambda_5$ at the critical point $\vec{k}_0$. Λ_4 and Λ_5 are states of $\pm \frac{3}{2} J_z$ symmetry which are compatible with the heavy-hole states at Λ_8. The light-hole states remain doubly degenerate as Λ_6 states. Far away from the zone center the $\Lambda_4 \Lambda_5$ states are believed to lie higher than Λ_6. The levels $\Lambda_4 \Lambda_5$ and Λ_6 may cross, but we only expect this behavior for pathological reasons which should show up in the optical spectra as well. We do not believe this to be the case for the data presented, and have assumed $\Lambda_4 \Lambda_5$ above Λ_6.[33] The spin-orbit splitting between $L_4 L_5$ and L_6 or $\Lambda_4 \Lambda_5$ or Λ_6 is approximately $\frac{2}{3}$ of the zone-center splitting. This "$\frac{2}{3}$ rule" has been important in the analysis of optical spectra.[34]

For materials with inversion symmetry $\Lambda_4 \Lambda_5$ do not split and are Kramers doublets. However, the lack of inversion symmetry can allow the twofold Kramers degeneracy to be lifted. Since the time-reversal operation connects Bloch states of opposite wave vector $k \to -k$ and spin $\sigma_z \to -\sigma_z$, those states will be degenerate. However, states at the same wave vector and opposite spin need not be, unless inversion symmetry is also present. We recall that inversion symmetry connects states of opposite wave vector and the same spin. Hence, except at certain places in the Brillouin zone, only
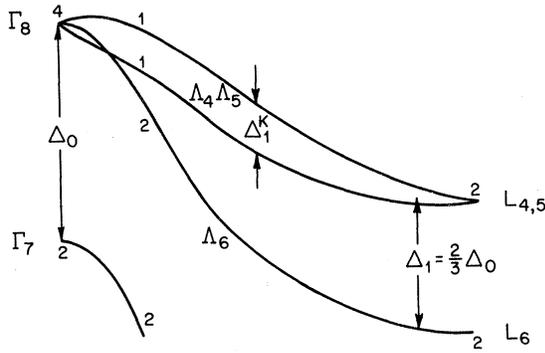
FIG. 9. Band structure in the valence band in the [111] direction of a zinc-blende semiconductor.

the combined operation of inversion and time reversal will flip a spin and ensure Kramers degeneracy. At the zone boundary inversion is not necessary, since translational invariance will connect states of the opposite wave vector. The same is true at the zone center where $\vec{k} = -\vec{k}$ automatically. Hence Λ_4 and Λ_5 can split except at the zone center and zone boundary. Dresselhaus gives a rough estimate of Δ_1^K as $\sim C(2\pi/a) \sim 20$ meV in InSb. Since the spin splittings introduced by the application of a magnetic field can be of comparable magnitude (~ 2 meV at 60 kG in InSb), the effect of including this quantity is important.

We can perform a $\vec{k} \cdot \vec{p}$ expansion about the critical wave vector $\vec{k}_0$ along Λ or at L. In the lowest conduction band the one-electron Bloch functions at $\vec{k}_0$ can be taken to transform as $\Lambda_1 \times D_{1/2}$ and will be given by

$$U_{c\dagger} = S\alpha , \quad U_{c\downarrow} = S\beta . \tag{16}$$

α and β are up- and down-spin functions and S is an s-like Bloch wave under C_{3v} symmetry. At the valence band edge the functions are taken to transform as $\Lambda_3 \times D_{1/2} = \Lambda_4 + \Lambda_5 + \Lambda_6$. They will be

$$U_{4\downarrow} = 2^{-1/2}(X - iY)\beta , \quad U_{5\dagger} = 2^{-1/2}(X + iY)\alpha$$

and $\tag{17}$

$$U_{6\dagger} = 2^{-1/2}(X + iY)\beta , \quad U_{6\downarrow} = 2^{-1/2}(X - iY)\alpha .$$

X and Y are Bloch functions which are p-like under C_{3v} symmetry. $U_{4\downarrow}$ and $U_{5\dagger}$, belonging to the representations $\Lambda_4 \Lambda_5$, will be mixed by Δ_1^K.

The $\vec{k} \cdot \vec{p}$ band structure about $\vec{k}_0$ may be obtained by calculating the matrix elements of

$$\vec{\pi} = \vec{p} + (\hbar/4m^2 c^2)(\vec{\sigma} \times \vec{\nabla} V)$$

between these states and carrying out the perturbation-theory computations. We have made the approximation in which only the interband momen-

tum matrix element $P = \langle S | p_x | X \rangle$ is nonzero.[35] The results are given and discussed in Appendix A.

The E_1 threshold in Fig. 1 consists of optical transitions between the valence bands $\Lambda_4 \Lambda_5$ and the conduction band Λ_1, so we must consider the effect of a magnetic field parallel to the longitudinal axis on these states. Those Bloch wave functions having wave vector $\vec{k}$ near the critical point at $\vec{k}_0$ will be mixed by the magnetic field and the Coulomb interaction and will form a Wannier excitonic wave packet. This will be examined in Sec. VI.

There is, in addition, a spin splitting of the states at $\vec{k}_0$, which can be treated by introducing the effective Hamiltonian $\mu^* \vec{\sigma} \cdot \vec{B}$.[29] Here $\vec{\sigma}$ is the Pauli spin operator, and μ^* is the magnetic moment obtained from the $\vec{k} \cdot \vec{p}$ perturbation analysis. In the conduction band $U_{c\dagger}$ and $U_{c\downarrow}$ are split by $2\mu_c^* B$. In the valence band including the inversion asymmetry splitting Δ_1^K gives

$$H_{\text{spin}} = \begin{pmatrix} \mu_v^* B & \frac{1}{2}i\Delta_1^K \\ -\frac{1}{2}i\Delta_1^K & -\mu_v^* B \end{pmatrix} , \tag{18}$$

where the basis states are $(U_{4\downarrow}, U_{5\dagger})$. This is easily diagonalized and gives the energy separation of $2[(\mu_v^* B)^2 + (\frac{1}{2}\Delta_1^K)^2]^{1/2}$ in the valence band. If $\Delta_1^K = 0$, the effective magnetic moment in the optical transition will be just $\mu^* = \mu_c^* - \mu_v^*$. If $\Delta_1^K \ne 0$, the situation is slightly more complicated and is discussed in Appendix A.

VI. EXCITON MODEL

Optical absorption in semiconductors proceeds by the simultaneous creation of electrons and holes, in pairs, which then scatter from one another via the Coulomb interaction. The scattering correlates the motion of the pairs to produce a new "quasiparticle" called an exciton. In a semiconductor in which effective masses are small and the dielectric constant of the material is large, bound exciton orbits will have dimensions much larger than the typical lattice spacing. Under such circumstances the effective-mass approximation[27] can usually be employed to analyze the motion.

The importance of including the Coulomb interaction in models of optical structure at critical points of various symmetries has been emphasized.[36] This has been necessary in order to interpret the experimental spectra in a variety of covalent and ionic solids. Kane has considered the Coulomb effect on transitions at an M_1 saddle point in which the longitudinal mass is negative.[37] Other investigators have computed electric field effects in the presence of the Coulomb poten-

tial.[38-40] Several calculations have also appeared in which a magnetic field has been included. [41-44]

Our results will be interpreted following a similar approach. We work in the framework of the model introduced by Elliott[23] and in the effective-mass approximation. Use is also made of the Born–Oppenheimer approximation since along Λ in the z direction the longtiudinal effective mass

is thought to be much larger than in the transverse direction.[37,45] This last approximation allows us to separate the Schrödinger equation and reduce it to a form which can be solved numerically.

The motion of each electron-hole pair, excluding the $\pm \mu^* B$ contribution to the Hamiltonian due to spin, will be given by the solution of the Schrödinger equation[43]

$$\left[\frac{\hbar^2}{2m_{te}} \left(-i\vec{\nabla}_{te} - \frac{e\vec{A}_e}{\hbar c} \right)^2 - \frac{\hbar^2}{2m_{th}} \left(-i\vec{\nabla}_{th} + \frac{e\vec{A}_h}{\hbar c} \right)^2 - \frac{\hbar^2\nabla_l^2}{2m_{le}} + \frac{\hbar^2\nabla_l^2}{2m_{lh}} - \frac{e^2}{\epsilon|r_e - r_h|} \right] \psi(\vec{r}_e, \vec{r}_h) = E\psi(\vec{r}_e, \vec{r}_h) .$$

(19)

We have assumed B parallel to the longitudinal direction and have separated out transverse (t) and and longitudinal (l) parts of the electron (e) and hole (h) motion. Moreover, we have taken the mass tensor to be diagonal and have assumed that each of the two transverse masses is the same. For simplicity a scalar dielectric constant has been chosen.

The center-of-mass transformation of the electron-hole coordinates $(\vec{r}_e, \vec{r}_h)$ has been discussed in detail[43] and will not be reproduced here. The inclusion of the center-of-mass motion will multi-

ply the wave function of the relative coordinates by a field-dependent phase factor $e^{-ie\vec{A}\cdot\vec{R}/\hbar c}$. $\vec{R}$ is the coordinate $(m_{te}\vec{r}_e + m_{th}\vec{r}_h)/(m_{te} + m_{th})$. Such a factor will not affect the optical absorption, which depends on the absolute square of a matrix element. We choose the cylindrical gauge so that

$$\vec{A} = \tfrac{1}{2}\vec{B} \times \vec{r} = \tfrac{1}{2} B(-y, x, 0)$$

and make the approximation that light has zero wave vector, so there is no center-of-mass exciton momentum. Equation (19) then reduces to

$$\left[- \frac{\hbar^2\nabla_t^2}{2\mu_t} - \frac{\hbar^2\nabla_l^2}{2\mu_l} + \frac{e\hbar}{2c} \left(\frac{1}{m_{te}} + \frac{1}{m_{th}} \right) BL_z - \frac{e^2}{\epsilon(\rho^2+z^2)^{1/2}} + \frac{e^2 B^2\rho^2}{8\mu_t c^2} \right] \psi_N(\rho, z) = E_N\psi_N(\rho, z) .$$ (20)

The relative electron-hole separation $r = (\rho^2+z^2)^{1/2}$, $\rho^2 = x^2+y^2$, and the reduced effective masses are as in Eq. (9). Equation (20) also occurs in the impurity problem in a magnetic field.

Because of cylindrical symmetry the azimuthal quantum number m remains good for all magnetic fields. States of different m are not mixed. In what follows only transitions for which $m = 0$ are considered, since only states with nonzero probability at the origin will contribute to the optical absorption. Since $L_z = 0$, B means $|B|$ and reversing the field will not affect the spectrum, in the absence of spin. Central cell effects will be ignored. Employing the Born–Oppenheimer approximation reduces (20) to the equations

$$\left(-\nabla_t^2 - \frac{2}{(\rho^2+z^2)^{1/2}} + \tfrac{1}{4}\gamma^2\rho^2 \right) \phi_n(\rho z) = E_n(z)\phi_n(\rho z) ,$$

(21a)

$$\left(-\frac{1}{M} \frac{\partial^2}{\partial z^2} + E_n(z) + W_n(z) \right) \psi_{nE}(z) = E\psi_{nE}(z) ,$$

(21b)

which may also be written as

$$\left(-\nabla_t^2 - \frac{2}{\gamma^{1/2}(\rho^2+z^2)^{1/2}} + \tfrac{1}{4}\rho^2 \right) \phi_n(\rho z) = E_n(z)\phi(\rho z) ,$$

(22a)

$$\left(-\frac{1}{M} \frac{\partial^2}{\partial z^2} + E_n(z) + W_n(z) \right) \psi_{nE}(z) = E\psi_{nE}(z) .$$

(22b)

Here $\psi_N(\rho z) = \phi_n(\rho z)\psi_{nE}(z)$. The quantity

$$W_n(z) = \frac{1}{M} \int d\rho \left| \frac{\partial\phi_n}{\partial z} \right|^2$$

is a term which represents the extra longitudinal kinetic energy of the transverse motion and is needed to ensure that the eigenenergies obtained will be upper bounds on the true energy. We will, in general, ignore this term, however (see Appendix B).

Equations (21) and (22) are written in dimensionless form. In Eq. (21), the unit of length is $a = \hbar^2\epsilon/\mu_t e^2$, the exciton Bohr radius, and the energies are in exciton rydbergs $\Re = \mu_t e^4/2\hbar^2\epsilon^2$. In Eq. (22), the corresponding quantities are the

cyclotron radius $\lambda = (\hbar/\mu_t\omega_c) = (\hbar c/eB)^{1/2}$ and the zero-point cyclotron energy $\hbar\omega_c/2$ where the cyclotron frequency $\omega_c = eB/\mu_t c$. For these calculations the transverse reduced mass μ_t has been assumed positive. Two parameters characterize the problem:

$$\gamma = \hbar\omega_c/2\Re = (a/\lambda)^2 \quad \text{and} \quad M = \mu_l/\mu_t \ .$$

It is important to consider both forms (21) and (22), since one wishes to work with the finest mesh for accuracy in numerical computations. Hence, Eqs. (21) are most useful for $\gamma \to 0$ and Eqs. (22) for $\gamma \to \infty$. It should be noted that (21a) and (22a) are just the equations of the two-dimensional Coulomb problem with a magnetic field $(z = 0)$.

The imaginary part of the dielectric function at the optical frequency ω is given in the linear-response approximation for direct allowed transitions by[23]

$$\epsilon_2(\omega) = \frac{4\pi^2 e^2}{\omega^2 m^2} \sum_N |\langle v|\epsilon \cdot p|c\rangle|^2$$
$$\times |\psi_N(0)|^2 \delta(E_\omega - E_N) \ , \quad (23)$$

where $E_\omega = \hbar\omega - E_g$ and ψ_N is the excitonic envelope function. The interband momentum matrix element is assumed to be constant and $\phi_n(\rho z)$ is so normalized so that

$$\int d\vec{\rho}\, |\phi_n(\rho z)|^2 = 1 \ . \quad (24)$$

This condition can always be met when there is a finite magnetic field, since all the states ϕ_n will then be bound states. In a zero magnetic field the continuum states must be treated separately, but fortunately do not have much oscillator strength.[37] Thus for our purposes they are relatively unimportant. $\psi_{nE}(z)$ is taken to be symmetric in z, and may be either a bound or continuum state. If we ignore the complication of bound states for the moment $\psi_{nE}(z)$ is a standing wave at infinity. It can then be written with an energy normalization such that

$$\int dz\, \psi_{nE'}\psi_{nE} = \delta(E - E') \ , \quad (25)$$

i. e., such that the asymptotic outward current is just $\frac{1}{2}\pi\hbar$. For an M_1 critical point this is no restriction since there are no bound states in the z direction. The approximation is worse in the case of an M_0 critical point where both bound and continuum states are possible.

For potentials in the z direction that are sufficiently weak, $\epsilon_2(\omega)$ may be written as

$$\epsilon_2(\omega) = \frac{4\pi^2 e^2}{\omega^2 m^2} |\langle v|\hat{\epsilon}\cdot\vec{p}|c\rangle|^2$$

$$\times \sum_n |\phi_n(0)|^2 \int_0^\infty dE'\, |\psi_{nE'}(0)|^2 \delta(E_3 - E')$$

$$= \frac{4\pi^4 e^2}{\omega^2 m^2} |\langle v|\hat{\epsilon}\cdot\vec{p}|c\rangle|^2 \sum_n |\phi_n(0)|^2 |\psi_{nE_3}(0)|^2 \ , \quad (26)$$

where $E_3 = E_\omega - E_n(0) - W_n(0)$. The quantity $E_n(0) + W_n(0)$ has been subtracted from E_ω to bring out the location of the singularities explicitly, and this amounts to a shift of energy scale in Eqs. (21b) and (22b). The location of peaks will occur when $E_3 \approx 0$ or for

$$E_\omega = E_n(0) + W_n(0) \approx E_n(0) \ . \quad (27)$$

The form (26) can be seen to reduce to the free-magnetic-field result, as given in Eq. (14), when the Coulomb interaction is turned off. For critical points of M_0 symmetry this is seen by noting that

$$|\psi_{nE_3}(0)|^2 \to \frac{1}{2\pi} \frac{(2M)^{1/2}}{\hbar} \frac{\theta(+E_3)}{(+E_3)^{1/2}} \ ,$$

$$|\phi_n(0)|^2 \to (1/2\pi)(eB/\hbar c) \ . \quad (28)$$

For M_1 symmetry the sign of E_3 is reversed. θ is the unit step function. One may, in addition, turn off the magnetic field to obtain the ordinary free-particle line shape

$$\epsilon_2(\omega) = \frac{e^2}{m^2\omega^2}\left(\frac{2|M|\mu_t^3}{\hbar^2}\right)^{1/2} |\langle v|\hat{\epsilon}\cdot p|c\rangle|^2$$

$$\times \begin{cases} (E_\omega)^{1/2}\theta(E_\omega) & (M_0) \\ (E_{co})^{1/2} - |E_\omega|^{1/2}\theta(-E_\omega) & (M_1). \end{cases} \quad (29)$$

E_{co} is a cut-off energy which appears for M_1 calculations. Its use is necessitated by the breakdown of the hyperbolic energy surface approximation for large wave vectors, as noted in the introduction.

The Cooley–Numerov method[46] has been applied to the solution of the transverse motion, after making the transformation $u \to \rho^{1/2}\phi$. Exact results to which the numerical calculations may be compared are, of course, available in both the zero-$(\gamma = 0)$ and high-field $(\gamma \to \infty)$ limits. A calculation using the Born–Oppenheimer approximation for zero magnetic field has been given by Ref. 37.

Numerical computations for the first four bound states of the two-dimensional Coulomb problem with a magnetic field are shown in Figs. 10 and 11 and in Table I. The dashed lines in the figures indicate the outcome from an approximate calculation in which the Coulomb potential was diagonalized in a harmonic-oscillator function basis using a 10×10 matrix expansion. This was performed as a consistency check. For comparison to the
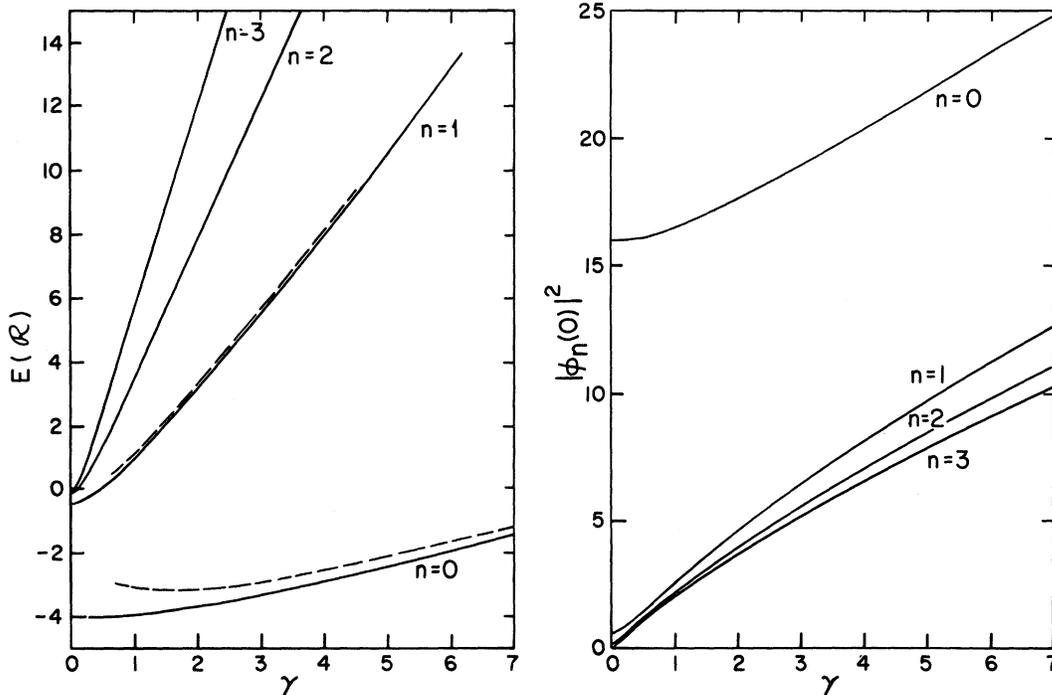
FIG. 10. Numerical solution to the two-dimensional Schrödinger equation with magnetic and Coulomb potentials. The dashed line indicates an approximate calculation in which the $1/r$ potential is diagonalized in a harmonic-oscillator basis. A 10×10 matrix representation has been used. The probability density at the origin for the first four eigenstates is also shown. Note that the ground-state amplitude is pushed up relative to the higher states by the attraction of the electron and hole.

table it should be noted that at zero magnetic field

$$E_n(0) = [-4/(2n+1)^2](\mathfrak{R})$$

and that

$$|\phi_n(0)|^2 = 16/2\pi(2n+1)^3 a^2 .$$

It is clear that the effect of the Coulomb interaction is left to quite high values of γ. For the experiment reported in this paper, however, we are primarily interested in the behavior of the solutions for the region in which $1 < \gamma < 3$ where magnetic and Coulomb energies are comparable.

The longitudinal motion in the z direction is obtained by solving the one-dimensional Schrödinger equation (21b) or (22b) after obtaining the adiabatic potential from the transverse motion. Either of these equations may be written in the subtracted form introduced in (26) to obtain

$$\left(-\frac{1}{M}\frac{\partial^2}{\partial z^2} + V_n(z)\right)\psi_{nE_3} = E_3\psi_{nE_3} , \qquad (30)$$

where

$$V_n(z) = E_n(z) + W_n(z) - E_n(0) - W_n(0) .$$

If $M < 0$ it may be written as

$$\left(-\frac{1}{|M|}\frac{\partial^2}{\partial z^2} + [-V_n(z)]\right)\psi_{nE_3} = (-E_3)\psi_{nE_3} . \qquad (31)$$

The effect of the negative mass has been to change the sign of the potential and energy. A one-dimensional bound- and continuum-state problem becomes the equivalent one-dimensional potential barrier scattering problem, as pointed out by Velický and Sak.[47] These authors consider this problem in the WKB approximation, whereas Kane has numerically integrated the longitudinal Schrödinger equation.

A summary of efforts to obtain solutions for the exciton in a magnetic field is given by Baldereschi and Bassani.[44] They have noted that good agreement to variational calculations to within a few percent for $\gamma > 1$ is attained by using the Born–Oppenheimer approximation. However, with this procedure the correct eigenenergies can be overestimated slightly. We present the eigenenergies (000) and (100) of Eq. (30) as a function of M in Fig. 12. The accuracy of the results will improve with increasing M. The notation is (n, m, i) of Elliott and Loudon. Here n is the Landau quantum number, m is the azimuthal quantum number, and i labels the states of $V(z)$. For very large mass ratios $M(M \to \infty)$ the longitudinal motion becomes
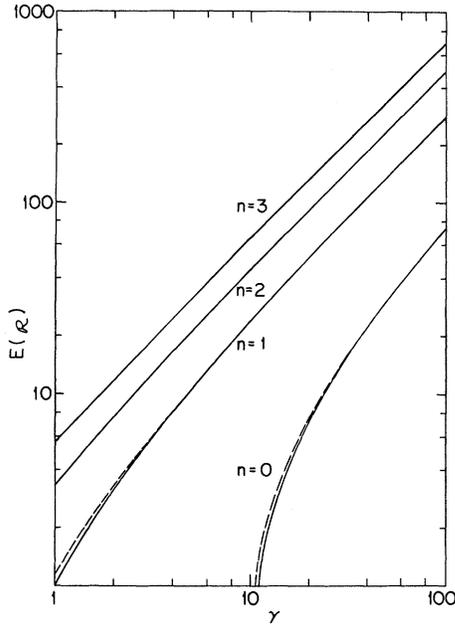
FIG. 11. First four eigenstates for the two-dimensional Coulomb magnetic field problem for large values of γ. The dashed line is the approximate solution.

free-particle-like. This may be seen by defining $z = z'/M^{1/2}$ in Eqs. (30) and (31) and noting that $V(z'/M^{1/2}) \to V(0) = 0$. However, we find that the eigenenergies do not approach those of the two-dimensional problem until mass ratios $M \gg 10^3$. But if we examine the difference energy, such as

between (000) and (100), we see that it agrees with the two-dimensional model for $M \gtrsim 20$ to within 27% for $\gamma = 1.5$, 16% for $\gamma = 2.5$, and 5% for $\gamma = 10$. The corrections for oscillator strength will be a few percent larger in each case. The energy difference between higher states $(n00)$ and $(n+100)$, where $n \geq 1$, will check much better with the two-dimensional difference energies. Within a given Landau subband, for mass ratios $M < 100$ at least 80% or so of the oscillator strength is in the $i = 0$ transition for $\gamma < 10$. We have checked this by examining some of the higher states i. These calculations are not included here, but the conclusion may be obtained by using the transition strengths of Ref. 44.

Hence we feel that a reasonable first approximation to a highly anisotropic critical point is to use a two-dimensional model. At an M_1 critical point there is no difficulty since there are no zero-point energy shifts. The zero-point energy of the bound states of an M_0 critical point *are* ignored, however. Moreover, the continuum approximation used to obtain Eq. (26) would seem to be poor, since all the transition strength is in the first bound state. In the case of highly broadened spectra, however, the line shapes due to bound states or continuum states are not readily distinguished. This is the limit of the present experimental data, and so we make the approximation. In addition there is the virtue that the model allows a determination of the excitonic rydberg independent of the sign of M.

A method of treating the motion in z is to take

TABLE I. Energies and probability densities at the origin for the first four bound states of the two-dimensional Schrödinger equation with Coulomb and magnetic fields.

	$E(\mathfrak{R})$				$2\pi|\phi_n(0)|^2$ (Bohr radius)$^{-2}$			
γ	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 0$	$n = 1$	$n = 2$	n = 1
0	−4.000	−0.444	−0.160	−0.082	16.0	0.59	0.13	0.047
0.3	−3.993	−0.210	+0.624		16.0	1.02	0.74	
0.5	−3.977	+0.081	1.353	2.486	16.1	1.45	1.16	1.08
0.7	−3.956	0.421	2.132	3.690	16.2	1.90	1.58	1.47
1.0	−3.911	0.989	3.354	5.543	16.5	2.57	2.17	2.01
1.5	−3.808	2.034	5.474	8.707	17.0	3.63	3.08	2.84
2.0	−3.674	3.154	7.657	11.93	17.6	4.62	3.93	3.62
2.4	−3.548	4.085	9.432	14.53	18.1	5.37	4.58	4.22
2.6	−3.479	4.560	10.33	15.83	18.3	5.73	4.89	4.51
3.0	−3.333	5.524	12.13	18.46	18.9	6.43	5.50	5.07
3.5	−3.135	6.752	14.40	21.75	19.6	7.29	6.25	5.77
4.0	−2.921	8.000	16.69	25.07	20.3	8.11	6.97	6.45
5.0	−2.454	10.54	21.30	31.72	21.8	9.68	8.37	7.78
6.0	−1.946	13.13	25.95	38.41	23.3	11.2	9.73	9.07
8.0	−0.836	18.40	35.34	51.86	26.2	14.0	12.4	11.6
10	+0.368	23.77	44.80	65.38	29.0	16.8	14.9	14.0
20	7.18	51.28	92.68	133.5	43.0	29.6	27.0	25.7
40	22.61	107.8	189.7	270.8	69.0	53.4	49.8	48.1
70	47.52	193.9	336.4	477.9	106	87.5	82.9	80.6
100	73.45	280.9	483.8	685.5	141	120.8	115.4	112.8
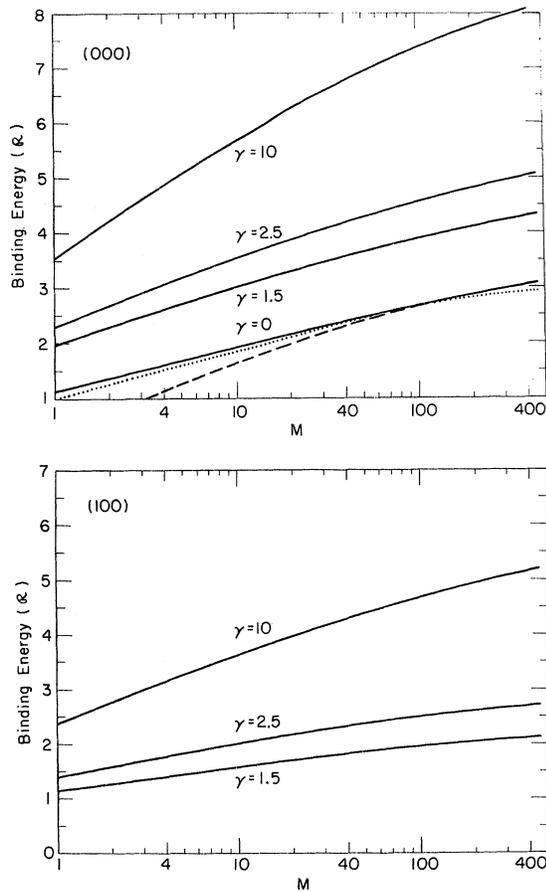
FIG. 12. Ground states (000) and (100) of Eq. (30). $V(z)$ has been computed at 100 well-chosen points. For the integration mesh, the potential values were calculated by interpolating quadratically from these starting values. The Cooley—Numerov technique was used. For the (000) state the dotted line is a calculation of the ground state using the variational method of Ref. 37. The dashed line is the Born—Oppenheimer result including the term $W_0(z)$.

a triangular barrier (linear) approximation. With such a simplification the problem can be solved exactly in terms of tabulated Airy functions and a simple expression for the effective mass in terms of the half-width of the line shape can be obtained. We expand $V_0(z)$ to first order in z to get $V_0(z) = 4\pi \mid \phi_n(0) \mid^2 \mid z \mid$. The coefficient is the perturbation-theory result for the nth state. We may then write Eq. (31) for the negative-mass case as

$$\frac{d^2\psi}{d\eta^2} - \eta\psi = 0 \;, \tag{32}$$

where

$$\eta = -[4\pi \mid \phi_n(0) \mid^2 \mid M \mid]^{1/3} \mid z \mid + \nu$$

and

$$\nu = \mid M \mid^{1/3} E_3 / [4\pi \mid \phi_n(0) \mid^2]^{2/3} \;.$$

The normalized solution can be obtained in terms of the two kinds of Airy functions by applying the condition $(\partial\psi/\partial z)(0) = 0$. This boundary condition assures that the same probability amplitude leaks out in both the position and negative z directions. The value of the function ψ_{nE_3} at the origin will determine the line shape in this approximation and for the nth state, it is[48]

$$\mid \psi_{nE_3}(0) \mid^2 = \frac{[M^2/4\pi \mid \psi_n(0) \mid^2]^{1/3}}{(2\pi)^2 [\mathrm{Ai}'(\nu)^2 + \mathrm{Bi}'(\nu)^2]} \;. \tag{33}$$

The line shape has been plotted in Fig. 13. For $\nu \ll 0$ we obtain $(1/2\pi)(M/ - E_3)^{1/2}$ which is just the free-particle line shape. This is illustrated in case A of Fig. 14. For $\nu > 0$ the exciton must tunnel out from the origin so the line shape will be exponentially damped as shown in case C. The maximum amplitude occurs for $E_3 = 0$ in case B. This conclusion is understandable, since the probability wave starts off with "zero" velocity or highest probability amplitude under this condition. Elliott and Loudon have derived line shapes similar to that given by Eq. (33), in the positive-mass case. However, they have assumed a more realistic potential $V(z)$, which unfortunately leads to greater analytic complexity. Our calculation is simplistic but involves well-known Airy functions, and it illustrates the source of broadening due to the Coulomb interaction.

An estimate of the longitudinal effective mass may be obtained from Eq. (33) by assigning a half-width to ν, even though the line shape is asymmetric. We take $\nu_{hw} \approx 2.9$ in which case

$$\mid M \mid \approx \frac{\nu_{hw}^3 \mid 4\pi\psi_n(0)^2 \mid^2}{\Gamma_{hw}^3} = \frac{25 \mid 4\pi\psi_n(0)^2 \mid^2}{\Gamma_{hw}^3} \;. \tag{34}$$

Γ_{hw} is the linewidth in exciton rydbergs. Unfortunately, since the initial slope near $z = 0$ is a poor approximation to the true potential, the relation (34) will vastly overestimate $\mid M \mid$. However, an order-of-magnitude estimate may still be made by choosing an effective slope for the linear potertial as we will do in Sec. VII.

Qualitatively, the inverse relation between longitudinal mass and half-width may be understood by noting that the negative mass makes the exciton unstable. Hence, the smaller the negative mass, the faster the probability amplitude can leak away from the origin, the shorter the lifetime of the state, and the larger the linewidth.

VII. COMPARISON OF EXPERIMENT TO EXCITON MODEL IN InSb

We wish to fit the absorptive part of the differential dielectric function $\Delta\epsilon_2$ with the model intro-
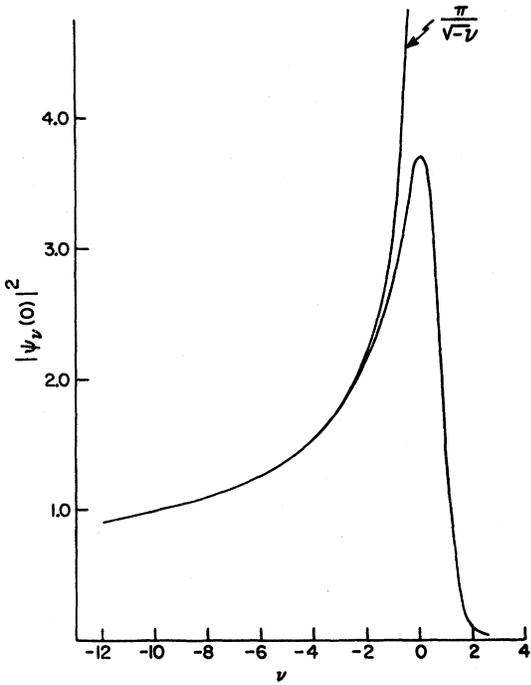
FIG. 13. Line shape for the linear approximation to the adiabatic potential barrier.

duced in Sec. VI. Results of this fit are given in Fig. 15. We have made the further simplifying assumption that the longitudinal motion can be approximated by a Lorentzian with a broadening parameter Γ so that
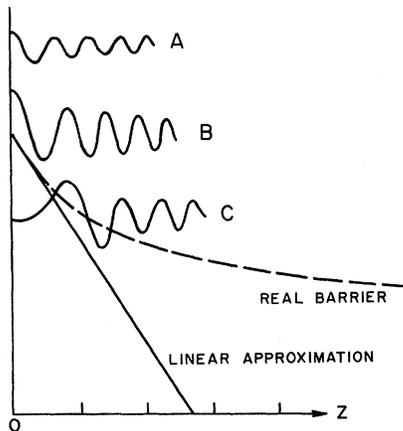
FIG. 14. Probability waves in the z direction leading to the line shape in Fig. 13. In case A the particle has more energy than the barrier height. For case B the particle energy is the same as the barrier potential at the origin so the particle starts off with the largest amplitude at the origin. In C, the particle must tunnel out from the middle of the barrier. Hence the wave amplitude at $z = 0$ is attenuated.

$$\frac{2\pi}{M^{1/2}}\,|\psi_{nE_3}(0)|^2 \rightarrow \mathrm{Re}\,\frac{1}{(-E_3+i\Gamma)^{1/2}}$$

$$= \left(\frac{(E_3^2+\Gamma^2)^{1/2}-E_3}{2(E_3^2+\Gamma^2)}\right)^{1/2}. \qquad (35)$$

This is equivalent to assuming free-particle motion in the longitudinal direction, which will not alter the line shape drastically from the exact solution. Moreover, since scattering mechanisms other than the Coulomb effect may broaden the structure, the
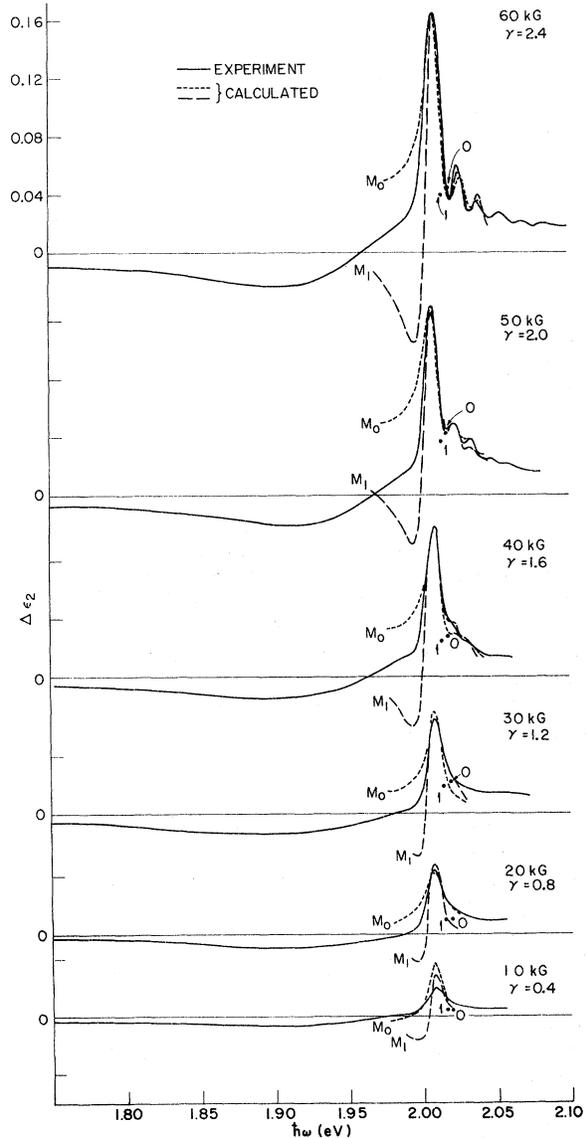
FIG. 15. Comparison of calculated and experimental curves for $\Delta\epsilon_2$. We have included the low-energy tails on the experimental curves to show how $\Delta\epsilon_2$ approaches zero. Infrared contributions to the absorption are omitted. The points marked 0 and 1 are coordinate origins for the M_0 and M_1 spectrum calculations.
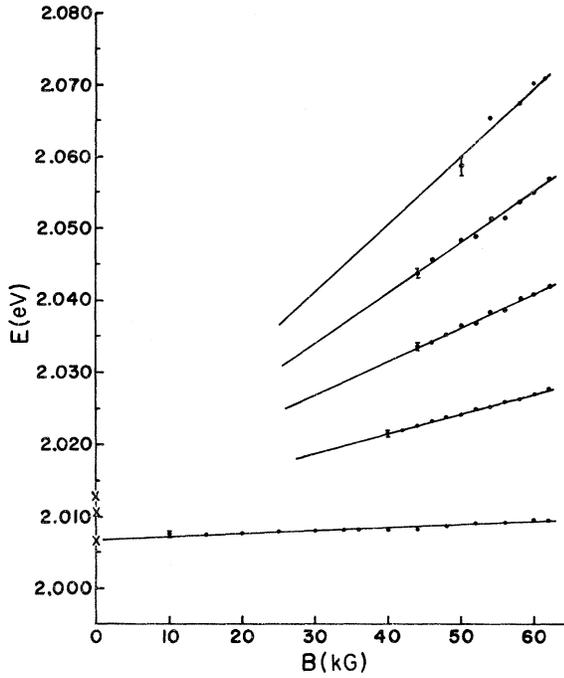
FIG. 16. Magnetic field dependence of peaks in the experimental magnetoreflectance structure. The error bars show typical uncertainty intervals in the measured data. The ×'s mark the intercepts for the $n = 0$, 1, and 2 states.

use of phenomenological broadening is not unreasonable. Explicitly, these could be due to a number of sources, say phonon or intraband. In addition, the sample preparation may introduce a surface potential which will bend the valence and conduction bands. Since the reflected signal is composed of scattered photons from near the surface, a suitable average over the surface potential is measured. This will have the effect of broadening the line shape.

The data are compared to (26) which may be written in differential form as

$$\Delta\epsilon_2 = \frac{4\pi^2 e^2}{\omega^2 m^2} |P|^2 \frac{1}{\Re a^3} \sum_n |\phi_n(0)|^2$$

$$\times \left[|\psi_{nE_3 + \mu^* B}(0)|^2 - |\psi_{nE_3 - \mu^* B}(0)|^2 \right] , \quad (36)$$

$$\Delta\epsilon_2 \approx 4\pi^2 \frac{e^2}{\omega^2 m^2} \frac{P^2}{\Re a^3} \sum_n |\phi_n(0)|^2 \frac{\partial |\psi_{nE_3}(0)|^2}{\partial E_3} 2\mu^* B , \quad (37)$$

where $\phi_n(0)$ is normalized as in Table I and ψ_{nE_3} as above. In the present work Δ_1^K is taken to be zero. This quantity appears to be small in InSb and may indicate that the transitions occur near L. $\Gamma = 3.5$ ×$\Re$ was determined experimentally from the condition $\omega_c \tau \approx 1$, where $\Gamma = \hbar/\tau$. Since the calculated spectrum was obtained numerically, the subtrac-

tion in Eq. (36) was computed directly instead of using the first-order expression (37). We plotted the M_0 line shape in this manner and then simply reversed the sign of E_3 in (36) to get the M_1 line shape in the equivalent approximation.

All fits were made with a single field-independent Γ. The calculated spectrum was scaled to fit the first two peaks of the measured spectrum at 60 kG. The relative zero of the theoretical curve is displaced upward by 0.04 at 60 kG from the experimental zero-signal value. This may indicate the presence of broad-band transitions from elsewhere in the Brillouin zone, or of transitions to degenerate Λ singularities made inequivalent due to magnetic field direction, which generate aperiodic exciton orbits. The background is roughly linear in B. It should be noted that the signal amplitude in Fig. 15 becomes nonlinear at high fields. This behavior is expected from (36) in both the excitonic and free-field models for the absorption. The spin splitting $\mu^* B$ is taken to be 0.29$\gamma(\Re)$, which has been calculated using Eqs. (46) in Appendix A and the measured cyclotron energy splitting.

The dip in the calculated curve below the 2-eV threshold for the M_1 fit arises from the subtraction of two $1/(\text{energy})^{1/2}$ tails (independent of the model used for the longitudinal motion) and does not seem to be reproduced in the data. The M_0 threshold produces no such dependence and in this sense can be made to fit the structure better, particularly if the free-particle model in the z direction is improved upon. We do not, however, believe that we can draw a definite conclusion from the present experiment about the sign of M.

The magnetic field dependence of the peaks in the measured $\Delta R/R$ spectrum is shown in Fig. 16 and is similar to the results of other magneto-optical experiments in which excitonic absorption is important.[49] We have performed a linear least-squares fit to this data and the parameters are displayed in Table II. If a linear extrapolation is used to the origin, a definite gap appears for the first few states. This procedure has been repeated using the results of our numerical calculation. The linear least-squares fit to peaks of the calculated spectra were made at values of the magnetic field equivalent to those of the experimental spectra. We used this method in an attempt to ensure consistency in our comparison of measured and computed quantities. The location of singularities is essentially determined by the maxima in $\partial |\psi_{nE_3}(0)|^2/\partial E_3$. These occur near the eigenenergies $E_n(0)$ in Table I, as discussed in the previous section. A calculation without the Coulomb interaction was also performed, and those results are listed in Table II as well. With both models there will be small shifts ($= \Gamma/10^{1/2}$) due to broadening of the

TABLE II. Magnetic field dependence of peak energies of the Landau-level spectrum.

n	Experiment $(\Delta R/R)$ $(E$ in eV$)$	Calculation with Coulomb interaction $(E$ in $\Re)$	Calculation without Coulomb interaction $(E$ in $\hbar\omega_c)$
0	$2.0068 + 3.9 \times 10^{-5} B$ kG	$-3.008 + 0.23\gamma$	$E_i + 0.5$
1	$2.0107 + 27 \times 10^{-5} B$	$-1.923 + 2.76\gamma$	$E_i + 1.5$
2	$2.0127 + 47 \times 10^{-5} B$	$-1.457 + 4.79\gamma$	$E_i + 2.5$
3	$2.0135 + 69 \times 10^{-5} B$	$-0.759 + 6.67\gamma$	$E_i + 3.5$
4	$2.008(?) + 10^{-3} B$	$-0.50 + 8.50\gamma$	$E_i + 4.5$
		$\Gamma = 3.5 \; \Re$ for these results	E_i depends on broadening and is $\approx \Gamma/10^{1/2}$

spectrum and due to the shape of the background on which each line is situated. We expect these to be unimportant, however, in any differential comparisons, since all the structure will shift together. The location of intercepts, in particular, will depend somewhat on the choice of broadening parameter.

Table III indicates the determination of the exciton rydberg from the intercepts when $\Gamma = 3.5 \; \Re$. The experimental differences are corrected by 1.0 meV in transforming from $\Delta R/R$ to $\Delta\epsilon_2$. The $n = 0$ line shifts by this amount in the transformation. The higher lines shift very little. Dividing by the calculated predictions gives a fairly consistent set of values for the rydberg whose average is 2.8 ± 0.6 meV. Using the value $\epsilon_0 = 15.5$ for the static dielectric constant[50] and the measured transverse mass $m/\mu_t = 19.7 \pm 1.3$ gives a value of 2.88 ± 0.2 meV. We use the static dielectric constant since $\hbar\omega_{LO} = 24$ meV[51] $\gg \Re$, $\hbar\omega_c$ for all magnetic fields in this experiment. For InSb this assumption is not very restrictive because the polaron coupling constant is small. At 60 kG we have $\gamma = 2.4 \pm 0.4$. The excitonic energy gap at the E_1 threshold is determined as $2.0068 + 3 \times 0.0028 = 2.015 \pm 0.001$ eV.

The comparison of experimental to theoretical peak energies vs magnetic field (slopes, Table IV) for both models is unfortunately inconclusive, since the $n = 0$ line does not seem to favor either model.

If we assume M_1 symmetry, we may use Eq. (36) to estimate $|M|$. A lower limit is obtained by taking $4\pi |\psi_n(0)|^2 = 2\gamma$ from the large-γ limit. We expect this to be low, since it overestimates the binding energy of the (000) ground state using the linear approximation in z, and since the linewidth due to Coulomb scattering $\lesssim 3.5$. Using $\gamma = 2.5$ and $\Gamma_{hw} = 3.5$ gives $|M| \gtrsim 15$. If we choose $4\pi |\phi(0)|^2$ to obtain the correct binding energy as calculated from the Born–Oppenheimer approximation, then $|M| \sim 100$.

The absolute order of magnitude of $\Delta\epsilon_2$ can be computed by calculating all the constants in Eq. (37) and is $\sim 10^{-2} NM^{1/2}$, where N is the number of equivalent valleys which contribute to the absorption. Experimentally, for $B = 60$ kG, $\Delta\epsilon_2 \sim 0.1$ so $NM^{1/2} \sim 10$. Since the Λ directions of the Brillouin zone are $\cos^{-1}(\frac{1}{3}) = 70.5°$ apart, excitons in those valleys whose major axis is not along k_z will be created in states whose structure will appear broadened out. Such a background may be due to aperiodic orbits,[52] as discussed earlier. This means we should take $N = 1$ at L and 2 in the interior of the zone. With the estimated value for M we get about the correct value for $NM^{1/2}$.

VIII. DISCUSSION ON GaSb MEASUREMENTS

GaSb is a material similar to InSb in that the E_1 and $E_1 + \Delta_1$ thresholds are well resolved be-

TABLE III. Determination of the exciton rydberg from intercepts.

Experiment $(\Delta R/R)$	Experiment $(\Delta\epsilon_2)$	Theory	Result of $\Re$
$E_1 - E_0 = 3.9$ meV ± 1	2.9	1.09 $(\Re)$	2.7 (meV)
$E_2 - E_0 = 5.9$ meV ± 1	4.9	1.55 $(\Re)$	3.1
$E_3 - E_0 = 6.7$ meV ± 1	5.7	2.25 $(\Re)$	2.5
			Av: 2.8 ± 0.6

TABLE IV. Comparison of models to experimental results.

n	With Coulomb interaction	Without Coulomb interaction
0	$\dfrac{3.9}{0.23} = 17$	$\dfrac{3.9}{0.5} = 7.8$
1	$\dfrac{27}{2.76} = 9.8$	$\dfrac{27}{1.5} = 18$
2	$\dfrac{47}{4.79} = 9.8$	$\dfrac{47}{2.5} = 19$
3	$\dfrac{69}{6.67} = 10.3$	$\dfrac{69}{3.5} = 20$
4	$\dfrac{100}{8.5} = 12$	$\dfrac{100}{4.5} = 22$
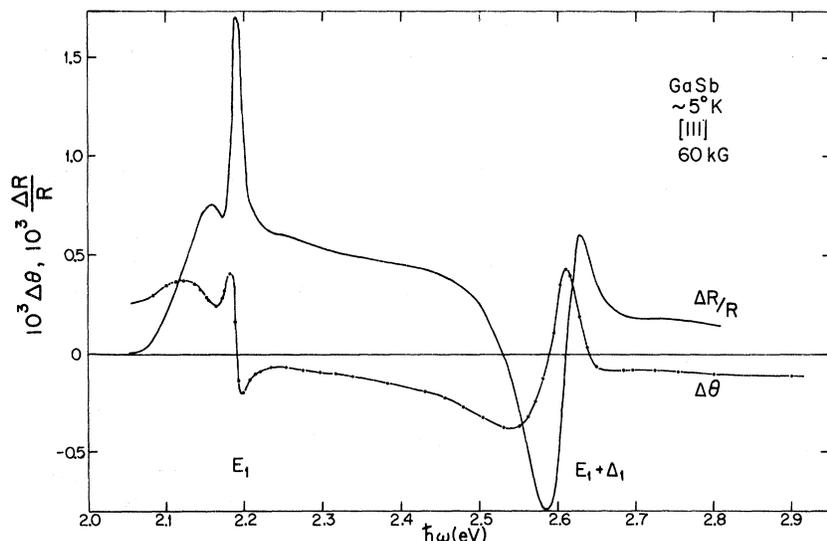
FIG. 17. GaSb magneto-optical spectrum taken at liquid-helium temperature. The dots show points at which the Kramers–Kronig integral is calculated.

cause of the large spin-orbit splitting. We attempted to look for Landau levels at the equivalent threshold in this material. Up to 62 kG they were not resolved. The measured signal was about $\frac{3}{4}$ of the InSb signal. Hence, if one assumes this ratio for the square of the interband momentum matrix element, then m/μ_t in GaSb is ~ 15, which means Landau levels should be seen for magnetic fields under 100 kG.

We show equivalent spectra for GaSb and InSb in Figs. 17 and 18. A peculiar feature of the GaSb spectrum is the extra structure appearing below the E_1 exciton line. The main exciton peak is at 2.189 eV and the subsidiary peak is at 2.156 eV. If this splitting is assumed due to inversion asym-

metry, then $\Delta_1^K = 33$ meV. We have checked to see that the splitting is at least second order in the magnetic field, but we have not observed a definite field dependence.

The difference in apparent amplitudes for the two lines can be accounted for in the model we have given. It is due to the large splitting of ϵ_2 by Δ_1^K. Under this condition the amplitude of the low-energy peak is lowered, since the tails of all the higher-lying structure must be subtracted in constructing the differential line shape. This interpretation for this peak is tentative. Neighboring critical points, for example, could lead to the same structure. It is puzzling that it is not observed in other modulation experiments.[53]
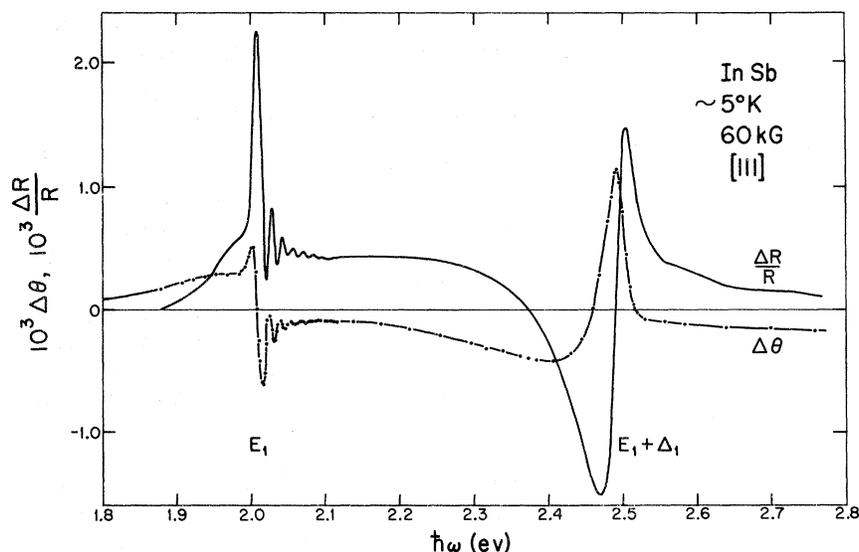
Fig. 18. Magneto-optical spectrum for InSb which may be compared to Fig. 17. The structure due to spin-orbit split-off valence bands $E_1 + \Delta_1$ has been included.

IX. CONCLUSIONS

A summary of information for InSb, which is determined from this experiment, is given in Table V.

At critical points of higher energy where effective masses may be larger than at the fundamental edge, the size of the Coulomb interaction relative to the cyclotron energy spacing becomes greater ($\propto \mu^2$). Hence, in order to interpret Landau-level transitions, Coulomb scattering between electrons and holes may need to be included. This is accomplished in an excitonic model by solving the effective Schrödinger equation including both Coulomb and magnetic effects. If the effective-mass tensor has cylindrical symmetry and the longitudinal mass is much larger than the transverse masses, then the Born–Oppenheimer approximation may be used to separate the Schrödinger equation when the magnetic field is in the longitudinal direction. In such a model the transverse Coulomb interaction and the magnetic field determine the relative amplitude and location of transitions, while the longitudinal Coulomb potential will determine primarily the line shape.

In materials lacking inversion symmetry it may prove possible to obtain information on the antisymmetric part of the periodic crystal potential. This asymmetry can warp energy bands by splitting the Kramers degeneracy and displace critical points from positions of high symmetry in the

TABLE V. New parameters at the E_1 edge of InSb determined from this experiment.

Quantity	Value	How obtained
$\hbar\omega_c(60\text{ kG}) = \hbar eB/\mu_t c$	13.7 ± 0.9 meV	Measured[a]
m/μ_t	19.7 ± 1.3	From $\hbar\omega_c$
μ_t/m	0.0508 ± 0.0032	
P^2	0.52 (au)	$\vec{k}\cdot\vec{p}$ analysis, Eqs. (46)[b]
m/m_{ct}	13.7	
m/m_{vt}	-6.0	
μ_c^*/μ_B	-0.36	
μ_v^*/μ_B	-6.0	
$\Re = \mu_t e^4/2\epsilon_0^2\hbar^2$	2.8 ± 0.6 meV	Measured,[c] and using a fit to exciton model
ϵ_0	15 ± 4	From $\Re$
$\gamma(60\text{ kG}) = \hbar\omega_c/2\Re$	2.4 ± 0.5	
$B(\gamma=1)$	25 ± 5 kG	
E_g	2.015 ± 0.001 eV	Measured, and using fit to exciton model
$|\mu_l/\mu_t|$	$\gtrsim 15$	Linewidth
Sign of μ_l/μ_t	?	Line-shape analysis
Δ_1^K	small (?)[d]	Line-shape analysis

[a] Average of spacing of higher-energy peaks of Landau-level spectrum. The error is the rms spread in the data used.

[b] May be compared to the zone-center value of 0.403 (a.u.) in Ref. 5; $\Delta_1 = 0.49$ eV was used in the computation.

[c] From intercepts in E-vs-B dependence of Landau energies. Compare to 2.98 ± 0.2 meV calculated using the measured transverse mass and $\epsilon_0 = 15.5$ from Ref. 50.

[d] $\ll 20$ meV of Ref. 31.

Brillouin zone. The fact that it exists can provide a means for deciding whether some critical points are interior to a zone boundary.

APPENDIX A

At the valence-band edge we introduce the spatial Bloch functions, X, $Y \sim e^{i\vec{k}_0\cdot\vec{r}} u_{\vec{k}_0}(r)$. These wave functions belong to the representation Λ_3 of C_{3v} and transform as x, y. The spin-orbit interaction

$$H_{so} = (\hbar/4m^2c^2)(\vec{\nabla}V\times\vec{p})\cdot\vec{\sigma} = \vec{L}\cdot\vec{\sigma} \qquad (38)$$

may then be diagonalized using as a basis set the product wave functions $X\alpha$, $X\beta$, $Y\alpha$, $Y\beta$. $\vec{L}$ has been defined as $(\hbar/4m^2c^2)(\vec{\nabla}V\times\vec{p})$ and transforms as an axial vector. Under conditions of spherical symmetry $\vec{L}$ becomes proportional to the orbital angular momentum.

Under C_{3v} symmetry there are two distinct non-zero matrix elements of $\vec{L}$:

$$\langle X|L_z|Y\rangle = i\,\text{Im}\,\langle X|L_z|Y\rangle$$

and $\qquad\qquad\qquad\qquad\qquad\qquad\qquad (39)$

$$\langle X|L_y|X\rangle = -\langle Y|L_y|Y\rangle = \langle X|L_x|Y\rangle = \text{Re}\,\langle X|L_x|Y\rangle \,.$$

The usual spin-orbit splitting in the [111] direction is just

$$\Delta_1 = 2\,\text{Im}\,\langle X|L_z|Y\rangle \,. \qquad (40)$$

This assumes that $\vec{k}_0$ is far enough from the zone center that the $\vec{k}\cdot\vec{p}$ splitting of the orbital degeneracy is larger than the spin-orbit energy. At the zone center the spin-orbit splitting[31]

$$\Delta_0 = 3\,\text{Im}\,\langle X|L_z|Y\rangle \,. \qquad (41)$$

In this case X and Y are Bloch states at the zone center which transform as x and y under tetrahedral operations. Since $\vec{\nabla}V$ and $\vec{p}$ are largest near the atomic cores of the crystal lattice, the bulk of the contribution to the matrix elements in Eqs. (40) and (41) occurs there. Hence we expect the spin-orbit splittings to be simply related to ionic spin-orbit parameters.[54] Moreover, the spin-orbit interaction is then rather insensitive to k vector, and to this extent we have a $\frac{2}{3}$ rule.[8,54]

The splitting $\Delta_1^K = 4\,\text{Re}\,\langle X|L_x|Y\rangle$, on the other hand, does depend on the k vector. It will only be nonzero due to the antisymmetric part of the

periodic crystal potential V, which vanishes when there is inversion symmetry. Symmetry arguments also tell us that whenever $X^* = TX = X$ and $Y^* = TY = Y$, Δ_1^K must vanish. Here T is the time-reversal operator which is antiunitary so that

$$\langle X | \vec{L} | Y \rangle^* = -\langle TX | \vec{L} | TY \rangle , \qquad (42)$$

where $T\vec{L}T^{-1} = -\vec{L}$. X and Y cannot be real at a general point along the [111] direction in the Brillouin zone, where a Bloch function at $\vec{k}_0$ is transformed to one at $-\vec{k}_0$. However, they are at the zone center or zone boundary. If we include inversion symmetry (I)

$$\langle X | \vec{L} | Y \rangle^* = -\langle ITX | \vec{L} | ITY \rangle , \qquad (43)$$

since $\vec{L}$ is invariant to inversion. Now Δ_1^K must vanish identically, since the combined operation IT transforms spatial Bloch wave functions into themselves.

The valence-band edge-states resulting from the spin-orbit interaction diagonalization are

$$U_4 = \tfrac{1}{2}[-i(X+iY)\alpha + (X-iY)\beta] ,$$

$$U_5 = \tfrac{1}{2}[(X+iY)\alpha - i(X-iY)\beta]$$

and   (44)

$$U_{6\uparrow} = 2^{-1/2}(X+iY)\beta ,$$

$$U_{6\downarrow} = 2^{-1/2}(X-iY)\alpha .$$

The latter two wave functions of Λ_6 symmetry are degenerate; U_4 and U_5 are split by Δ_1^K, and $U_{4,5}$ and $U_{6\uparrow,\downarrow}$ are split by Δ_1.

The $\vec{k}\cdot\vec{p}$ effective masses and magnetic moments in the approximation in which only $P = \langle S|p_x|X\rangle = \langle S|p_y|Y\rangle$ is nonzero may be calculated from[35]

$$\left(\frac{m}{m^*}\right)^{\mu\nu} = \delta_{\mu\nu} + \frac{2}{m}\sum_{l\neq i}\mathrm{Re}\left(\frac{\langle i|p^\mu|l\rangle\langle l|p^\nu|i\rangle}{E_i - E_l}\right) , \qquad (45)$$

$$\mu^*\rho = \frac{e}{2mc} + \frac{e}{m^2 c}\,\epsilon^{\mu\nu\rho}$$

$$\times \sum_{l\neq i}\mathrm{Im}\left(\frac{\langle i|p^\mu|l\rangle\langle l|p^\nu|i\rangle}{E_i - E_l}\right),$$

where $\epsilon^{\mu\nu\rho} = 1$ if $\mu\nu\rho$ are cyclic, and zero otherwise.

The results which we use are

$$\left(\frac{m}{m^*}\right)_{c\uparrow} = 1 + \frac{p^2}{m}\,\frac{2E_g + \Delta_1}{E_g(E_g + \Delta_1)} ,$$

$$\left(\frac{m}{m^*}\right)_{4,5\uparrow} = 1 - \frac{p^2}{mE_g} ,$$

$$\left(\frac{m}{m^*}\right)_{6\uparrow} = 1 - \frac{p^2}{m(E_g + \Delta_1)} ,$$

$$\left(\frac{\mu_c^*}{\mu_B}\right)_z = 1 - \frac{p^2}{m}\,\frac{\Delta_1}{E_g(E_g + \Delta_1)} ,$$

$$\left(\frac{\mu_v^*}{\mu_B}\right)_{4,5z} = \left(\frac{m}{m^*}\right)_{4,5\uparrow} ,$$

$$\left(\frac{\mu_v^*}{\mu_B}\right)_{6z} = \left(\frac{m}{m^*}\right)_{6\uparrow} , \qquad (46)$$

where μ_B is the Bohr magneton ($= e\hbar/2mc$). Corrections of order Δ_1^K/E_g are ignored.

The eigenstates obtained by diagonalizing H_{spin} in (18) are just

$$U_{4m} = (1/D)\{[\mu_v^*B + [(\mu_v^*B)^2 + (\tfrac{1}{2}\Delta_1^K)^2]^{1/2}]U_{4\downarrow} + i(\tfrac{1}{2}\Delta_1^K)U_{5\uparrow}\},$$

$$U_{5m} = (1/D)\{-\tfrac{1}{2}i\Delta_1^K U_{4\downarrow} + (\mu_v^*B + [(\mu_v^*B)^2 + (\tfrac{1}{2}\Delta_1^K)^2]^{1/2})U_{5\uparrow}\}$$
(47)

where

$$D = \{(\mu_v^*B + [(\mu_v^*B)^2 + (\tfrac{1}{2}\Delta_1^K)^2]^{1/2})^2 + (\tfrac{1}{2}\Delta_1^K)^2\}^{1/2} .$$

For $\mu_v^*B \gg \Delta_1^K$, $U_{4m} \to U_{4\downarrow}$ and $U_{5m} \to U_{5\uparrow}$. In the opposite limit we recover Eq. (44).

In optical transitions these states interact with the radiation field of light. If $\Delta_1^K = 0$, the situation is particularly simple: A spin $\uparrow(\downarrow)$ electron in the valence-band state $U_{5\uparrow}(U_{4\downarrow})$ is promoted to the conduction-band state $U_{c\uparrow}(U_{c\downarrow})$ leaving behind a hole. The transition matrix element for circularly polarized radiation $\langle C|\hat{\epsilon}_\pm\cdot\vec{p}|V\rangle$ is

$$\tfrac{1}{2}\langle S|p_x \pm ip_y|X \mp iY\rangle = \langle S|p_x|X\rangle .$$

Hence, following Eq. (23),

$$\epsilon_{2\pm} = \epsilon_2(\omega \pm \mu^*B)$$

$$= \frac{4\pi^2 e^2}{\omega^2 m^2}\sum_N |\psi_N(0)|^2 |P|^2\,\delta(\hbar\omega - E_N \pm \mu^*B) , \qquad (48)$$
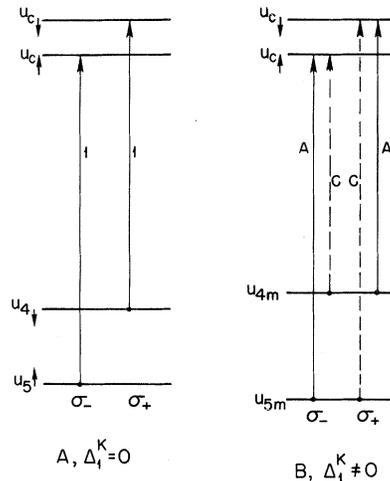
FIG. 19. Transitions between the valence and conduction bands for $\Delta_1^K = 0$ and for $\Delta_1^K \neq 0$.
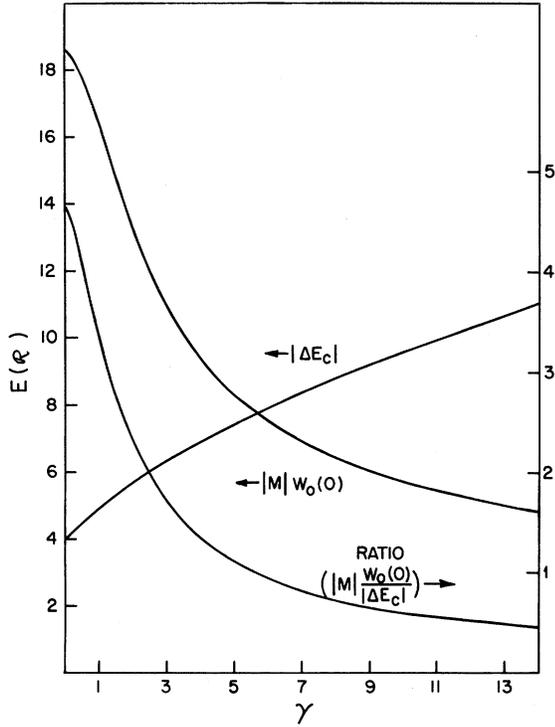
FIG. 20. Computation of $W_0(0)$ for small γ and comparison to the Coulomb contribution to the ground-state energy. The result for $W_0(0)$ was obtained by summing over the first fifteen terms in Eq. (53) and estimating the remainder by an appropriate integral.

where the effective magnetic moment $\mu^* = \mu_c^* - \mu_v^*$. If $\Delta_1^K \neq 0$, we define

$$A = \{(\mu_v^* B + [(\mu_v^* B)^2 + (\tfrac{1}{2}\Delta_1^K)^2]^{1/2})/D\}^2$$

and $C = (\Delta_1^K/2D)^2$. Then Eq. (48) becomes

$$\epsilon_{2+} = \frac{4\pi^2 e^2}{m^2 \omega^2} \sum_N |\psi_N(0)|^2 |P|^2$$

$$\times [A \, \delta(\hbar\omega - E_N \pm \mu_1 B) + C \, \delta(\hbar\omega - E_N \pm \mu_2 B)] , \quad (49)$$

For $z \to 0$, note that

$$\rho \frac{d}{dz} \left(\frac{1}{(\rho^2 + z^2)^{1/2}} \right) = \frac{z\rho}{(\rho^2 + z^2)^{3/2}} \to \delta(\rho - 0^+) ,$$

so that

$$W_0(0) = \frac{4}{M} \sum_{n \neq 0} \left| \frac{2\pi \phi_n(0) \phi_0(0)}{E_n(0) - E_0(0)} \right|^2 . \quad (53)$$

where

$$\mu_1 B = \mu_c^* B - [(\mu_v^* B)^2 + (\tfrac{1}{2}\Delta_1^K)^2]^{1/2}$$

and

$$\mu_2 B = \mu_c^* B + [(\mu_v^* B) + (\tfrac{1}{2}\Delta_1^K)^2]^{1/2} .$$

These cases are illustrated in Fig. 19. When B is reversed, both $\mu_1 B$ and $\mu_2 B$ reverse sign although this is not immediately apparent in the square root factor. This is essential if $\epsilon_{2+} - \epsilon_{2-}$ is to be odd in B. The difficulty is that diagonalizing Eq. (17) produces a sign ambiguity, which can be recognized by examining the eigenvalues when $\Delta_1^K \to 0$.

Since $\Delta_1 \ll E_g + \Delta_1$, the results (46) show that $|\mu_c^*| \ll |\mu_v^*|$. Hence $\mu_1 \approx -\mu_2$ so that

$$\Delta\epsilon_2 = \epsilon_{2+} - \epsilon_{2-} = (A - C)[\epsilon_2(\omega + \mu_1 B) - \epsilon_2(\omega - \mu_1 B)] ,$$

$$(50)$$

where ϵ_2 is given by (48). We emphasize the difference $\epsilon_2(\omega + \mu_1 B) - \epsilon_2(\omega - \mu_1 B)$. It is only equal to $(2/\hbar)(\partial\epsilon_2/\partial\omega)\mu_1 B$ when $\mu_1 B$ is small compared to the energy characterizing the structure of ϵ_2 (the cyclotron energy, for example); otherwise we expect higher-order terms to be important.

APPENDIX B

The effect of the kinetic energy of the z dependence of the transverse motion $W_n(z)$ may be considered by differentiating the expressions (20a) or (21a) for the transverse motion with respect to z. An inhomogeneous equation results which may be formally solved by introducing the Green's function for the problem. Since $\phi_n(\rho z)$ must remain normalized,

$$\frac{d}{dz} \int d\vec{\rho} \, \phi_n^2(\rho z) = 2 \int d\vec{\rho} \, \phi_n(\rho z) \frac{d\phi_n}{dz} (\rho z) = 0 , \quad (51)$$

which means that there is no homogeneous part to the solution. The result for $W_0(z)$ is

$$W_0(0) = \frac{4}{M} \sum_{n \neq 0} \left| \frac{2\pi \int_0^\infty d\rho \, \rho \phi_n(\rho z) \, \phi_0(\rho z) \, (d/dz)[1/(\rho^2 + z^2)^{1/2}]}{E_n(z) - E_0(z)} \right|^2 (\Re) . \quad (52)$$

This is in agreement with perturbation theory.

Our main interest in investigating the term $W_0(z)$ is to check that it is "small" compared to the main contribution of the Coulomb potential. Hence we have computed $|M|W_0(0)$ and compare it to the Coulomb contribution $\Delta E_c = \gamma - E_0(0)$ to the total energy (Fig. 20).

For large γ the ratio goes as $\gamma^{-1/2}$. This may

be verified by using perturbation theory and high-field results:

$$\phi_n(\rho) = \frac{1}{2^{1/2}\lambda n!}\, e^{-\rho^2/4}\, L_n(\tfrac{1}{4}\rho^2) \,,$$

$$|\phi_n(0)|^2 = 1/2\pi\lambda^2 = \gamma/2\pi a^2 \,. \tag{54}$$

Small z:

$$E_n(z) = (2n+1)\gamma - 4\pi\gamma^{1/2}\int_0^\infty d\rho\, |\phi_n(\rho)|^2$$

$$+ 4\pi|\phi_n(0)|^2|z|(\Re) \,, \tag{55}$$

$$|M|W_0(0) = \tfrac{1}{6}\pi^2(\Re) \,.$$

Large z:

$$E_n(z) = (2n+1)\gamma - 2/|z| + O(1/|z|^3)(\Re) \,.$$

L_n are Laguerre polynomials.

However, when $\gamma \sim 6$ or less, which is the region of interest to us, $|M|W_0(0)$ becomes larger than the Coulomb energy ΔE_c. The worst case is for zero field when $|M|W_0(0)/\Delta E = 4.65$. Since W_0 is in a sense a correction term this indicates poor convergence for small mass ratios $|M|$. For the Born–Oppenheimer approximation to be valid to 1%, say, we expect W_0 to have to be of order 1% so $|M|$ would have to be > 500. Numerical computations show the actual situation to be somewhat better if the W_n term is omitted entirely. This is what we have done. The eigenenergies for the potential $V(z)$ will not then be proper upper bounds on the true energies but will approximate them to order 10% or better. Take a worse case: The ground state for $B = 0$ and $|M| = 1$ is computed as $-1.127(\Re)$ which is 13% low. If $W_0(z)$ is included in the computation, the eigenenergy obtained is high but much worse.

†Work supported by the National Science Foundation and a grant from the E. I. duPont de Nemours Co., Inc.

*Present address: Department of Physics, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

[1]G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. <u>92</u>, 827 (1953).

[2]R. J. Keyes, S. Zwerdling, S. Foner, H. Holm, and B. Lax, Phys. Rev. <u>104</u>, 1804 (1956); E. Burstein, G. Picus, H. Gebbie, and F. Blatt, *ibid.* <u>103</u>, 825 (1956).

[3]B. Lax and J. G. Mavroides, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1960), Vol. 11, p. 261; B. Lax and J. G. Mavroides, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1967), Vol. 3.

[4]R. L. Aggarwal, Phys. Rev. B <u>2</u>, 446 (1970); M. Reine, R. Aggarwal, B. Lax, and G. Wolfe, *ibid.* <u>2</u>, 458 (1970).

[5]C. R. Pidgeon and R. N. Brown, Phys. Rev. <u>146</u>, 575 (1966).

[6]Y. Nishina and B. Lax, J. Appl. Phys. Suppl. <u>32</u>, 2128 (1961); A. Moore, quoted by M. Cardona, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1967), Vol. 3.

[7]S. O. Sari, Phys. Rev. Letters <u>26</u>, 1167 (1971).

[8]M. Cardona, *Modulation Spectroscopy* (Academic, New York, 1969).

[9]K. L. Shaklee, J. E. Rowe, and M. Cardona, Phys. Rev. <u>174</u>, 828 (1968); Y. Hamakawa, F. Germano, and P. Handler, J. Phys. Soc. Japan Suppl. <u>21</u>, 111 (1966).

[10]G. Dresselhaus, A. F. Kip, C. Kittel, and G. Wagoner, Phys. Rev. <u>98</u>, 556 (1955).

[11]R. L. Aggarwal, Bull. Am. Phys. Soc. <u>12</u>, 100 (1967).

[12]S. N. Jasperson and S. E. Schnatterly, Rev. Sci. Instr. <u>40</u>, 761 (1969).

[13]A. B. Callender, Ph.D. thesis (Princeton University, 1970) (unpublished); R. E. Palmer, Ph.D. thesis (Princeton University, 1971) (unpublished).

[14]S. Sari, Ph.D. thesis (Princeton University, 1971) (unpublished).

[15]Sample data: InSb: undoped, $n = 1.7\times10^{14}$ cm^{-3}, $\rho = 0.07$ Ω cm. GaSb: undoped, $n = 1.9\times10^{17}$ cm^{-3}, $\rho = 0.06$ Ω cm.

[16]H. C. Gatos and M. C. Lavine, in *Progress in Semiconductors*, edited by A. F. Gibson and R. E. Burgess (Temple, London, 1965), Vol. 9, p. 1. InSb and GaSb were etched with 1:10 I$_2$:CH$_3$OH for ~ 40 sec. HNO$_3$:HF: CH$_3$COOH 25:15:40 (dilute CP4) for $\sim 15-30$ sec.

[17]I. M. Boswarwa, R. E. Howard, and A. B. Lidiard, Proc. Roy. Soc. (London) A269, 125 (1962); see also R. Z. Bachrach and F. C. Brown, Phys. Rev. B <u>1</u>, 818 (1970).

[18]G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. <u>100</u>, 618 (1955).

[19]The fact that the measurements are near normal will not substantially affect the results, since changes thereby introduced will be second order in the angle of incidence. The off-axis problem is considered by W. Voigt, *Magneto und Elektrooptik* (Teubner, Leipzig, 1908); and also in Ref. 14.

[20]S. E. Schnatterly, Phys. Rev. <u>183</u>, 664 (1969).

[21]H. R. Phillipp, quoted by B. O. Seraphin and H. E. Bennett, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1967), Vol. 3, p. 535.

[22]R. Zucca and Y. R. Shen, Phys. Rev. B <u>1</u>, 2668 (1970).

[23]R. J. Elliott, Phys. Rev. <u>108</u>, 1384 (1957).

[24]T. K. Bergstresser and M. Cohen, Phys. Rev. <u>141</u>, 789 (1966)

[25]M. Cardona and F. Pollak, Phys. Rev. <u>142</u>, 530 (1966).

[26]G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. <u>98</u>, 368 (1955).

[27]J. M. Luttinger and W. Kohn, Phys. Rev. <u>97</u>, 869 (1955).

[28]J. M. Luttinger, Phys. Rev. <u>102</u>, 1030 (1956).

[29]L. M. Roth, B. Lax, and S. Zwerdling, Phys. Rev. <u>114</u>, 90 (1959).

[30]E. O. Kane, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), Vol. 1, p. 75.

[31]G. Dresselhaus, Phys. Rev. 100, 580 (1955).

[32]E. O. Kane, J. Phys. Chem. Solids 1, 249 (1957).

[33]The reader should compare Fig. 5 of Ref. 31 and Fig. 2 of R. J. Elliott, Phys. Rev. 96, 266 (1954).

[34]Reference 8, p. 71, also see our Appendix A.

[35]M. Cardona and D. L. Greenaway, Phys. Rev. 125, 1291 (1962).

[36]R. S. Knox, *Theory of Excitons* (Academic, New York, 1963).

[37]E. O. Kane, Phys. Rev. 180, 852 (1969); Phys. Rev. B 2, 4305 (1970).

[38]J. D. Dow and D. Redfield, Phys. Rev. B 1, 3358 (1970).

[39]H. I. Ralph, J. Phys. C 1, 378 (1968).

[40]D. F. Blossey, Phys. Rev. B 3, 1382 (1971).

[41]Y. Yafet, R. Keyes, and E. Adams, J. Phys. Chem. Solids 1, 137 (1956).

[42]D. M. Larsen, J. Phys. Chem. Solids 29, 271 (1968).

[43]R. J. Elliott and R. Loudon, J. Phys. Chem. Solids 15, 196 (1960); 8, 382 (1959).

[44]A. Baldereschi and F. Bassani, in *Proceedings of the Tenth International Conference on Physics of Semiconductors, Cambridge, Mass.,* 1970 (U. S. AEC, Oak Ridge, Tenn., 1971), p. 191.

[45]Y. Petroff and M. Balkanski, Phys. Rev. B 3, 3299 (1971).

[46]J. W. Cooley, Math. Tabl. Comput. 15, 363 (1961).

[47]B. Velicky and J. Sak, Phys. Status Solidi 16, 147 (1966).

[48]M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

[49]P. F. Edwards and V. J. Lazazzera, Phys. Rev. 120, 420 (1960).

[50]M. V. Hobden, Phys. Letters 16, 107 (1965).

[50]Reference 8, p. 37.

[51]D. M. Larsen, in Ref. 44, p. 145.

[52]Excluding Coulomb effects, this will occur for $\theta > \tan^{-1}|M|^{1/2} \approx 76°$. See, A. Baldereschi and F. Bassani, Phys. Rev. Letters 19, 66 (1967).

[53]C. W. Higginbotham, F. H. Pollak, and M. Cardona, *Proceedings of the Ninth International Conference on the Physics of Semiconductors, Moscow,* 1968, edited by S. M. Ryvkin (Nauka, Leningrad, 1968), p. 57.

[54]M. Cardona, *Semiconductors and Semimetals,* edited by R. Willardson and A. Beer (Academic, New York, 1967), Vol. 3.

Spectral Distribution of the Photomagnetoelectric Circulating Current in Semiconductors

Jan Hlávka

*Department of Solid State Physics, University J. E. Purkyně,
Brno, Czechoslovakia*

(Received 17 November 1971)

The investigation of the spectral distribution of the photomagnetoelectric (PME) circulating current is made possible by the new method of measurement based on the magnetic flux created by this current. The formula for PME-circulating-current magnetic flux is derived and the results of calculations of the magnetic flux as a function of absorption coefficient are presented in graphical form. The measured spectral distributions are in good qualitative agreement with theoretical distributions.

INTRODUCTION

The idea has been introduced that the so-called circulating current flows in a semiconductor sample when the photomagnetoelectric (PME) effect takes place.[1,2] The photomagnetomechanical effect gives us some information about PME circulating current.[3] This effect is not suitable for measurement of spectral distribution because of a spurious effect caused by the diamagnetism of the sample. The new method enabling measurement of spectral distribution is based on the following idea: Circulating current can be understood as a current winding. The PME circulating current is created and disappears in a sample placed in a magnetic field when the sample is illuminated by a chopped radiation; the magnetic flux created by this current induces a measurable voltage in coils.

Results presented here are the first direct proof of the existence of the PME circulating current.

MAGNETIC INDUCTION FLUX OF PME CIRCULATING CURRENT

The PME circulating current has been shown[2] to flow in a sample as illustrated in Fig. 1(a). If a long enough sample or a sample with end electrodes is considered the current flow approaches the case illustrated in Fig. 1(b).

The current flow illustrated in Fig. 1(b) is considered in our derivation. The resulting magnetic flux is a sum of the magnetic flux of elementary current windings as in Fig. 2. The magnetic flux of the elementary winding is divided into two parts, a part for $y < y_0$ and a part for $y > y_0$, and each part is expressed separately. We have

$$d\phi = \frac{J_x^{(oc)}(y)\,v}{R + R_0}\,dy ,\qquad (1)$$