Electrical Conduction in *n*-Type Germanium at Low Temperatures

SEYMOUR H. KOENIG*
IBM Research Laboratory, Adliswil-Zurich, Switzerland

AND

RODNEY D. BROWN, III, AND WALTER SCHILLINGER
IBM Watson Research Laboratory, Columbia University, New York, New York

(Received May 14, 1962; revised manuscript received August 24, 1962)

A thorough study has been made of the electrical conductivity (including Hall effect) of high-purity *n*-type Ge in the temperature range 4–25°K. Both the Ohmic region and non-Ohmic or "hot electron" region were studied, the latter by both dc and high-speed pulse techniques. From measurements in the Ohmic region, accurate values for the impurity concentrations, in particular the compensating impurities, and for the donor activation energies are obtained; these must be known in order to convert the pulse data to cross sections for certain elementary processes that were investigated. As a subsidiary result, it is confirmed that for P and As donors there is a difference between the thermal and optical activation energies which is proportional to the singlet-triplet splitting of the ground state. By comparing experimental results for the electric field dependence of the drift mobility with the (hot electron) theory for acoustic phonon scattering including anisotropy, it is shown that the electrons appear to be hotter at lower temperatures than the theory would predict. One must conjecture a "forward scattering effect" that reduces the average collision rate for hot electrons moving in the heavy-mass direction, i. e., they can scatter only through small angles at low temperatures because of a dearth of the appropriate phonons. The conjecture is confirmed by a measurement of the mobility anisotropy parameter K (=4.2 for phonon scattering), obtained from data on the magnetic field dependence of the breakdown field. The major aim of the work was to determine values for the (velocity- and temperature-dependent) cross sections for two recombination processes: the recombination of a single electron with an ionized donor (the inverse of thermal ionization) and the Auger recombination of an electron with an ionized donor (the inverse of ionization of a donor by collision). Cross sections have been determined for these processes, and for the inverse processes as well. Criticisms made of some of our early work by Ascarelli and Brown are shown to be invalid. In particular the *ad hoc* assumption by Ascarelli and Brown that Auger processes do not exist is what led them to false conclusions; their data, reinterpreted, give values for the Auger recombination probability in agreement with ours. Both the direct and Auger recombination cross sections are very large ($\sim 10^{-12}$ cm^2 and $\sim 10^{-24}\ n$ cm^2, respectively); they agree well, as regards absolute magnitude, and temperature and energy dependence, with the "giant trap" cascade mechanism proposed first by M. Lax. Moreover, the inverse of the giant-trap mechanism would be a mechanism for impact ionization of neutral impurities that would have a rapid dependence on lattice temperature and would have no threshold for the energy of the colliding electron. This has been observed. Lastly, the hot electron behavior for orientations of the applied electric field other than along a ⟨100⟩ direction have been studied; the electrons in different valleys then have different "temperatures" and densities. It is shown that, though this is the case, no net transfer of carriers from one valley to the other occurs. A simple theoretical treatment for the breakdown field in terms of impurity content and for the variation of the breakdown field with crystallographic orientation and with magnetic field has been formulated, using as a basis Price's criterion for breakdown, which is in excellent agreement with the data.

I. INTRODUCTION

THE major objective of the research reported here was to experimentally determine the cross sections for various kinetic processes that, in turn, determine the temperature and electric field dependent conduction band carrier density in *n*-type Ge below ~20°K. Preliminary results have been reported,[1–3] along with a general description of the experimental procedure involved. To improve both the accuracy and reliability of the early results and to resolve some apparent disagreement in the literature,[4] it was necessary to extend the early results and to perform many ancillary experiments. Several of these proved to be interesting in their own right; all the results will be presented here.

The type of measurements made on the germanium samples involve: (a) dc measurements (usually at very high impedance levels[5]) of the linear conductivity and the Hall effect; (b) pulsed-dc Hall and conductivity measurements to further extend the range of data without undue heating of the sample; and (c) measurements of the time dependence of the conductivity upon application of fast rise-time pulses. The set (c) above yields values for the rates of the kinetic processes that it is desired to measure[1,2]; the set (a) provides information on the concentration of impurities that is required to convert the kinetic data to average cross sections; the data (b) yield information on the variation of the electron velocity distribution with electric field (the "hot electron" problem) as well as providing verification for the model used in interpreting the data (c).

The bulk of the measurements reported here were taken on two samples cut from different regions of one single, nominally antimony-doped crystal of germanium. The acceptor concentration in this crystal is believed lower than any previously reported in the

* Permanent address: IBM Watson Research Laboratory, Columbia University, New York 25, N. Y. All experimental work reported here was performed at the New York laboratory.

[1] S. H. Koenig, Phys. Rev. **110**, 986 (1958).

[2] S. H. Koenig, Phys. Rev. **110**, 998 (1958).

[3] S. H. Koenig, J. Phys. Chem. Solids **8**, 227 (1959).

[4] G. Ascarelli and S. C. Brown, Phys. Rev. **120**, 1615 (1960); *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961) p. 271.

[5] See S. H. Koenig and G. R. Gunther-Mohr, J. Phys. Chem. Solids **2**, 268 (1957).

literature, so that the accuracy of many measurements and the certainty of their interpretation are optimized (as will be clear from subsequent sections). The donor concentrations for the two samples differed by a factor of about 2, while their acceptor concentrations were roughly equal. The data are, however, completely consistent with results (both published and unpublished) on a large number of other crystals from different sources, so that there is no reason to believe that the results are peculiar to this crystal.

II. LINEAR CONDUCTIVITY AND IMPURITY CONCENTRATIONS

A. Apparatus; Samples

The cryostat used is shown schematically in Fig. 1. The sample (A) is enclosed in a light-tight isothermal copper container (B), which, in turn, is contained within a brass outer jacket (C), on the outside of which is wound a niobium solenoid. The entire assembly is inserted into a 30-liter Superior helium storage vessel with a $1\frac{1}{4}$-in. neck. A heater (D) is mounted on the outside of the isothermal container. Helium gas at ~ 2 mm pressure in the space (E) provides limited thermal contact between (B) and (C). Three watts of heater power, which corresponds to a liquid-helium loss rate of ~ 3 liters/h, is sufficient to maintain the sample at 25°K. With no power input the loss rate is $\sim\frac{1}{3}$ liter per day, so that under typical conditions, one filling of the Dewar suffices for 1–$1\frac{1}{2}$ months of experimentation, even allowing for removal of the apparatus from the helium several times.

Thermal contact between the sample and the container (B) is made both by the exchange gas and by mounting the sample with silicone grease on a sapphire block which is similarly attached to (B). The temperature is measured by a $\frac{1}{10}$-w Allen-Bradley 1000-Ω carbon resistor (F) mounted deep within the copper body of (B) and used as a resistance thermometer. Typical thermal time constants range from tenths of a second below ~ 7°K to several seconds at the higher temperatures. The resistance of the thermometer is determined by the balance of a 30-cps bridge. The sensitivity of the system at the lower temperatures is such that temperature variations of one millidegree can be detected with less than 10^{-6} W dissipation in the thermometer.

Our experience from the reproducibility of the electrical data on the sample has been that with repeated temperature cycling of the thermometer, reasonable variation of exchange gas pressure, various locations and mountings of the sample, soldering or not soldering the cap (B) to the body, etc., the temperature of the sample is reproducible with a maximum scatter of no more than ± 25 mdeg below ~ 12°K. At higher temperatures, the electrical parameters of the sample become less temperature-sensitive.

The samples used were prepared from x-ray oriented

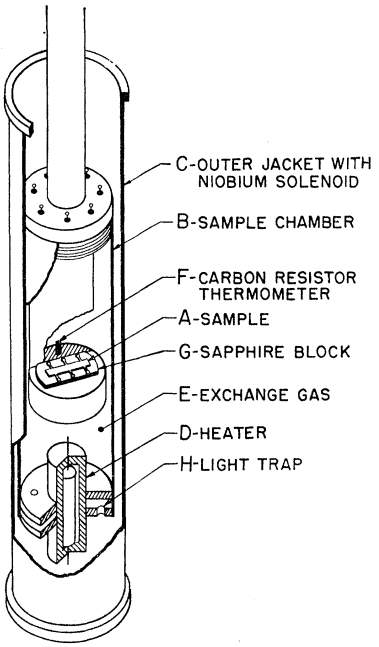


FIG. 1. Cut-away view of the apparatus.

slices of single-crystal germanium cut into a bridge shape with an ultrasonic cutter. The steady-state dc measurements on the sample involved the measurement of the potential difference between either two potential arms or the two Hall arms as a function of current through the sample. The voltage measurements were made using two Applied Physics Corporation vibrating reed electrometers connected as a differential voltmeter. The current was measured by reproducing the voltage drop across a high-impedance standard resistor with a high-gain, unity-feedback dc amplifier that used an electrometer tube for its input stage. The current was supplied by the output of a motor driven Helipot fed by an adjustable battery source. The outputs of the ammeter and voltmeter were recorded by a Leeds and Northrup X-Y recorder. The limitations on the accuracy of the measurements were 0.5% due to the ammeter feedback resistors and 0.3% imposed by the recorder, except for the Hall voltage at low electric and/or magnetic fields, which was often limited by noise to $\sim 1\%$. The absolute values for quantities derived from the measurements are limited in addition by uncertainties in sample geometry of $\sim 1\%$.

The superconducting solenoid is wound of Formvar-coated, 0.005-in., niobium wire. A magnetic field of ~ 2500 G may be obtained when no heat is applied to the sample container. When the sample is maintained at ~ 15°K, the heat flow is sufficient to heat the inner layer of the solenoid so the maximum field then obtainable is reduced to ~ 1000 G. The calibration constant of the magnet was determined from the geometry of the solenoid and a measurement, with a traveling microscope, of the linear turns density of a single layer. This procedure should give the field per unit current

to much better than 1%. An accidental independent check of this procedure was had when a second solenoid had to be wound after a short developed in the first. Overlapping sets of data agreed within the stated experimental uncertainty.

B. General Considerations

The main purpose in measuring the conductivity σ and Hall constant R in the "Ohmic" range (electric fields $\lesssim 0.1$ V/cm) is to obtain values for the donor and acceptor concentrations. For samples in the range of concentration considered here, the principle contributions to the scattering of carriers are from acoustic phonons and ionized impurities. Below ~ 10°K, the only ionized impurities are the acceptors and the donors they compensate, so that a measurement of the Hall mobility $\mu_H = R\sigma$ allows one to order different samples according to their acceptor concentration. However, as the inadequacy of current theories of impurity scattering is most pronounced at low temperatures,[6] quantitative values for N_A cannot be obtained by this procedure.

An accurate way of determining the impurity concentration (in principle) is by fitting the temperature-dependent Hall constant data to the appropriate statistical formula. The expression for the temperature and impurity dependence of the carrier density n in a "multi-ellipsoid" conduction band such as is appropriate to germanium with only one type donor impurity is[7]

$$\frac{n(n+N_A)}{(N_D-N_A-n)}=\frac{2(2\pi m^* kT)^{\frac{3}{2}}\nu}{h^3\sum_i g_i \exp(-\mathcal{E}_i/kT)}, \qquad (1)$$

where N_D is the donor concentration, N_A is the volume density of acceptor states (e.g., for a triple acceptor such as copper, N_A would be three times the copper density), ν is the number of equivalent conduction band minima, $\mathcal{E}_i$ is the energy of a bound electronic state of the donor measured with respect to the band edge, g_i is its degeneracy (including spin), and m^* is the density-of-states mass (the geometric mean of the three principle value masses for one valley). The summation is over all bound donor electronic states. Equation (1) is valid so long as the Fermi level is several kT below the conduction band edge, which is the case for all data reported here (and almost always the case in reasonably pure material at all temperatures). In the limit of large T, the "extrinsic region," $n=N_D-N_A$.

For temperatures $\lesssim 10$°K, only the singlet and triplet (excluding spin) ground-state multiplets, which are formed because of central core effects (deviations from the effective mass approximation), need be included in the summation for Sb-doped germanium.[8–10] Equation (1) then simplifies to

$$\frac{n(n+N_A)}{(N_D-N_A-n)}=\frac{4(2\pi m^* kT)^{3/2}\exp(-\epsilon_2 kT)}{h^3[3+\exp(\delta\epsilon/kT)]}. \qquad (2)$$

Here $\delta\epsilon$ is the singlet-triplet separation or "chemical shift," which for antimony in germanium is[11] $(0.57\pm 0.03)\times 10^{-3}$ eV, and ϵ_2 is the (positive) energy separation between the conduction band edge and the triplet level. The density-of-states effective mass $m^*=(m_{\|}m_{\perp}^2)^{\frac{1}{3}}=0.224m$,[12] where m is the free-electron mass.

In principle, the value of n may be determined from the "infinite field" Hall constant,[13,14] $R_\infty=(nec)^{-1}$; the value for (N_D-N_A) may be obtained from the Hall constant at ~ 77°K and the known drift to Hall mobility ratio.[15]

It is clear from Eq. (2) that ϵ_2 may be determined from the slope of a plot of $\ln(\alpha)$ vs T^{-1}, where $\alpha=nT^{-\frac{3}{2}}\times[3+\exp(\delta\epsilon/kT)]$, at temperatures sufficiently low such that $n\ll N_A$. Using this value of ϵ_2, a value for N_A may be computed using Eq. (2) in the temperature range considered. Consistency may then be checked by plotting $\ln\chi_1$ vs T^{-1}, where

$$\chi_1\equiv nT^{-3/2}(N_A+n)\{3+\exp(\delta\epsilon/kT)+\sum_{i>2}(g_i/2)\exp[(-\mathcal{E}_1+\mathcal{E}_2)/kT]\}/(N_D-N_A-n),$$

for the range $n\ll N_A$ to $n\sim N_D-N_A\gg N_A$, or roughly from 4–25°K; the plot should yield a straight line.

C. Experimental Procedures and Results

The temperature variation of the Hall mobility of the two ⟨100⟩ samples mentioned above, $n45-2a$ and $n45-10a$, (hereafter referred to as samples 2 and 10) is shown in Fig. 2. Also shown are the data for $n45-10c$, a ⟨110⟩ oriented sample cut from the same crystal. For comparison, the data for crystal No. 135, taken from Fig. 6, of Morin, Geballe, and Herring,[16] are shown.

[6] C. Herring, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961), p. 60. See C. Herring, T. H. Geballe and J. E. Kunzler, Bell System Tech. J. **38**, 657 (1959), discussion p. 702 ff.

[7] K. S. Shifrin, J. Tech. Phys. (U.S.S.R.) **14**, 43 (1944). H. Y. Fan, in *Solid-State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1955), Vol. I, p. 283; E. M. Conwell, Phys. Rev. **99**, 1195 (1955); P. T. Landsberg, Proc. Phys. Soc. (London) **B69**, 1056 (1956); E. H. Putley, *ibid.* **72**, 917 (1958); E. H. Putley, *The Hall Effect and Related Phenomena*, (Butterworths Scientific Publications, Ltd., London, 1960), p. 123 ff.

[8] E. M. Conwell, reference 7, p. 1197 ff; W. Kohn and J. M. Luttinger, Phys. Rev. **97**, 883, 1721 (1955); **98**, 915 (1955).

[9] P. J. Price, Phys. Rev. **104**, 1223 (1956).

[10] See W. Kohn, in Solid-State Physics, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5, p. 257 for a review.

[11] H. Fritzsche, Phys. Rev. **120**, 1120 (1960).

[12] G. Dresselhaus, A. F. Kip and C. Kittel, Phys. Rev. **98**, 368 (1955); R. N. Dexter, H. J. Zieger, and B. Lax, *ibid.* **104**, 637 (1956).

[13] J. A. Swanson, Phys. Rev. **99**, 1799 (1955).

[14] H. Brooks, *Advances in Electronics and Electron Physics* (Academic Press Inc., New York, 1955), Vol. VII, p. 135 ff.

[15] F. J. Morin, Phys. Rev. **93**, 62 (1954); C. Herring, Bell System Tech. J. **34**, 237 (1955); J. J. Hall (1960) (unpublished).

[16] F. J. Morin, T. H. Geballe, and C. Herring, Phys. Rev. **105**, 525 (1957).

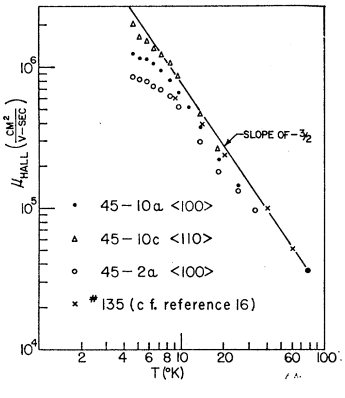


FIG. 2. Hall mobility vs temperature for several samples. The solid line with a slope of $-3/2$ represents the theoretical variation of the mobility due to acoustic phonon scattering. It is normalized to the data at 77°K.

This latter sample, with stated acceptor concentration of $3\times10^{11}/\text{cm}^3$, has the highest low-temperature mobility published to date. A line with slope $-\frac{3}{2}$, passing through the measured Hall mobility at 77°K is drawn to indicate the mobility expected from acoustic phonon scattering alone, (at 77°K, optical phonon and intervalley scattering make negligible contributions to the mobility). For sample $10-c$, which has the lowest donor concentration of the three, and thus the least neutral impurity scattering, it is seen that theoretical lattice mobility is maintained as low as 6°K. From Fig. 2, it is clear that N_A for sample 2 is essentially the same as for No. 135, and that for sample 10 is less. It is, however, difficult to be more quantitative.

The accuracy of the determination of N_A, utilizing Eq. (2), depends directly on the certainty with which ϵ_2 can be determined. It is readily seen that an error $\sim2\%$ in ϵ_2 can produce an error of a factor of 1.5 in N_A, since $\epsilon_2/k\sim100°$. Therefore, special thermometry procedures were used in obtaining the data from which a value for ϵ_2 was deduced.

In Fig. 3 is plotted the quantity $[3+\exp(\delta\epsilon/kT)]\times T^{-3/2}(eR_\infty)^{-1}$ vs T^{-1} for both samples 2 and 10 (the former having the greater donor concentration) for the temperature range 4.0–5.1°K. Here R_∞ is the "infinite field" Hall constant. Over this entire temperature range $n\ll N_A$. The magnetic field H, ~1000 G, which corresponds to $\mu H\sim10$, was well in the high-field limit; on the other hand, H was not so great as to produce a variation of the activation energy[17] of more than $\sim0.1\%$. For this data, region E (Fig. 1) was made common with the space above the helium in the Dewar and the exchange gas allowed to condense. The temperature was determined from the vapor pressure using the T_{58} scale.[18] The T_{58} scale is thought to be within ±0.001 deg of the true thermodynamic scale, and the data points of Fig. 2 do not scatter about the best straight line through them by more than this amount. The fit to the T_{55} scale[19] is not as good.

The values of ϵ_2 for samples 2 and 10 are $(9.57\pm0.03)\times10^{-3}$ eV and $(9.72\pm0.03)\times10^{-3}$ eV, respectively. The quoted uncertainty is felt to be conservative. It includes estimates of any systematic errors due to temperature gradients, scatter in the experimental points, and a presumed uncertainty of 10% in the value of $\delta\epsilon$, which is twice the stated probable error.[11] This relative insensitivity of ϵ_2 to variations of $\delta\epsilon$ occurs because $\delta\epsilon\simeq kT$, and therefore more donors are in the triplet than the singlet state for the temperature range considered. The difference between the values for ϵ_2 for the two samples is real and will be discussed below.

The data of Fig. 2 also yield a value for $C=N_A/(N_D-N_A)$ since all the other parameters in Eq. (2) are now known. Here N_D is the density of those donors that contribute electrons to the conduction band over the temperature range considered. (This remark will be amplified below.) The results are $C_2=0.036$, $C_{10}=0.067$.

To accurately extend the measurements to higher temperatures requires careful calibration of the carbon-resistance thermometer. Our procedure was to calibrate the thermometer at the normal boiling point of liquid argon (~87°K), in the range from the triple point to the boiling point of equilibrium hydrogen (~14–20°K) and from 4.0–5.0°K in the liquid-helium range. The three constant interpolation equation of Clement and Quinnell[20] was used to compute a resistance vs temperature curve for the thermometer, using the normal boiling point of the three gases as fixed points. In addition, a

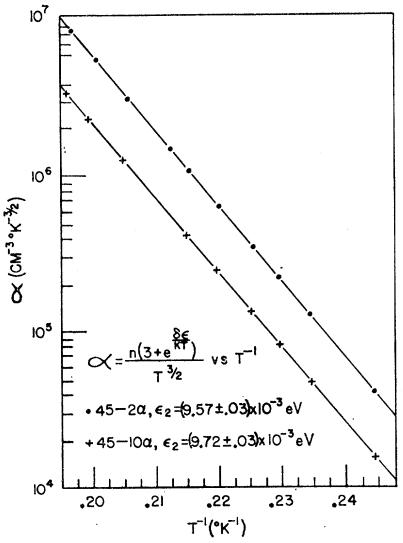


FIG. 3. Data from which the "activation energy" of the donors may be determined. The quantity ϵ_2 is the separation between the conduction band and the triplet levels derived from the unperturbed 4 fourfold degenerate ground state.

[17] Y. Yafet, R. W. Keyes, and E. N. Adams, J. Phys. Chem. Solids **1**, 137 (1956); W. Kohn, reference 10, p. 315. Note that the estimate after Eq. (14.3) of the importance of the magnetic field in affecting the activation energy is too great by a factor of ~4.

[18] H. van Dijk and M. Durieux, Physica **24**, 920 (1958).

[19] H. van Dijk and M. Durieux, *Progress in Low-Temperature Physics* (North-Holland Publishing Company, Amsterdam, 1957), Vol. II, p. 431.

[20] J. R. Clement and E. H. Quinnell, Phys. Rev. **85**, 502 (1952).

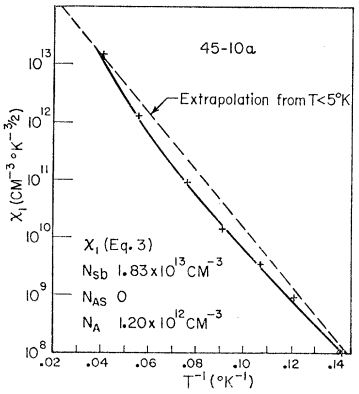


FIG. 4. Data showing that the temperature-dependent carrier density for sample 45–10a cannot be fit to the appropriate statistical formula if it is assumed that the sample contains only Sb donors. The experimental points would otherwise fall along the dashed line.

four-constant equation due to Hall[21] was used by including the hydrogen triple-point data. Both curves fit the calibration data over the hydrogen range within the uncertainty of the vapor-pressure measurements. However, a significant discrepancy, with a maximum spread of 0.25° in the region of ~7°K, where standard fixed points are most difficult to obtain, exists between the two curves. The final calibration curve for the thermometer was obtained by simply drawing a smooth curve, between the two analytic curves, that fit the vapor pressure data to 5.1°K and passed through a known point at 7.14°K obtained by measuring the superconducting transition temperatures of two samples of lead.[22] It is felt that all temperatures quoted are correct to within $\frac{1}{2}\%$.

Using the value for (N_D-N_A) obtained for sample 10 from the Hall constant data taken at 77°K (utilizing the known ratio of drift to Hall mobility,[15] the quantity χ was plotted (Fig. 4) vs T^{-1}. The value used for N_A $(1.2\times10^{12}/\text{cm}^3)$ was obtained from N_D-N_A and the value for C_{10} above. It is clear that in the region $n\sim N_A \ll(N_D-N_A)$, the curve deviates significantly from linearity. The deviation is in a direction that suggests that a smaller value of N_A would be appropriate. However, at higher temperatures where $n\sim(N_D-N_A) \gg N_A$, the deviation goes to zero. Since the ratio of acceptors to donors active at ~4°K(C_{10}) is fixed, it would appear that perhaps a deeper lying donor (e.g., As or P) is present in the sample. The activation energy of these donors is ~30% greater than that of Sb[23–25]; therefore if present in amounts similar to Sb, they would not measureably affect the conductivity below ~8°K. However, for $T\geq20$°K, these donors would be indistinguishable from Sb.

The circles of Fig. 5 are a plot of $\ln\chi_2$ vs T^{-1}. χ_2 (see Eq. A6, Appendix) is a combination of experimental quantities including N_{Sb} and N_{As} (the antimony and arsenic density, respectively), and the singlet-triplet ground-state splitting for both As and Sb donors. When plotted as in Fig. 5, it should yield a straight line with slope ϵ_2/k and coincide with the data of Fig. 3 at lower temperatures. The deeper donors were assumed to be arsenic since, judging from the distribution coefficients of the various donors[26] and the history of the crystal, this is a more likely contaminant. However, the results in Fig. 5 would be changed only in a trivial way were the deeper level assumed to be phosphorus. In computing χ_2, sums over all the excited states of both Sb and As donors had to be made. The "effective mass values" as computed by Kohn and Luttinger[8,10] (and experimentally verified[25]) were used for all levels with principal quantum number $n>1$. The measured ground-state splitting[27] of 4.0×10^{-3} eV and activation energy of [23,28] 12.7×10^{-3} eV for As were used.

The relative values of N_{Sb} and N_{As} were obtained from an analysis of data kindly obtained by Hall on two samples, cut near 10, that were subjected to large uniaxial ⟨111⟩ compressive stress. In the limit of large strain, the difference in ground-state–conduction-band separation is altered in a known manner[9] so that by

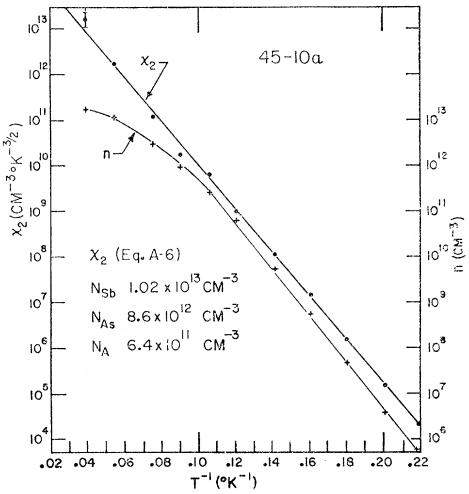


FIG. 5. Data showing that the temperature-dependent carrier density can be fit to the appropriate statistical formula if it is assumed that the sample contains both Sb and As (or P) donors in the concentrations indicated. The straight line for χ_2 is the theoretical result.

[21] J. J. Hall (unpublished).
[22] W. B. Pearson and J. M. Templeton, Phys. Rev. **109**, 1094 (1958).
[23] T. Geballe and F. J. Morin, Phys. Rev. **95**, 1085 (1954).
[24] P. Debye and E. M. Conwell, Phys. Rev. **93**, 693 (1954).
[25] H. Y. Fan and P. Fisher, J. Phys. Chem. Solids **8**, 270 (1959); W. S. Boyle, *ibid.* **8**, 321 (1955).
[26] F. A. Trumbore, Bell System Tech. J. **39**, 205 (1960).
[27] H. Fritzsche, Phys. Rev. **115**, 336 (1959); G. Weinreich and H. G. White, Bull. Am. Phys. Soc. **5**, 60 (1960); D. K. Wilson and G. Feher, *ibid.* **5**, 60 (1960).
[28] There is a significant difference between the thermal activation energy of 12.7×10^{-3} eV (footnotes 23, 24) and the optical value of 14.0×10^{-3} eV (footnote 25). This point will be discussed in detail in the final sections of this paper.

TABLE I. Pertinent parameters for the various n-type germanium samples studied.

Sample[a]	Source	Nominal dopant	(N_D-N_A)[b] (10^{13} cm^{-3})	N_{Sb} (10^{13} cm^{-3})	N_{As} (10^{13} cm^{-3})	ϵ_2 (meV)	ϵ_1 (meV)	N_A (10^{12} cm^{-3})	$N_{A,I}$[c] (10^{12} cm^{-3})	μ_I (10^5 cm^2/Vsec)	E_B(5°K) (V/cm)
45–10	GE	Sb	1.84	1.02	0.86	9.72±0.03		0.64	0.6	32.2	4.0
45–2	GE	Sb	3.87	2.11	1.83	9.57±0.03		0.73	1.5	13.2	3.6
44–1	BTL	Sb	6.41	···	···	9.41		6.8	5	4.2	5.5
49–1	BTL	As	15.6	···	···	···	12.6	16.9	···	···	6.6
46–2	BTL	P	0.43	···	~0	···	12.1	2.9	2.5	8.2	15.1
31	BTL	Sb	0.70	···	···	···		···	2	11.0	14.9
28–6	BTL	Sb	1.26	~0.7	~0.7	9.79±0.03		1.4	1.7	12.1	4.8
42–1	BTL	Sb	2.5	···	···	···		···	8	2.5	8.05
41–15	BTL	Sb	3.1	···	···	···		···	6.5	3.1	7.6
50–1	BTL	Sb	16.1	···	~0	9.30±0.03		13.3	···	···	···

[a] The first pair of numbers is a chronological numbering for the crystals, in the order received. The number to the right of the hyphen indicates the slice taken from the parent crystal. Samples cut from the same slice would have a letter following the numbers.
[b] N_D-N_A was obtained from Hall constant measurements at 77°K, when all donors are ionized.
[c] $N_{A,I}$ is the value of N_A deduced from the partial mobility due to ionized impurities, μ_I, assuming the scatterers are singly charged. The results are normalized to sample 45–10.

comparing the large-strain and zero-strain conductivity and Hall coefficient, a value for $(N_{As}+N_{Sb}-N_A)/(N_{Sb}-N_A)$ may be obtained. This experiment was performed at three temperatures. The results were consistent with the conjecture of a deeper impurity and yielded the ratio

$$(N_{As}+N_{Sb}-N_A)/(N_{Sb}-N_A)=1.95\pm0.1.$$

It was assumed that this ratio is appropriate for both samples 10 and 2. It is seen that these values for donor concentration greatly reduce the discrepancies in Fig. 4, but do not completely eliminate them. It is tempting to blame the remaining discrepancies, which are, in fact, quite small, on such things as impurity gradients and inhomogeneities, internal strains, perhaps some unique, as yet unknown, type of donor structure associated with antimony impurities,[29] or a dependence of ϵ_2 on the density of ions.[24]

The values for impurity concentration of these samples, and several others, are included in Table I. Sample 44–1 is an antimony-doped zone-leveled crystal of higher impurity concentration than either 2 or 10. The variation of χ_1 with T^{-1} is shown for it in Fig. 6. There is again a significant deviation of the plot from the expected straight line, but in a direction opposite to that for sample 10 (and 2) as shown in Fig. 4 and much smaller in magnitude. No correction for possible As contamination was made (the crystal history makes this reasonable) which would, in fact, adversely affect the fit of the curve. The deviation in this case sets in before n becomes of the order of N_A; Eq. (1) would then require the trouble to be related to some unknown structure of the low-lying excited donor states,[29] were the crystal homogeneous.

The relative values of N_A listed in Table I for all the crystals are quantitatively consistent with all other data that will be described. The absolute values for N_A, though larger than for example the value stated[16] for sample No. 135 (which we feel is incorrect), are approximately a factor of two less than some of our previous estimates obtained by less exact procedures. This will require the reinterpretation of some previously published[2,3] values for absolute cross sections and will be discussed in the appropriate sections of this report.

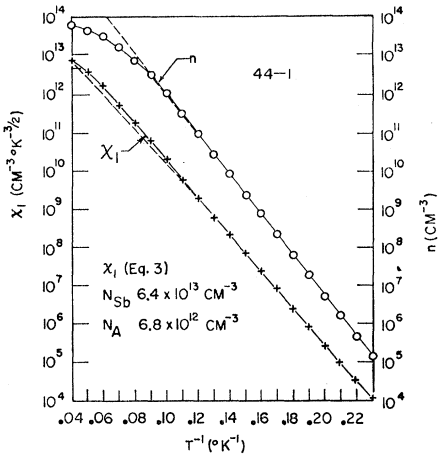


FIG. 6. Data showing that the temperature-dependent carrier density for a somewhat more impure sample than 45–10, almost certainly containing only Sb donors, cannot be fit precisely to the appropriately statistical formula (the dashed extrapolation for χ_1). Assuming deeper donors would make the fit worse. The deviations set in for $n \ll N_A$, suggesting that the discrepancy is related to some unknown structure of the low-lying excited donor states.

D. Discussion

Figure 2 shows the typical behavior of the T dependence of the Hall mobility to be an increasing deviation from the $T^{-3/2}$ acoustic phonon scattering law as T is reduced from 77°K. The deviation (greater for larger impurity concentrations), is caused by ionized impurity scattering from the N_D+N_A ionized sites until ~15°K, when the electrons begin to freeze out and neutralize the donors. The mobility then inflects upward toward the $T^{-\frac{3}{2}}$ line, reaching it at ~10°K for the sample of

[29] R. E. Pontinen and T. M. Sanders, Jr., Phys. Rev. Letters 5, 311 (1960); R. W. Keyes and P. J. Price, *ibid.*, 5, 473 (1960).

lowest N_A. At still lower temperatures, the mobility is determined by N_A, as discussed previously, which dependence will contribute a positive term to the temperature derivative of the mobility. This general behavior has also been discussed by Debye and Conwell.[24] However, a recent "anomaly" in the cyclotron resonance linewidth found by Bagguley *et al.*[30] in the range 15–30°K makes necessary a quantitative discussion to ascertain whether or not a similar anomaly exists in the Hall mobility. *Note added in proof.* M. Fukai, H. Kawamura, I. Imai, and K. Tomishima [J. Phys. Soc. Japan **17**, 1191 (1962)] have repeated the experiment of Bagguley *et al.*[30] and have found no anomaly. The resonance anomaly consists of a contribution to the linewidth, the temperature dependence of which is qualitatively the same as the contribution to the mobility of the scattering by ionized impurities; quantitatively, however, the magnitude would appear to be too great to be accounted for by the presumed (temperature dependent) number of ionized centers[30,31] present in the sample.

The contribution to the data of Fig. 2 of neutral impurity scattering can be shown to be small. From the results of Erginsoy,[32] the partial mobility due to N_N neutral, hydrogenlike donors is

$$\begin{aligned}\mu_N &= m_N e^3/(6000 N_N K\hbar^3), \\ &= 8.95(m_N/m)N_N^{-1}\times 10^{20}\ \text{cm}^2/\text{Vsec}.\end{aligned} \tag{4}$$

Here m_N, which should be the "density-of-states" effective mass,[14] has been treated by Debye and Conwell[24] as a parameter to be experimentally determined. They found for neutral As donors that a value for m_N/m between 0.25 and 0.5 gave the best fit to their data. For Sb donors, the data in Fig. 3 of Koenig and Gunther-Mohr[5] show that 1.7×10^{15} neutrals per cm^3 contribute a partial mobility of 1.3×10^5 cm^2/Vsec, corresponding to $(m_N/m)=0.25$. Since the Sb ground state radius is somewhat larger than that of As, we have taken the value 0.25, for which

$$\mu_N = 2.2 N_N^{-1}\times 10^{20}\ \text{cm}^2/\text{Vsec}. \tag{5}$$

For sample 2, the partial mobility $\mu_N \simeq 6\times 10^6$ cm^2/Vsec then contributes less than 10% to the total mobility in the temperature range of the microwave anomaly.

The contribution of ionized impurity scattering to the mobility, μ_I, at 5°K can now be estimated from the measured Hall mobility. The results for several samples, listed in Table I, were obtained by assuming (1) the acoustic phonon scattering at 5.0°K to be 2.16×10^6 cm^2/Vsec (from the extrapolation of the $T^{-3/2}$ line of Fig. 2), and (2) that the various partial mobilities add strictly as sums of reciprocals. It is apparent that though μ_I varies monotonically with N_A^{-1}, the variation is not linear. This is not unexpected, however, since the nature of the acceptor center is not known. N_A as used in Eq. (1) measures the number of compensated donors. If the acceptors are (triply ionized) Cu centers, the Cu density is $N_A/3$. Since the scattering strength goes as the square of the charge, the equivalent number of singly ionized sites is therefore $3N_A$ negative and N_A positive sites. The partial mobility μ_I would then be proportional to $(4N_A)^{-1}$. If, on the other hand, the acceptors are singly ionized, μ_I would vary as $(2N_A)^{-1}$. Within the uncertainty set by these considerations, μ_I as tabulated in Table I varies as N_A^{-1}.

From Table I, 10^{12} singly ionized donors/cm^3 would contribute a partial mobility at 5°K of 2×10^6 cm^2/Vsec. One may compare this value with the Brooks formula, appropriately modified for the "freeze out" region.[14]

The theoretical result is 2.5×10^6 cm^2/Vsec for 10^{12} donors/cm^3. Though this close agreement, in view of the basic uncertainties in the theory, is fortuitous, it (together with the fact that the theory fits the data reasonably at much higher temperatures) suggests that one may use the approximate theoretical temperature dependence ($T^{3/2}$) to extrapolate the 5°K values of μ_I to temperatures in the region of the anomaly. Thus sample 2, which at 15°K is essentially extrinsic, would have $\sim 3.8\times 10^{13}$ ionized scatters contributing a partial mobility $\mu_I = 2.5\times 10^6\times(15/5)^{3/2}/38 = 3.4\times 10^5$ cm^2/Vsec. Combining this with the theoretical lattice mobility (summing reciprocals) of 4.18×10^5 cm^2/Vsec give $\mu = 1.9\times 10^5$ cm^2/Vsec. This compares very well with the experimental value (Fig. 2) of 2.1×10^5 cm^2/Vsec.

From the above analysis, it is quite unlikely that any anomalous scattering exists that contributes to the mobility. In particular, it would appear that no scattering other than that due to the electrically active impurities known to be present in the samples exists. For example, oxygen which is often reputed to be present in significant amounts in germanium, and conjectured to have an important effect on the low-temperature mobility,[4] would appear not to influence the present data. Similar results were found by Logan and Peters for silicon.[33]

It is interesting to note that, from Table I, the observed value of ϵ_2 is still concentration dependent at relatively low concentrations of donors. (It should be recalled that ϵ_2 is the energy separation of the conduction and triple levels.) In addition, the value of ϵ_2 is somewhat greater than ϵ_{em}, the value of the ionization energy calculated in the effective mass approximation by a variational technique (9.2±0.2 meV) ignoring the core effects which give rise to the singlet-triplet splitting.[10] One might make the argument, as did Kohn and Luttinger[8] for Si, that since the triplet levels penetrate the central cell least, ϵ_2 should be close to ϵ_{em}. However, this argument is only approximate and not appropriate

[30] D. M. S. Bagguley, R. A. Stradling, and J. S. S. Whiting, Proc. Roy. Soc. (London) **A262**, 340 (1961).
[31] D. M. S. Bagguley (private communication).
[32] C. Erginsoy, Phys. Rev. **79**, 1013 (1950).
[33] R. A. Logan and A. J. Peters, J. Appl. Phys. **31**, 122 (1960).

for the accuracy considered here. The difference between ϵ_2 and ϵ_{em} is given by the difference between the expectation value of the core perturbation over the ground-state, "effective-mass" wave function made from Bloch functions from the region of the band edge of a given valley, and the off-diagonal matrix element between such functions for two different valleys (ϵ_{Id} and Δ_c in the notation of Price[9]). There is no *a priori* reason for these matrix elements to be equal.

A more accurate solution of the effective mass equation and more measurements of donor activation energy, particularly as a function of strain, will be necessary to better understand the nature of the deviations from the simple effective mass theory in germanium.

The variation of ϵ_2 with donor density cannot be due to an overlapping and subsequent broadening of the ground state, since the ground state diameter[10] is $\sim$150 Å and the mean donor separation is $\sim$5000 Å. Moreover, the measurements by Fritzsche[11] of the singlet-triplet splitting in Sb donors made at concentrations one to two orders of magnitude larger than those considered here show no effects due to possible overlapping of the $n=1$ levels. The variation must be due to overlapping of the highest excited states, and a subsequent merging of these states with the bottom of the conduction band. One model[34] would be to consider only the ionized donors, and to realize that their separation is $\sim 10^4$ Å, so that the Coulomb fields of adjacent ionized donors would overlap at energies $\sim 10^{-4}$ eV below the band edge. Electrons within energies $\sim 10^{-4}$ eV below the conduction band then could not be considered to be bound to a particular donor, but rather would be in the conduction band. The variation of ϵ_2 from sample to sample is also of the order of 10^{-4} eV, and we feel that the variation does in fact result from a distortion of these high states.

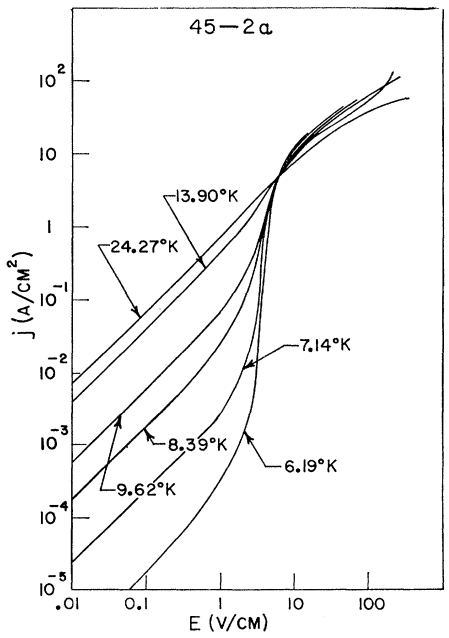


FIG. 7. Current-density–electric-field variation for sample 45–2a for several temperatures.

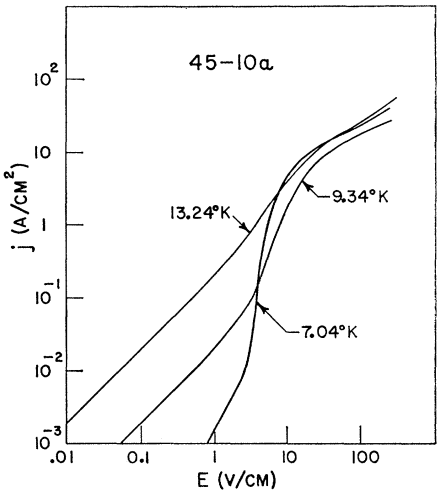


FIG. 8. Current-density–electric-field variation for sample 45–10a for several temperatures.

III. NONLINEAR STEADY STATE CONDUCTIVITY

A. General Considerations

The nature of the variation of conductivity with applied electric field for n-germanium at low temperature has been known for some time.[5,35] For electric fields $\gtrsim$0.2 V/cm, deviations from Ohm's law occur; the conductivity increases monotonically until a critical or "breakdown" field (typically in the range 4–10 V/cm) is reached at which the current increases by many orders of magnitude for a small additional increase in field. Figures 7 and 8 indicate this variation, for different temperatures, for samples 2 and 10.

The deviations from Ohm's law set in because the electrons get "hot," i.e., the rate of energy absorption by the electron distribution from the field cannot be dissipated to the lattice for electric fields $\gtrsim$0.2 V/cm, unless the average energy of the distribution deviates from its thermal equilibrium value. In the region of the breakdown field, it is generally assumed that the mean electron energy is of the order of 10^{-2} eV, i.e., of the order of the donor ionization energy. The large increase in conductivity is then associated with an increase in carrier density caused by impact ionization of neutral impurities by the hot carriers.

The variation with electric field of the Hall mobility μ_H, current density j and carrier density n (obtained

[34] This model was suggested to us by T. N. Morgan. Other mechanisms that could cause the activation energy to vary with impurity concentration have been discussed by P. Debye and E. M. Conwell, reference 24.

[35] W. Turner, J. W. Davisson, and E. Burstein, Proceedings of the Schenectedy Cryogenics Conference, October, 1952 (unpublished).

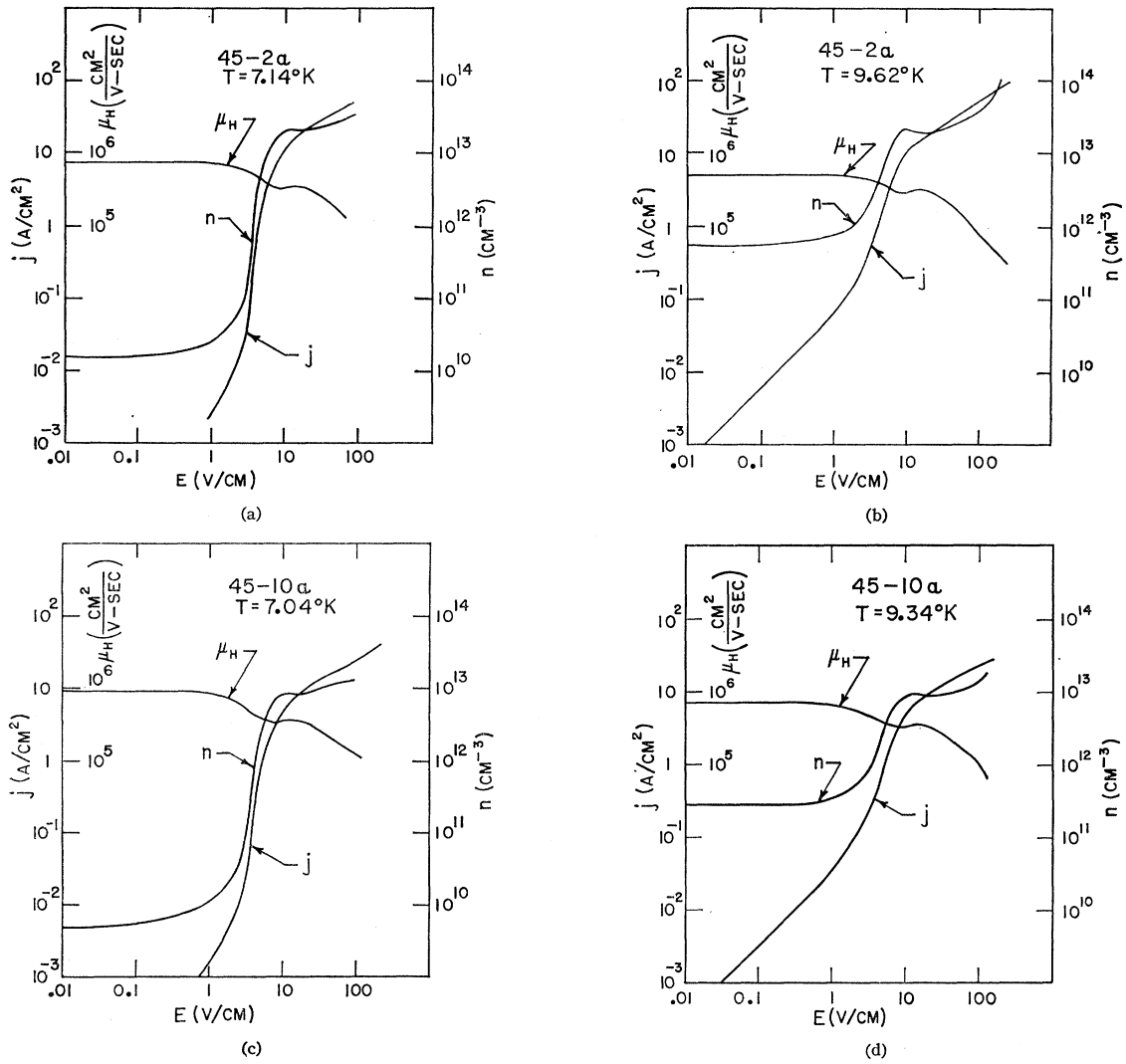


FIG. 9. (a)–(d). Variation of current density, carrier density, and Hall mobility vs electric field for samples 45–10*a* and 45–2*a* for various temperatures. The carrier density n was obtained from the Hall constant assuming a value of unity for the Hall to drift mobility ratio. Data for $j>5$ A/cm^2 were obtained using 3-μsec pulses; the other data were obtained by dc measurements.

from the Hall mobility and conductivity, assuming the drift and Hall mobility equal) is shown in Fig. 9(a), (b), (c), (d). Data for current densities greater than $\sim$5 A/cm^2 were taken with pulses $\sim$3-μsec duration. For the larger values of electric field ($\sim$100 V/cm) a combination of impedance of the arms of the sample and decreasing Hall mobility made the results progressively more unreliable. The tendency of n to increase to a value greater than the donor density is spurious; it has been shown to be due to injection in one case, and probably is for all data above 100 V/cm. It is seen from Fig. 9 that n increases rapidly in the region of breakdown and attains an approximately constant value of $\sim(N_D-N_A)/2$. This is in agreement with the earlier results of Ryder *et al.*[36] The behavior of the Hall mobility μ_H is more complex. Between 0.2 and 1 V/cm there is a slight peak in the mobility which cannot quite be discerned on the scale of Fig. 9, but becomes more apparent at lower temperatures and in more impure samples.[5] This peak is associated, as per the early suggestion of Conwell,[37] with an initial decrease in impurity scattering as the electrons begin to heat followed by a subsequent increase in lattice scattering which again lowers the mobility. The minimum and following peak in μ_H in the region of $\sim$6 V/cm is of a similar origin. As n increases, the number of ionized scattering centers increases, lowering the mobility. When n reaches its maximum value, μ_H again increases since the carrier energy is still increasing, until once again the mobility is determined by the lattice scattering alone.

[36] E. J. Ryder, I. M. Ross, and D. A. Kleinman, Phys. Rev. **95**, 1342 (1954).

[37] E. M. Conwell, Phys. Rev. **90**, 769 (1953).

B. Mobility Variation and Electron "Temperature"

That the above interpretation of the field dependence of the mobility is correct can be demonstrated in two ways. Firstly, the values of μ_H in the regions of the mobility curve which we attribute to lattice scattering have, for equal lattice temperatures, the same magnitude and variation for samples 10 and 2. Therefore the electron distribution function in these regions is independent of impurity concentration. Secondly, it is possible to measure the drift mobility as a function of electric field without allowing the impurities to ionize. The times that characterize the establishment of the carrier energy distribution and the carrier density after a change of applied electric field are grossly different, the first being $\sim 10^{-9}$ sec, the second, one or two orders of magnitude longer.[2] Therefore, after the electric field is altered quickly, the initial (after 10^{-9} sec) change in current depends, in an obvious manner, only on the initial and final value of field and the relative drift mobility for the distribution function corresponding to each field. (The experimental procedures, especially for the higher values of field, are somewhat involved because of the high-speed pulses required; they are described in Appendix B.) Results are shown in Fig. 10, appropriately normalized. It is seen that the variations in the region of breakdown associated with impurity scattering do not appear in the drift mobility.

The variation of the mobility at higher fields, quite closely as $E^{-\frac{1}{2}}$, as predicted theoretically for acoustic phonon scattering (with the simplifying assumption that the phonons involved in scattering obey equipartition),[38] suggests that application of this theory should give a good indication of the electron energy as a function of field. At these low temperatures, optical phonons do not exist to be absorbed, and as will be apparent, the electron distribution in the region of breakdown probably is not so energetic as to make spontaneous emission of optical phonons even from the tail of the distribution an important consideration.

To make quantitative comparisons with theory, the standard theories, which assume an isotropic mass and relaxation time,[38,39] must be extended. This has been done by Shibuya[40] for the limit of the electric field sufficiently high such that the variation of μ as $E^{-\frac{1}{2}}$ is a good approximation. The result for the distribution function in a single valley (which he does not state) is

$$f(\epsilon)=A\,\exp[-\epsilon^2/2p(kT)^2], \tag{6}$$

from which

$$\langle\epsilon\rangle/kT=(2p)^{1/2}\Gamma(5/4)/\Gamma(3/4)=1.045p^{1/2}. \tag{7}$$

Here $\langle\epsilon\rangle$ is the average energy of the carriers. The quantity p is proportional to the square of the electric field, and in the anisotropic case contains the mass and scattering anisotropy. The constant A is for normalization. The ratio of (field-dependent) mobility to Ohmic mobility is

$$\mu/\mu_0=\pi/[2^{5/4}\Gamma(3/4)p^{1/4}]. \tag{8}$$

The current, in general no longer collinear with the applied field, is given by

$$\mathbf{j}=ne\mu\left(\frac{3}{\mathrm{Tr}(\mathbf{m}^{-1})}\right)\mathbf{m}^{-1}\cdot\mathbf{E}, \tag{9}$$

where $\mathbf{m}^{-1}$ is the reciprocal effective mass tensor.

The form of the results (6)–(8) are identical to those obtained by Yamashita and Watanabe[39] for the isotropic case, the differences between the two cases being in the expression for p. The anisotropy contribution to p may be derived very simply. The steady state is achieved as a result of a balance between the rate of energy gained by the distribution from the field $\mathbf{E}$ and the energy loss rate to the lattice. The rate of energy gain per electron is

$$\begin{aligned} e\mathbf{E}\cdot\mathfrak{y}\cdot\mathbf{E}&\simeq(e\mathbf{E}\cdot\mathfrak{y}_0\cdot\mathbf{E})(kT/\langle\epsilon\rangle)^{1/2}\\ &=(kT/\langle\epsilon\rangle)^{1/2}e^2\tau_0\mathbf{E}\cdot\mathbf{m}^{-1}\cdot\mathbf{E}\\ &=3(kT/\langle\epsilon\rangle)^{1/2}e\mu_0[\mathrm{Tr}(\mathbf{m}^{-1})]^{-1}\mathbf{E}\cdot\mathbf{m}^{-1}\cdot\mathbf{E}; \end{aligned} \tag{10}$$

Here τ is the scattering time for acoustic scattering (assumed isotropic, since its anisotropy is small compared to the mass anisotropy), $m_\perp$ and $m_\parallel$ the usual transverse and longitudinal effective mass components; the subscript zero refers to the Ohmic value. Since phonons are fairly "heavy" compared to electrons,[38] the energy loss will be greater when collisions occur when the electrons are heavy. The loss term, if proportional to the average mass, would then be[38]

$$[(\mathrm{Tr}\mathbf{m}/3)c^2/kT]\langle\epsilon\rangle/\tau. \tag{11}$$

Equating Eqs. (10) and (11) gives

$$\begin{aligned} &(\langle\epsilon\rangle/kT)^2\\ &\quad=(\mu_0/c)^2[\mathrm{Tr}(\mathbf{m}^{-1})/3]^{-2}(3/\mathrm{Tr}\mathbf{m})\,\mathbf{E}\cdot\mathbf{m}^{-1}\cdot\mathbf{E}. \end{aligned} \tag{12}$$

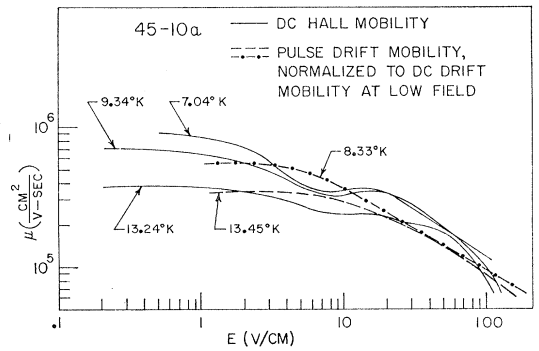


FIG. 10. Electric field dependence of Hall and drift mobility. The Hall mobility was measured after the carrier density had achieved its steady-state value. The dips in this data are due to the increased ionic scattering as more donors ionize. The drift mobility was measured by fast-pulse techniques, before the carrier density increased from its value in the Ohmic range.

[38] W. Shockley, Bell System Tech. J. **30**, 990 (1951).
[39] J. Yamashita and M. Watanabe, Progr. Theoret. Phys. **12**, 443 (1954).
[40] M. Shibuya, Phys. Rev. **99**, 1189 (1955).

TABLE II. Values for the average electron energy $\langle\epsilon\rangle$ inferred from data of Figs. 9 and 10 and also computed from first principles using the Shibuya theory (reference 40). The quantity μ/μ_0 is the ratio of the mobility at E to the theoretical lattice mobility in the Ohmic region. p is defined in Eq. (6) of the text.

T (°K)	E (V/cm)	μ/μ_0	$p^{1/2}$ Theory	$p^{1/2}$ Experimental	$\langle\epsilon\rangle/k$ (°K) Theory	$\langle\epsilon\rangle/k$ (°K) Experimental
7.04	$3.2=E_B$	0.36	2.7	9.25	20	68
8.3	40	0.16	26	45	230	380
13.5	40	0.33	13	11	183	155

The terms on the right-hand side of Eq. (12) containing the mass tensor constitute the correction to the simple Shockley[38] picture for mass anisotropy and are identical to the factors obtained by Shibuya [his Eq. (12)]. Here c is an appropriate average of the transverse and longitudinal sound velocities that contains a factor (S in Shibuya's notation) that is a complicated algebraic function of the mass anisotropy and the ratio of the two deformation potentials. This quantity appears to be of the order of unity, within about 20%, but unfortunately, there are several (subtle) misprints in the expression and the graph that Shibuya gives for S. We are, therefore, unable to make as quantitative a comparison of the data with theory as had been hoped.

The resulting physical picture is that for a given low-field mobility, the more anisotropic the carriers, the more difficult they are to heat, since both the mobility and energy gain go as the light mass, but the losses are weighted by the heavy mass.

In Table II, values for p and $\langle\epsilon\rangle$ computed using Shibuya's theoretical expressions, assuming $S=1$, are listed for representative fields and temperatures. Experimental values for p and $\langle\epsilon\rangle$, obtained by using Eq. (8) and the data of Figs. 9 and 10, are also tabulated. Though at 13.5°K, the experimental and theoretical values are in quite good agreement, as the temperature is lowered the electrons appear to be hotter for a given field than the theory allows. In addition, it is seen in Fig. 10 that the 8.3° and 13.5° mobility curves coincide for $E=40$ V/cm. This is in contradiction to the simple theory[38] and again implies that for a given field, $\langle\epsilon\rangle$ at the lower temperature is greater, and the mobility lower, than one would expect. The explanation of this point is straightforward, and is to be found in a breakdown of the phonon equipartition assumption.

Consider an energy losing collision for a carrier moving in the heavy-mass direction with a kinetic energy $\sim 10^{-2}$ eV, i.e., a "hot" electron. The most energetic process would be for it to reverse its momentum; in fact, the theory involves implicitly the condition that the particle scatter with equal probability to all points of the final energy surfaces. Since the heavy mass is ~ 1.6 times the free-electron mass, the momentum change Δq would be $\sim 1.4\times 10^{-20}$ g cm/sec. The phonon involved would have an energy $(\Delta q)c \sim 5\times 10^{-15}$ erg ~ 30°K. This energy is greater than kT for the temperature range considered, so that the possibility for backward scattering of the carrier in question is progressively reduced as the lattice temperature is lowered. (This is the same phenomenon that in metals contributes to the T^5 behavior of the low-temperature conductivity.) Though this "forward scattering effect" will not reduce the rate of energy loss from the distribution (except insofar as it alters the form of the distribution) since the net energy loss is only determined by spontaneous phonon emission, this reduction in total scattering rate for hot electrons at lower temperatures will effectively increase the mobility and, hence, the power input into the distribution. This conjectured reduction in the scattering anisotropy has been measured directly and will be discussed in a later section of this paper.

That the value for $\langle\epsilon\rangle$ obtained by this procedure is essentially correct has been verified by observing the strain dependence of the breakdown field.[41] A more complete discussion of this point will be published at a later date.

C. Carrier Concentration vs Electric Field

The variation of steady-state carrier density with electric field can be described by the kinetic or rate equation[42,43]

$$\begin{aligned} 0&=A_T(N_D-N_A-n)+nA_I(N_D-N_A-n) \\ &\quad -nB_T(N_A+n)-n^2B_I(N_A+n) \\ &=A_T(N_D-N_A)+n[A_I(N_D-N_A)-A_T-B_TN_A] \\ &\quad -n^2(B_IN_A+A_I+B_T)-n^3B_I. \end{aligned} \tag{13}$$

Here A_T and A_I are the generation rate of carriers due to thermal excitation and impact ionization, respectively; B_T and B_I are the respective inverse processes, the second being an Auger process. A_T, B_T, and B_I are functions of both the lattice temperature T and the electron distribution function $f(\epsilon)$ whereas A_I is presumably a function only of $f(\epsilon)$. We make the assumption that $f(\epsilon)$ is a function of E and T, and only at low values of E, of N_D and N_A (because of impurity scattering). $f(\epsilon)$ is assumed not to be a function of the state of ionization of the donors. In the earlier discussion of the data in Fig. 9, this latter point was shown to be somewhat incorrect for large values of n, but this will not be of significance for what follows. [For highly doped, highly compensated samples, the situation is much changed. Qualitatively different isothermal[44]

[41] J. J. Hall and S. H. Koenig, Bull. Am. Phys. Soc. **5**, 194 (1960).

[42] Equation (13) as the basic equation governing the carrier density and Eq. (15) as the criterion for breakdown were first proposed by P. J. Price (unpublished).

[43] The same rate equation has been used to study photoconductivity by N. Sclar and E. Burstein, Phys. Rev. **98**, 1757 (1955).

[44] It is possible for thermal oscillations to occur under typical experimental conditions which manifest themselves as an apparent negative resistance when measurements are made with dc equip-

phenomena, in particular negative resistances,[45,46] can in principle occur because of the effect on $f(\epsilon)$ of different ionization conditions of the impurities.]

For values of current density $\lesssim 10^{-3}$ A/cm^2, which corresponds to $n<10^{10}$/cm^3, the terms in n^2 and n^3 in Eq. (13) can be neglected. In addition, since the principle of detailed balance requires

$$A_T(N_D-N_A)=nB_TN_A \tag{14}$$

at thermal equilibrium, it is clear that for $n\ll(N_D-N_A)$, the term A_T in the coefficient of n, Eq. (13), may be neglected. Within this limit,

$$n=\frac{A_T(N_D-N_A)}{B_TN_A-A_I(N_D-N_A)}. \tag{15}$$

For small values of electric field, the second term in the denominator of Eq. (15) may be neglected. The initial variation of n with E then occurs because of a decreasing recombination rate as the electrons heat.[2] For greater values of electric field, A_I increases as B_T continues to decrease, "breakdown" occurring when the denominator becomes very small.[42] The increase of n continues rapidly until other terms in Eq. (13) become important.

It is clear from Eq. (15) that the breakdown field E_B is a decreasing function of the ratio $(N_D-N_A)/N_A$. Figure 11 is a plot, for an assortment of samples *sufficiently pure so that at the onset of breakdown only lattice scattering is important*, of the breakdown field vs $\mu_I(N_D-N_A)$. Here μ_I is the partial mobility in the Ohmic region due to ionized impurity scattering and is taken to vary as N_A^{-1}. (Though of course the "breakdown" field is not really a uniquely defined quantity, when the data for current vs applied field are plotted on a *linear* scale, a critical field at which the current first starts to increase rapidly can be defined within several percent. This is entirely adequate for the present discussion.) Similar behavior has been observed by Bok *et al.*[47] for silicon.

The variation of breakdown field for samples sufficiently doped so that impurity scattering affects the mobility at breakdown is somewhat more complicated, though the physical principle implicit in Eq. (15) is still appropriate. This point has not been appreciated in several instances.[48,49] Yamashita[45] and Chuenkov[50] have discussed the impure case theoretically in some detail.

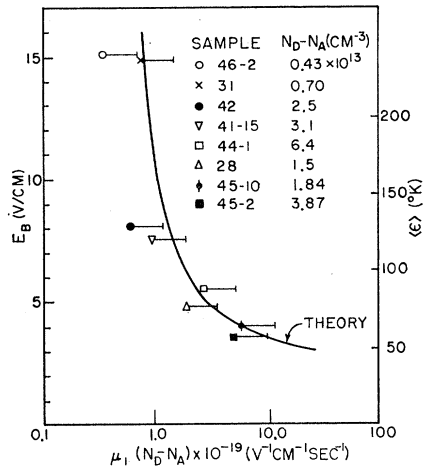


FIG. 11. Variation of breakdown voltage with a quantity that should be proportional to $(N_D-N_A)/N_A$. The quantity μ_I is the contribution to the mobility due to ionized impurity scattering, obtained by subtracting (of reciprocals) from the Hall mobility measured in the Ohmic region the known contriubtions of lattice and neutral impurity scattering. The "tails" on each point represent the uncertainty that arises when assuming $\mu_I \propto N_A$, since the scattering centers may be multiply charged. The theoretical curve, discussed in the text, has as an adjustable parameter its horizontal position.

As discussed in the previous subsection, there is a possible uncertainty of a factor of two in the relation between μ_I and N_A, because the acceptors may be multiply charged. This uncertainty is indicated in Fig. 11 by a horizontal "tail" on each data point. The right end of the tail would correspond to triply-ionized acceptors. It should be noted that the data of Fig. 11 suggest that sample 2 has mainly triple acceptors, while sample 10 has single ionized acceptors. This is consistent with the data for these two samples in Table I from which it is seen that though the values for N_A for the two differ by $\sim$20%, the ionic scattering is different by more than a factor of two.

The variation of breakdown field with impurity concentration, Fig. 11, is seen to be in excellent qualitative agreement with the model of Price.[42] Also shown in Fig. 11 is a theoretical curve. To obtain this, the experimental value for the mobility for sample 2 at breakdown was used to obtain the value for p_B (the subscript refers to the value at breakdown) that corresponds to the 3.6 V/cm breakdown field. The theoretical relation[40] between E and $p(E\propto p^{1/2})$ was then used to

ment [cf. S. H. Koenig and R. D. Brown III, J. Phys. Chem. Solids **10**, 201 (1959)].

[45] J. Yamashita, J. Phys. Soc. Japan **16**, 720 (1961).

[46] R. H. Rediker and A. L. McWhorter, *Proceedings of the International Conference on Solid-State Physics, Brussels, June, 1958* (Academic Press Inc., New York, 1960), Vol. 2, p. 939.

[47] J. Bok, J. C. Sohm, and A. Zylbersztejn, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961), p. 138.

[48] R. H. Rediker, A. L. McWhorter, and C. R. Grant, Quarterly Progress Report, Solid-State Research, Lincoln Laboratory, August 1, 1958 (unpublished), p. 4 ff. Their Table I-1 lists the breakdown field for a variety of (p-type) samples for which N_A and N_A–N_D are given. If the data is rearranged so as to be in terms of $(N_A-N_D)N_A$ and allowance made for the variation of mobility with compensation, then the ordering of the samples with breakdown field is completely consistent with the discussion associated with our Eq. (15).

[49] N. Sclar and E. Burstein, J. Phys. Chem. Solids **2**, 1 (1957).

[50] V. A. Chuenkov, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961), p. 109; Soviet Phys.—Solid State **2**, 734 (1960).

obtain values for p_B vs E_B. It was assumed that every carrier with energy sufficiently great to ionize an impurity does so with equal probability; therefore the integral I of the distribution function, Eq. (6), over all energies greater than the donor ionization energy yields a number proportional to A_I. Since the breakdown condition, Eq. (15), is that $A_I(N_D-N_A)/N_A$ is constant (for a given T), the desired theoretical curve is a plot of E_B vs I^{-1}. This is the curve of Fig. 11, with the (undetermined) horizontal position adjusted for best fit to the data.

It is seen from Figs. 7 and 8 that E_B decreases with an increase in lattice temperature T. Since for a given value of E, $\langle\epsilon\rangle$ and presumably A_I decreases as T increases, B_T would also decrease with increasing lattice temperature (as well as vary with $\langle\epsilon\rangle$). This is consistent with the "giant-trap" mechanism postulated[51] for the capture process (which will be discussed at length in a later section).

The variation with temperature of the current density at the onset of breakdown is determined by the T dependence of A_T (see Eq. (15)), which contains the Boltzmann factor $\exp(-\epsilon_2/kT)$. Figure 7 shows this variation very clearly. In addition, it is seen that the "end" of the breakdown, i.e., the region of current density at which the electric field starts to increase significantly again, is relatively insensitive to the lattice temperature. Referring to Eq. (13), it is seen that for large n, the coefficient of n becomes positive and all terms become large compared to A_T. Then n is given approximately by the expression

$$n=[A_I(N_D-N_A)-B_TN_A]/(B_IN_A+A_I+B_T), \quad (16)$$

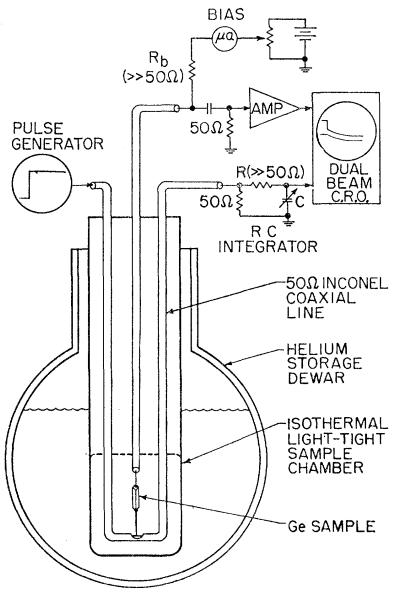


FIG. 12. Schematic diagram of the apparatus used for pulsed measurements to determine values for the kinetic coefficients.

[51] M. Lax, Phys. Rev. **119**, 1502 (1960).

and the field dependence of n for $E\gtrsim 6$ V/cm is determined by the (comparatively slow) variation of the rate coefficients with electric field. The range of current density change for small variations in E near E_B is then seen to decrease exponentially with increasing lattice temperature.

IV. KINETIC COEFFICIENTS

A. B_T

(i) General Considerations

The earliest experimental estimates of B_T were deduced by Burstein *et al.*[52] from the magnitude of the steady-state photoconductive response of germanium (and silicon) at 4.2°K. Though only approximate, it was clear then that the recombination cross section was larger by a factor $\sim 10^3$ than predicted by the theory[53] for recombination with emission of a single phonon.

The accurate determination of B_T for thermal electrons, as a function of temperature, was the original motivation for the present series of experiments; preliminary results have already been published, and their significance discussed in some detail.[1-3] However, B_T is a very difficult quantity to measure by any technique, and data obtained by Ascarelli and Brown[4] using a method different from ours are in some disagreement with our results. It is therefore appropriate to discuss the experimental problems in some detail.

In the present experiments, advantage is taken of the fact that the electron energy distribution follows changes in applied electric field much faster than the carrier density does. The sample is initially biased at a field E_i and current $I_i\sim 50\ \mu$A, or ~ 5 mA/cm^2, corresponding to $n_i\sim 10^{10}$/cm^3. Below 6°K, this is well in the breakdown region. The magnitude of I_i is so chosen that the power dissipation is always less than 250 μW, for which power the sample remains within 0.1°K of the ambient, and/or so that only terms linear in n [Eq. (13)] are important. A fast-falling pulse is applied to the sample to reduce the electric field to some value E_f in the Ohmic region. The current "instantaneously" changes to a value $I_f=I_i(E_f/E_i)(\mu_f/\mu_i)\lesssim I_i/20$, and then decays as the excess carriers recombine. This time dependence is given by the solution of

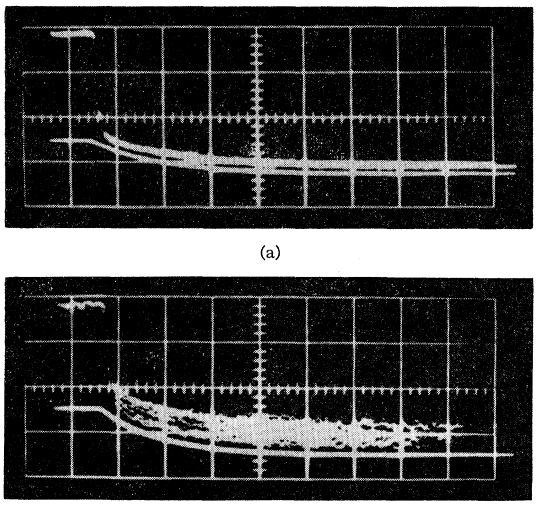
$$dn/dt=A_T(N_D-N_A)-n[B_TN_A-A_I(N_D-N_A)] \quad (17)$$

[derived in an obvious manner from Eq. (13)], which is an exponential decay of n (and, hence, the current) towards the thermal equilibrium value n_T, with a time constant[54]

[52] E. Burstein, G. Picus, and N. Sclar, *Proceedings of the Conference on Photoconductivity, Atlantic City, 1954* (John Wiley & Sons, New York, 1956), p. 401 ff.

[53] H. Gummel and M. Lax, Phys. Rev. **97**, 1469 (1955).

[54] The symbol τ is used in this paper to represent both the scattering relaxation time and the recombination time, in accordance with the practice in the literature. However, these two quantities are sufficiently different physically so that no confusion should arise.



(a)

(b)

FIG. 13. The upper pair of traces shows: (1) the time variation of the current through sample 45–10a when the voltage across the sample is reduced, by means of a fast falling pulse, from the breakdown value to a value of $\sim\frac{1}{4}$ this amount; and (2) the voltage pulse integrated by the variable time constant RC network. By superimposing both curves, using a greater gain setting, the time constant for decay of the current after the initial decrease can be readily measured. These data were taken at 6.35°K. The lower pair of curves shows similar data taken at 5.03°K. Here the measurements are obscured by noise, presumably associated with intrinsic thermal instabilities of the sample (cf. footnote 44). These curves illustrate the worst conditions under which data must at times be taken. The time scale for both sets of curves is 200 nsec/cm; the bias settings, 60 μA. The situation in the lower traces can be somewhat improved by using a larger value of bias.

$$\tau=[B_T N_A-A_I(N_D-N_A)]^{-1}, \tag{18}$$

that is independent of I_i.

The contribution of $A_I(N_D-N_A)$, which will be estimated subsequently, can be neglected in thermal equilibrium below $\sim$10°K. The experimental problem would then appear to reduce to that of measuring a current, with a peak amplitude of $\sim$2 μA, (corresponding to $\sim$0.2 mV in a 100 Ω system) that decays exponentially with a time constant $\gtrsim$50 nsec. The difficulties are: (1) the voltages that need be measured approach within an order of magnitude the thermal noise in the required bandwidth; (2) the samples, when biased in breakdown, are very noisy for reasons that are not completely understood[55]; (3) the amplifiers must be such that overshoot, ringing, etc., must be $\ll$5% of the initial pulse, since the magnitude of the exponential component of the total current change is that small.

The block diagram of the circuit and apparatus used is shown in Fig. 12. The pulse was generated by a mercury switch discharging a charged length of cable.[56] The amplifier was either a series of Hewlett-Packard type 460 wide band amplifiers, appropriately modified to improve the low-frequency response and stability and run from batteries, feeding a Tektronix 541 oscilloscope, or else it was the distributed amplifier of one channel of a Tektronix 551 dual-beam oscilloscope that fed directly into the normal input of the second channel. The time constant was measured by superimposing on the oscilloscope screen the output of a calibrated variable-time-constant RC integrator, also fed by the pulse generator, together with the signal from the sample. Figure 13 shows the two signals when the pulse reduces the field to $\sim E_i/4$. The decay is clear and simple to measure. The difficulties that can arise, however, when the relative amplitude of the exponential is further reduced by a factor of five and the time constant shortened by a similar factor are readily apparent. It was, in fact, not generally possible to obtain reliable results below $\sim$6°K by pulsing down into the Ohmic region, and an extrapolation procedure had to be used.

(*ii*) *Results and Discussion*

Data for τ were obtained as a function of pulse field E_p from $E_p=0$ to $E_p\sim 2E_i$, yielding values for $\tau(E)$ from as close to the linear region as could be measured to values of field $\sim\pm E_i$, i.e., for both directions of current flow through the sample. (The results should be, and were, quite closely, independent of the sign of E.) Since $\tau(E)$ is proportional to $n(E)$ in the range considered [compare Eqs. (15) and (18), as well as reference 2 where this is discussed in detail], the curve for $\tau(E)$ vs E was compared with the curve of $n(E)$ vs E as obtained from Hall effect. The comparison was made by plotting both sets of data on separate logarithmic graphs and sliding one upon the other (with E fixed) until they matched in the region of overlap. A value for τ for thermal electrons was then read off the graphs in an obvious manner. This procedure is illustrated in Fig. 14. The only intrinsic error in this procedure is due to the possible field dependence of the drift to Hall mobility which is ignored.

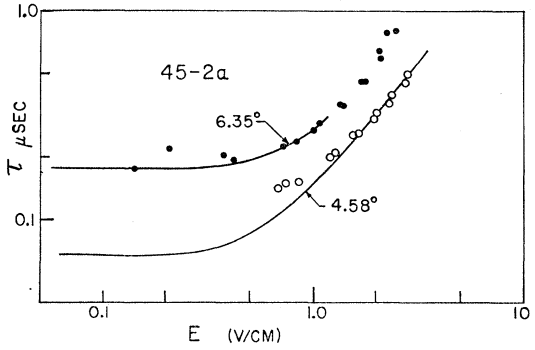


FIG. 14. The extrapolation procedure used to obtain the values of recombination time τ in the linear region is shown here. The solid lines are the field-dependent carrier density curves as obtained from measurements of the Hall constant. The circles are the measured values of τ. The solid curves are moved *vertically* until they best fit the pulse data; the value of τ in the Ohmic region is then read off directly. The 6.35° data, which require no extrapolation, illustrate the reliability of the procedure.

[55] G. Lautz and M. Pilkuhn, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Science, Prague, 1961), p. 141.
[56] R. L. Garwin, Rev. Sci. Instr. **21**, 903 (1950).

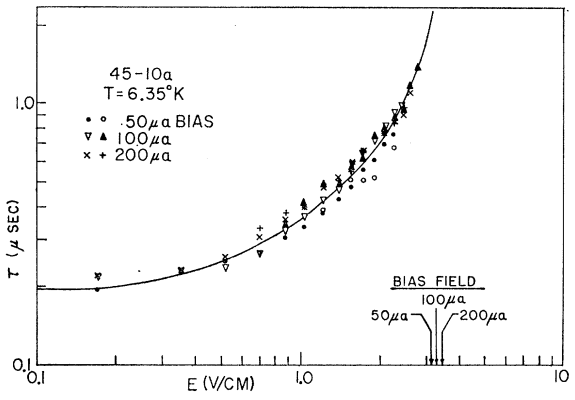


FIG. 15. Values of recombination time taken for various values of initial bias curent, each with either polarity, to show that the results are essentially independent of the initial bias and any associated heating. Since sample 45–10*a* contains both Sb and As (or P) donors, any significant difference in capture cross section for the two types of donors would show up here as a bias dependent value of τ (as discussed in the text). The electric field across the sample for each value of bias is indicated.

In Fig. 15 values for τ vs E obtained for various magnitudes and polarities of the initial bias are shown. The purpose of this procedure was to ascertain if $\tau(E)$ was independent of I_i as expected from Eq. (18). A deviation from this behavior could occur: (1) since the sample contains both As and Sb donors which will be ionized in different relative amounts for different I_i (as the activation energies are different), so that τ would be a function of I_i if the capture cross sections of Sb and As were different; and (2) if heating of the sample by I_i were significant. There is no indication in Fig. 15 of any grossly anomalous behavior, though there appears to be a small systematic variation with I_i. The data of Fig. 14 suggest that heating of $\sim$0.2°K at the larger I_i would be sufficient to explain the variation. An order-of-magnitude difference in the cross section for capture by Sb and As centers[4] is, we feel, precluded.

The results for recombination time extrapolated to the linear region for samples 2 and 10 as a function of temperature are shown in Fig. 16, along with older results[2,3] for sample nWLB 28–6. The latter results have been corrected by improved extrapolation procedures; the disagreement with the measurements made on this sample by Ascarelli and Brown[4] is thereby eliminated. Values for cross section σ_T, obtained from τ using values of N_A determined carefully by the procedures discussed in Sec. II are also plotted. Although for sample 28–6, the effects of strain on the conductivity were not measured, there is evidence from plots of χ_1 vs T (which had the same general shape as that in Fig. 4) that the n-type impurities are a mixture of Sb and As (or P). Fitting a plot of χ_2 vs T to the extrapolated straight line (Fig. 5) and varying the relative values of N_{Sb} and N_{As} (and hence N_A) yields $N_{As} \simeq N_{Sb} \simeq N_D/2 = 7\times10^{12}$ cm^{-3} and $N_A \simeq 1.4\times10^{12}$ cm^{-3}. This is in agreement with the estimate from mobility (Table I). σ_T was obtained by dividing B_T by the rms velocity $(3kT/m^*)^{\frac{1}{2}}$, where m^* is taken rather arbitrarily as the density-of-states mass. Theoretical results are also included. [It should be noted in this connection that Eq. (16) of Lax[51] should read $\sigma_{rms} = (8/3\pi)^{\frac{1}{2}}\sigma$.]

It should be stressed that the major uncertainty in obtaining values for τ for thermal electrons in the region from $\sim$4–6°K is the extrapolation procedure used. However, it is felt that this cannot produce a systematic error greater than $\pm$15%. The uncertainty in the procedure for determining N_A would of course affect the absolute value of σ_T; the spread of the data for several samples suggests that N_A can be determined within an uncertainty of 30%. In addition, *thermal effects play no role*, since the sample temperature never varies more than a tenth of a degree during the pulse. This variation would alter by a small amount the value of the thermal carrier density n_T to which n tends to decay. [Ascarelli and Brown (AB) have suggested that thermal effects are the cause of an apparent discrepancy between their results and those reported here.] However, since the

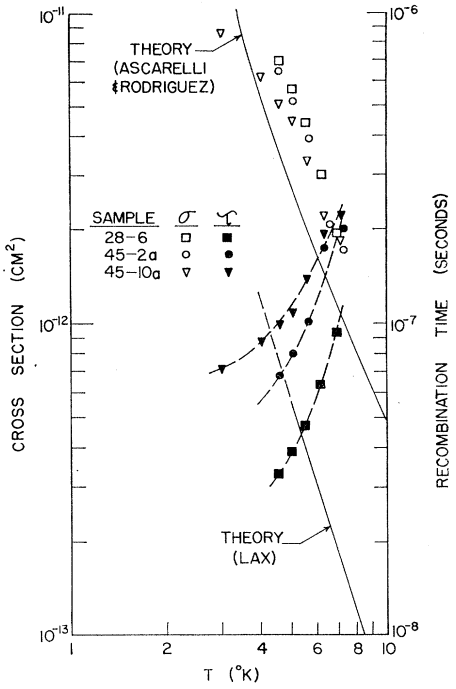


FIG. 16. Data for the recombination time for thermal electrons are shown for several samples vs temperature. The cross section computed using the rms velocity and the density-of-states mass is also shown. The spread at low temperatures between the values of σ for samples 2 and 10 is not to be regarded as significant; it may well be due to the fact that the drift to Hall mobility ratio in the neighborhood of the Ohmic range was assumed to vary with field in a similar manner for both samples. The theoretical results of Lax (reference 51) and the more recent results of Ascarelli and Rodriguez (reference 75, 76) are also shown. The latter papers appeared in print after the bulk of the present paper had been written; they are discussed in Appendix D. (The correction to the results of Lax, suggested in reference 75, was not applied because: (1) the factor of four to correct for the multiplicity of the conduction band is not necessary; and (2) the correct dielectric constant for Ge was used by Lax to obtain his Fig. 1, contrary to the statement in the caption.)

magnitude of n_T is so small as to be unobservable, such variations during the pulse are insignificant.

The results of AB were obtained by pulsing (~ 2 μsec) the sample to a high conductivity, then measuring the conductivity of the sample vs time after the end of the pulse by observing its influence on the Q of a microwave cavity. The major advantage of this method is that the microwave "probe" field can be kept sufficiently low so as to be always within the linear conductivity range. The main difficulty with their experiments, as performed, is that there is no independent procedure for determining N_A. It is claimed that if the exponential decay in Q, observed $\sim 10^{-7}$ sec after the dc pulse has been switched off, is extrapolated backward to zero time, that the intercept is proportional to

$$N_A/[(N_A/n_i)+1]. \tag{19}$$

Therefore if n_i, the initial carrier density which the pulse produces, is $\gg N_A$, a value for N_A can be obtained (after relating the scale for Q to carrier density). This procedure is, in fact, incorrect for several reasons:

(1) Equation (19), corresponding to Eq. 4(b) of the first reference[4] to AB, is obtained by solving an equation equivalent to our Eq. (17) with the term [see Eq. (13)]n^2B_T included. The other term in n^2, namely, $n^2B_IN_A$, has been *arbitrarily* dropped, with the argument that the experimental data fit the solution so obtained. The term $n^2B_IN_A$ must be included, and if $n_i \gg N_A$, so must the term n^3B_I. Solving the resulting equation (cf. Appendix C) replaces expression (19) by

$$n_i\left[1+\frac{B_In_i}{B_T}\right]^{-B_IN_A/(B_IN_A-B_T)} \times\left[1+\frac{n_i}{N_A}\right]^{B_T/(B_IN_A-B_T)}, \tag{20}$$

which, if $B_IN_A \gg B_T$, reduces to

$$N_A/[(N_A/n_i)+(B_IN_A)/B_T]. \tag{21}$$

N_A then drops out of Eq. (19) altogether; the intercept does not depend on N_A at all. It will be subsequently shown from our data and that of AB that $B_IN_A \gg B_T$, even for samples with lower N_A than those measured by AB, so that the values for N_A quoted by AB are an underestimate.

(2) In the region of high currents and nonexponential decay, as well as the beginning of the exponential region (Fig. 4 of AB), the conductivity of the sample is sufficiently high so that skin effect distorts the results.

(3) The value of current to which the samples are pulsed corresponds to $\sim 10^2$ W/cm^3 dissipation, or during the 2-μsec pulse, to a temperature rise of the order of several degrees Kelvin. This has been discussed by AB; all data required corrections for this heating.

The somewhat arbitrary (and incorrect) method of estimating the acceptor concentrations in samples BTL–1 (Sb-doped) and LL–2 (As-doped) have led AB to the conclusion that the recombination cross section for electrons with ionized Sb donors is a factor of ten larger than for As donors. The experimental recombination times for both samples differed only by 25%, but the alleged acceptor concentrations differ by a factor of eight. It is considerably more probable that the acceptor concentrations are in fact approximately equal, and, therefore, that the cross sections for As and Sb are as well. We have, for example, measured the Hall mobility ($\sim 5\times 10^5$ cm^2/Vsec) on a sample cut from the same crystal as BTL–1, kindly furnished by Dr. Ascarelli. We find $N_A \sim 3$ times that for our samples 2 and 10, or $\sim 2\times 10^{12}$/cm^3. AB quote 4.6×10^{11}/cm^3 for BTL–1, and $\sim 3.5\times 10^{12}$/cm^3 for the As-doped sample.

The estimate for N_A by AB for their arsenic-doped sample is probably nearly correct, though somewhat high, judging by our experience with large numbers of crystals in the 20–40 Ω cm range from various sources. It is the rare crystal (such as that from which 10 and 2 were cut) that has a mobility at 4.2°K outside the range 3.5 to 7×10^5 cm^2/Vsec, corresponding to a typical acceptor variation from sample to sample of about a factor of 2. Within this uncertainty, the reinterpreted data of AB are entirely consistent with the statement that the capture cross sections of As and Sb donors are equal.

We have attempted a complete set of measurements, both dc and pulsed, on several As-doped samples and one P-doped one (46–2, Table I). However, to date, we have not been successful in obtaining an As-doped sample of sufficiently low compensation. N_A for the P-doped sample was marginal. However, the compensation was so close that the resultant high breakdown field required excessive power dissipation when biasing the sample for pulse measurements. The capacitive feedthrough associated with the large pulse voltages also caused excessive ringing of the circuits; consistent, reliable measurements were not possible.

Expression (21) can be used together with the value for N_A quoted by AB and our conjecture for its proper value to obtain a value for B_IN_A/B_T. For sample BTL–1, the value given by AB for N_A first must be reduced by a factor of ~ 3 since the *ad hoc* assumption of a mobility at breakdown of 10^5 cm^2/Vsec made by AB to convert current to carrier density is low by that amount. Then, using values from the preceeding discussion,

$$\frac{N_A}{(B_IN_A)/B_T} \simeq \frac{2\times 10^{12}}{(B_IN_A)/B_T} = (4.6/3)\times 10^{11} = 1.5\times 10^{11}, \tag{22}$$

and

$$(B_IN_A/B_T) \simeq 12. \tag{23}$$

From this, and the result for B_TN_A measured by AB,

one obtains for thermal electrons at 4.2°K,

$$B_I \simeq 0.7\times10^{-16}\ \text{cm}^6/\text{sec}. \quad (24)$$

This result is larger by a factor of ~4 than our results (in the next section) for hotter electrons, as it should be. The agreement is, in fact, remarkably good.

In the "giant trap" mechanism of Lax,[51] which would appear to be the appropriate description of the recombination mechanism, excited donor states within $\sim kT$ of the conduction band edge make the main contribution to the capture cross section. Since these states, as distinct from the ground state, are the same for As and Sb donors, the capture cross sections should be the same. The experiment of Koenig and Hall on p-type germanium,[57] in which the acceptor ionization energy was varied by the application of a large uniaxial stress, shows that at least the energy dependence of B_T is independent of the impurity ionization energy.

Data of ours to 2°K, purporting to show that the recombination time tends to saturate[58] below ~4°K are reproduced in the paper of Lax.[51] However, it is somewhat difficult to make a strong argument for the validity of the extrapolation procedure used; it is not clear what procedure to follow. Below ~3.8°K, the linear conductivity becomes, to all purposes, immeasurably low so that it is difficult to determine the extent of the Ohmic range. The mobility tends toward a constant value (Fig. 2) and the energy losses due to lattice scattering decrease [cf. Eq. (11), noting that the acoustic relaxation time $\tau \propto T^{-\frac{3}{2}}$] as T decreases, so that it might appear that the Ohmic range would decrease steadily. Were this the case, our extrapolation procedure would not be appropriate, since it assumes that below 4°K the extent of the Ohmic region remains constant. On the other hand, another energy loss mechanism exists which has not been considered, and which will become more significant as impurity scattering increases: an energy loss due to intervalley scattering cause by ionized impurities. If the mechanism suggested by Weinreich *et al.*[59] for intervalley scattering by ionized impurities, to wit: a capture of electrons in excited orbits with subsequent reemission in another valley, be correct, the re-emitted electrons will have lost any memory of their initial energy and thereby become thermalized in one collision. (The collision probability may be energy-dependent, however.) Such collisions are more frequent than captures (the rate of which increases as T decreases), and below 4°K begin to approach within an order of magnitude of the lattice collision frequency. But lattice collisions are not as lossy per collision so that this intervalley mechanism cannot be neglected.

Optical excitation of carriers might be used to produce additional carriers, but it would be hard to ensure that one were measuring a thermal distribution.

There are other arguments which would, however, be consistent with a saturation of the cross section at lower temperatures. At the end of Sec. II D, the variation of ϵ_2 with impurity concentration was discussed and ascribed to the overlap or distortion of excited impurity states near the conduction band edge. From Figs. 3 and 6, for example, it is seen that the variation is of the order of 2×10^{-4} eV corresponding to thermal energies ~2°K. Since states $\sim kT$ from the band edge do contribute most to capture, and these states (for $T \lesssim 4$°K) are apparently distorted to the extent of no longer being bound states, it is believable that the cross section does stop increasing as T is lowered below several degrees Kelvin.

The variation of B_T with electron energy has been discussed previously, but several additional points deserve mention. Figure 17 shows the variation of τ [Eq. (18)] with electric field for a typical case. In the region delineated by arrows, the Hall mobility varies quite closely as $E^{-1/2}$ so that the mean electron energy $\langle\epsilon\rangle$ is increasing essentially linearly with E. The region between 2 and 4 V/cm, where τ varies as a power of E, is thought to be the range where the term $A_I(N_D-N_A)$ in Eq. (18) may be neglected, and, therefore, to represent the variation of B_T with electron energy. The slope corresponds to a 1.8 power (in reasonable agreement with the giant trap theory[51]), while the variation with T of B_T at thermal equilibrium is closer to T^3. The difference, of course, is the separate contribution to B_T of the lattice temperature which was invoked in Sec. III C to explain the temperature dependence of the breakdown voltage.

It is believed that impact ionization does not make a significant contribution to τ below ~10°K until fields such that deviations from a simple power law (Fig. 17) set in. This is justified by the data on p-germanium referred to earlier.[57] In this work the form of the varia-

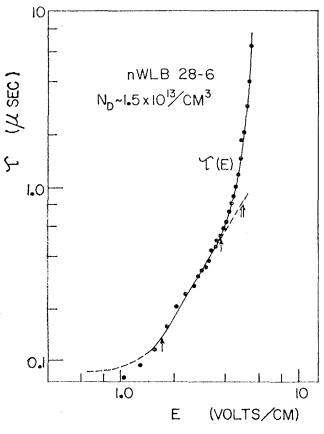


FIG. 17. The variation of recombination time τ with electric field E. In the region delineated by single arrows, the Hall mobility varies quite closely as $E^{-\frac{1}{2}}$, so that the mean electron temperature is increasing essentially linearly with E. It is felt (see text) that impact ionization may be neglected in this region; the energy dependence of B_T is then given by the slope of τ. The double arrow is the point where, by extrapolation, the impact ionization rate roughly equals half the recombination rate.

[57] S. H. Koenig and J. J. Hall, Phys. Rev. Letters **5**, 550 (1960).

[58] It has recently been reported by R. S. Levitt and A. Honig that the "giant trap" recombination time in Si is temperature independent between 1.2 and 4.2°K [Bull. Am. Phys. Soc. **6**, 482 (1961); J. Phys. Chem. Solids **22**, 269 (1961)].

[59] G. Weinreich, T. M. Sanders, Jr., and H. G. White, Phys. Rev. **114**, 33 (1959).

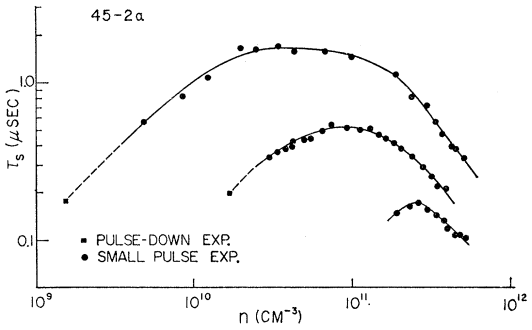


FIG. 18. Variation with carrier density (as determined from Hall effect) of the time constant for response of the sample to a small voltage pulse. The temperatures for the three curves, reading from top to bottom, are 6.19, 7.14, and 8.39°K.

tion of current density with electric field is only altered by a variation of acceptor binding energy for electric fields greater than ~ 7 times the Ohmic limit. [It is clear that the ideal way to separate the B_T and A_I in Eq. (18) is to vary the relative concentration of N_D and N_A, maintaining, however, sufficient purity to have the scattering due only to phonons. This is, in principle, technically feasible; in practice, we have been unable to obtain samples which have simultaneously sufficient homogeneity, the required impurity concentration and the proper orientation to perform these experiments.]

B. A_T

The thermal generation rate of carriers per neutral donor, A_T, may be obtained by the principle of detailed balance from a knowledge of n_T, the equilibrium carrier density, and $\tau=(B_T N_A)^{-1}$ for thermal electrons. The result at 4.2°K is $A_T=n_T/\tau(N_D-N_A)=1.6\times 10^{-2}$ sec^{-1} for Sb donors. That is, the mean time for a given donor to be thermally ionized is 60 sec. (The characteristic time for n_T to change if T is changed is, of course, given by τ which is $\ll 60$ sec). This time decreases exponentially to $\sim 10^{-4}$ sec at 10°K. The mean time for ionization of As donors at 4.2°K is a factor $\sim 10^4$ longer than for Sb. This long time is of significance below ~ 6°K in the transient response of the conductivity of a two-terminal, initially unbiased, sample to a pulse large enough to cause breakdown. The response, a "delayed breakdown," cannot be understood in terms of the time-dependent form of Eq. (13); it is related to the time necessary for the space charge layer at the contact, through which electrons enter the sample, to alter in size. This phenomenon has been discussed at length elsewhere.[60]

Again, by detailed balance, the process of thermal ionization must be the exact inverse of thermal capture, i.e., a multistep process in which the electron climbs the ladder of excited donor states, ultimately entering the conduction band from a fairly large orbit.

[60] S. H. Koenig, *Proceedings of the International Conference on Solid State Physics, Brussels, June, 1958* (Academic Press Inc., New York, 1960), Vol. 1, p. 422.

C. B_I, A_I

(i) *General Considerations and Results for* B_I

B_I represents recombination due to an Auger process, i.e., two carriers interacting in the vicinity of an ionized donor, one being captured, the other carrying off the excess energy. The process is in detail the inverse of impact ionization.

Consider the time-dependent form of Eq. (13):

$$dn/dt=a-bn-cn^2-dn^3, \tag{25}$$

where

$$a=A_T(N_D-N_A),$$
$$b=B_T N_A+A_T-A_I(N_D-N_A)\simeq B_T N_A-A_I(N_D-N_A),$$
$$c=B_T+A_I+B_I N_A\simeq B_T+B_I N_A,$$
$$d=B_I.$$

Since at breakdown, $B_T N_A\sim A_I(N_D-N_A)$, one has $A_I\ll B_T$ [see Eq. (15)].

If d could be measured, or if c could be determined and B_T shown to be negligible, a value for B_I would be obtained. Since both c and d are coefficients of terms nonlinear in n, and since these terms must be significant in Eq. (25) for the coefficients to be measurable, the appropriate solution to Eq. (25) will not be a simple exponential unless the equation is somehow linearized. The experimental procedure used to accomplish this was to bias the sample at some current, corresponding to a carrier density n_0, and then to apply a pulse sufficiently small so that the change in carrier density δn is $\ll n_0$. The experimental arrangement is essentially that of Fig. 12. The expression for the time rate of change of $\delta n=n-n_0$ is readily shown to be

$$d(\delta n)/dt=(-b-2cn_0-3dn_0^2)\delta n, \tag{26}$$

since

$$a-bn_0-cn_0^2-dn_0^3=0. \tag{27}$$

Terms nonlinear in δn have been neglected, which for small pulses is justified. The response is now exponential with a time constant τ_s given by

$$\tau_s^{-1}=(b+2cn_0+3dn_0^2), \tag{28}$$

or by substituting Eq. (27) in (28),

$$\tau_s^{-1}=a/n_0+cn_0+2dn_0^2, \tag{29}$$

Experimental results for τ_s are shown in Fig. 18 for sample 2. Values for B_I deduced from the data of Fig. 18 with the aid of Eq. (29), assuming $c\simeq B_I N_A$, (this still remains to be justified) are graphed in Fig. 19. The range of n_0 spanned is such that the sample is for the most part in the breakdown region. (Biasing at higher values of n_0 produces too much heating, at lower values the signal becomes too small.)

From Eq. (29) it is clear that τ_s as a function of n_0 must pass through a maximum. Since the electric field and therefore the distribution function change very

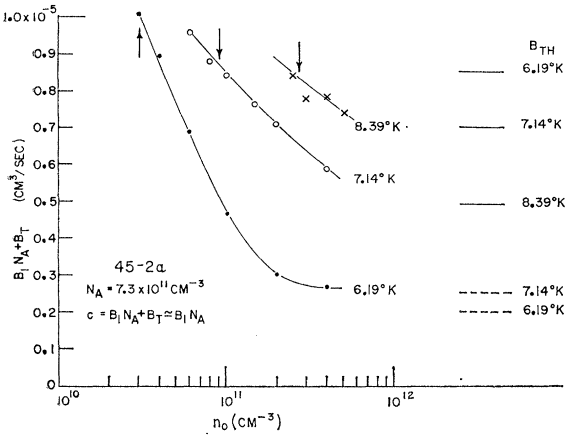


FIG. 19. Variation of $c=B_I N_A+B_T$ with carrier density n_0 for several temperatures. These data can be used to show that B_T can be neglected compared to $B_I N_A$, that c therefore is just the Auger recombination time. The solid lines at the right show the value of B_T for thermal electrons at the temperatures indicated, the dashed lines the value at breakdown, corresponding to $n_0 < 10^{11}$ cm^{-3}. Since B_T decreases monotonically as n_0 increases, it is clear that B_T makes but a small contribution to c. The arrows indicate the position of the maximum of τ_s (cf. Fig. 18).

little over the range of n_0, one may differentiate Eq. (29) with respect to n_0 considering c constant. [Only b varies rapidly with n_0 and this was eliminated to obtain Eq. (29).] If d is neglected, the result for the maximum τ_s and the respective value of carrier density simplifies to

$$(\tau_{s_{\max}})^{-1}=2(ac)^{\frac{1}{2}}, \quad n_{\max}{}^2=a/c, \quad \text{and} \quad b=0. \tag{30}$$

(This is valid for the two lower temperatures of Fig. 18. If d cannot be neglected, the equations become more complex and it is more difficult to follow the arguments given below. However, all data have been reduced with the complete formulas.)

The temperature variation of $\tau_{s_{\max}}$ is contained mainly in the exponential dependence of $a=A_T(N_D-N_A)$ on T^{-1}. It has previously been shown[1,60] that the (rapid) T dependence of $\tau_{s_{\max}}$ is in strict accord with the predictions of Eq. (30).

The sense of the variation of B_I with lattice temperature T and "electron temperature" T_E may be obtained from Fig. 19. For fixed T, it is clear that B_I decreases quite rapidly as T_E increases. Since at $n_{\max}$ (indicated by arrows in Fig. 19) $b=B_T N_A-A_I(N_D-N_A)=0$, and since B_T for fixed T_E decreases as T increases (cf. Sec. III C), T_E at $n_{\max}$ must decrease with increasing T. The fact that $\tau_{s_{\max}}$ occurs at lower electric fields for greater values of T of course implies the same result. Since B_I at $n_{\max}$ is reasonably independent of T, the foregoing requires that B_I decrease also as T increases. The qualitative dependence of B_I on energy and lattice temperature is thus the same as that of B_T.

(ii) General Considerations and Results for A_I

Values for A_I in the region of breakdown can be roughly estimated by extrapolating (cf. Fig. 16) the contribution that $B_T N_A$ makes to

$$\tau^{-1}=B_T N_A-A_I(N_D-N_A)$$

which for $n_0 \ll n_{\max}$ equals τ_s^{-1}. (As pointed out earlier, the appropriate assortment of samples to make a direct separation of B_T and A_I is not available.) From the data of Fig. 17 at the point indicated by a double arrow, one has $A_I(N_D-N_A)\simeq B_T N_A/2=10^{+6}$ sec, or $A_I=5\times10^{-8}$ cm^3/sec. Converting this to a cross section, assuming the electron "temperature" to be $\sim$50–100°K (cf. section III B; this means, considering the increase of density of states with energy, that the larger fraction of electrons have sufficient energy to ionize donors) gives

$$\sigma_I \simeq 5\times10^{-8}\, v^{-1} \simeq 5\times10^{-15} \text{ cm}^2 = 2\times10^{-2}\sigma_0,$$

where σ_0 is the geometric cross section of the first donor Bohr orbit and v is the appropriate velocity.

A more significant and unexpected property of A_I can be obtained from the variation of B_I with T for constant $n_0 \gg n_{\max}$ Fig. 19. From Eqs. (29), (30) the terms $cn_0+2dn_0^2$ then become $\gg(a/n_0)$. For Eq. (27) to hold now requires $b<0$; therefore $n_0 A_I \gg A_T$. Solving for n_0 gives

$$n_0 \simeq \frac{A_I(N_D-N_A)}{B_I N_A[1+(n_0/N_A)]}. \tag{31}$$

For fixed n_0, T_E decreases significantly as T increases (this was true even at $n_{\max}$ vs T), and B_I increases; therefore A_I must increase quite rapidly with T in this range (even as T_E decreases) to maintain n_0 constant. A typical magnitude for A_I in this region is, using Eq. (31), $A_I \sim 10^{-7}$ cm^3/sec, of the same order as the previous estimate.

It has been tacitly assumed all along that the mechanism for impact ionization is one of impact excitation from the ground state directly to the conduction band. A_I would then be a function only of the electron distribution and not explicitly of T, contrary to the above results. A mechanism whereby A_I varies with T will be discussed below. None of the arguments that have been made in earlier sections, however, are influenced significantly by the T dependence of A_I.

(iii) Discussion

The quantity $c\simeq B_T+B_I N_A$ can be shown to be essentially equal to $B_I N_A$. The solid lines on the right side of Fig. 19 indicate the values for B_T for thermal electrons. Extrapolating, as in the preceeding subsection, to obtain values for B_T near $\tau_{\max}$ yields results which are shown in Fig. 19 by dashed lines. It is clear that B_T at $n_{\max}$ is small compared to $B_I N_A$ and may be neglected without significant error. For values of $n_0 > n_{\max}$, B_T and B_I both decrease, and presumably B_T may be ignored over the entire range of the data in Fig. 19. This demonstrates that we are, in fact, measuring the Auger recombination process when determining

the coefficient of the n^2 term in the rate equation, and not simply a bi-molecular rate for thermal recombination. The *ad hoc* procedure of AB referred to above (cf. Sec. IV B) of setting $B_I=0$ is therefore without justification.

The experimental values for B_I are in good agreement with the theoretical estimates for at least one recombination mechanism. Lax[51] has estimated the cross section for Auger recombination, assuming the mechanism to be an initial capture of one of the colliding electrons into a large orbit with a subsequent cascade, identically as in the mechanism for B_T. For $T=6°$K and an electron energy $\sim 50k$ ergs, the estimated Auger capture cross section is $\sigma\sim 5\times 10^{-24} n_0$ cm^2. For this set of parameters, the experimental value of B_I when converted to a cross section gives $\sigma\sim 2\times 10^{-24} n_0$. If the cascade capture mechanism is the main one, as the closeness of theory and experiment might suggest, the implications for the mechanism of impact ionization are interesting. By detailed balance, impact ionization by thermal electrons would take place from the upper most levels, rather than from the ground state. For lattice temperatures such that the excited donor states are populated with reasonable probability, impact ionization by hot electrons could also occur in the same manner. For this mechanism, the typical energy of an electron involved in an ionizing collision would be much less than the donor ground-state binding energy; therefore A_I ought not be a particularly rapid function of the mean energy of the distribution, but rather of the lattice temperature (which alters exponentially the population of the upper levels). The dependence on lattice temperature is in agreement with the experimental findings for $T\gtrsim 6°$K discussed above. For lower temperatures, the "traditional" impact ionization mechanism should be dominant. The agreement of the theoretical curve with the experimental data in Fig. 11 is evidence for this.

In principle, impact excitation of donors to higher bound states may occur, with subsequent ionization from these levels, i.e., that the distribution of excitation of donors may be a function of electron density and "temperature," as well as of lattice temperature. This would introduce into the rate equation a positive term in n^2; it would also imply that part of the cascading during the capture process was influenced by electron collisions. However, even if the impact ionization rate were $\sim 10^3$ times greater than the thermal generation at, say, 6°K (cf. Sec. IV A) and the excitation rate to excited states of the same order of magnitude, any particular donor would only be excited once every millisecond. This time is much longer than any reasonable value for the lifetime of an excited state,[61] so that this type of cascade ionization is negligible.

It is possible to make an estimate of the rate of Auger recombination, for thermal electrons, due to the mechanism that is the inverse of the "traditional" ionization process. At thermal equilibrium,

$$n_T A_I(N_D-N_A-n_T)=n_T^2 B_I(N_A+n_T), \quad (32)$$

and

$$A_I\simeq\langle A_I\rangle \exp(-\epsilon_2/kT), \quad (33)$$

where $\langle A_I\rangle$ is an average impact ionization rate for carriers with sufficient energy to ionize. From the previous subsection, $\langle A_I\rangle\sim 10^{-7}$ cm^3/sec. Combining Eqs. (1) and (32),

$$B_I\simeq 10^{-7}h^3/2(2\pi m^* kT)^{\frac{3}{2}}=2.3\times 10^{-22}T^{-\frac{3}{2}} \text{ cm}^6/\text{sec}. \quad (34)$$

At 6°K, $B_I\simeq 1.5\times 10^{-23}$ for thermal electrons and perhaps an order of magnitude less for the distribution near breakdown. The experimental values of B_I, from Fig. 19, are $\sim 5\times 10^{-6}/N_A=0.7\times 10^{-17}$ cm^6/sec. This result is orders of magnitude greater than the result of Eq. (34); the mechanism for Auger recombination, as for direct recombination, is then one involving a "giant trap."

V. MISCELLANEOUS

A. The Conwell-Ryder Results[37,62,63]

Conwell[37] had initially interpreted some early data of Ryder[62] (Fig. 20) as a variation of mobility with electric field. Subsequently, the fairly sharp rise was reinterpreted as evidence for the onset of impact ionization.[63] These data may be compared with the 24°K data of Fig. 7, which are for a sample that at 24°K has very closely the theoretical lattice mobility. Ryder's sample was more impure, having a mobility $\sim\frac{1}{2}$ as great.

The deviations from Ohm's law for sample 10 are much gentler than those for Ryder's sample. This difference can be readily understood if Conwell's initial ideas on the field-dependent mobility are correct.

From the data of Fig. 19 and the associated discussion, it is seen that for fields of a few V/cm, as T increases, the electron energy decreases, A_I increases, and the total impact ionization rate remains at all times much greater than the thermal rate. By extrapolation, therefore, the initial rapid rise in conductivity in the Ryder data cannot be due to the sudden onset of impact ionization; *this is the mechanism that generates carriers in the Ohmic region.* Rather, the initial

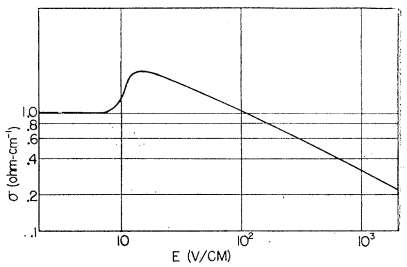


FIG. 20. Early data of Ryder (reference 62) taken at 21°K.

[61] E. O. Kane, Phys. Rev. **119**, 40 (1960).

[62] E. J. Ryder, Phys. Rev. **90**, 766 (1953).

[63] E. M. Conwell, Phys. Rev. **94**, 1068 (1954).

increase in conductivity at 24°K for sample 10, and partly for Ryders' sample, is, we feel, due to a steady increase in A_I coupled with a decrease of B_I as the mean electron energy increases. The remaining contribution to the initial conductivity increase in Ryder's sample is due, as per Conwell's initial suggestion, to the decrease of impurity scattering as the electrons heat. The subsequent decrease in conductivity at the higher fields in both samples is, of course, due to the mobility finally decreasing (as $E^{-\frac{1}{2}}$).

B. The Magnetic Field Dependence of the Breakdown Field

The condition for breakdown [cf. Eq. (15) and related discussion] should depend on magnetic field H only through the dependence of the "electron temperature" on H, provided H is not so large that quantization effects need be considered. Thus the probability, A_I, that an electron colliding with a neutral donor will ionize it should not depend on the fact that the electron path before impact was somewhat curved. Similarly, so long as the curvature of the electron paths is small compared to that of the orbits that contribute most to B_T, B_T should be independent of H. These conditions are met at 4°K for H less than a few thousand oersteds. For a given distribution function, the rate of energy loss to phonons should also be independent of the presence of H, since the mean time for an electron to collide with a phonon is not effected by the curvature of the electron path. The only way that H can influence the value of E needed to produce a given carrier density n is by reducing the rate at which the distribution gains energy from E. The effect would enter in the mobility in the left-hand side of Eq. (7) identically as it would in the magneto resistance, since both are a measure of the change in the drift velocity of the distribution along the electric field direction with H. One may then write

$$\mu[1-\xi(\mu H/c)^2]E_H^2=\mu E_B^2, \tag{38}$$

where

$$E_H\equiv E_B+\Delta E$$

is the electric field that will maintain the energy gain in magnetic field H the same as that at E_B when $H=0$. The constant ξ, of the order of unity, will depend, in general, upon whether H is parallel (longitudinal case) or transverse to the applied electric field. The mobility μ (now scalar) is the component of the Hall mobility tensor measured when E and H are mutually perpendicular and both along a cube-edge direction. Equation (38) will not be valid beyond the region of quadratic magnetoresistance, i.e., for $(\mu H/c)>\frac{1}{4}$. In addition, the implicit assumption is made in writing Eq. (38) that the distribution function is not altered in form by the magnetic field.

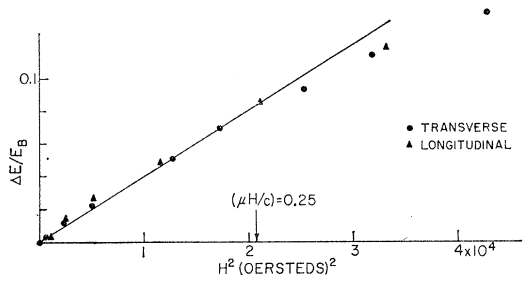


FIG. 21. Variation of the fractional change in breakdown field ($\Delta E/E_B$) with both longitudinal and transverse magnetic field.

The current for fixed carrier density will decrease because of the mobility decrease. Experimentally, however, it is most convenient to plot the variation of E_H with H at constant current rather than at constant carrier density. If the current is restricted to the region where it rises "vertically" with E_H, i.e., where the variation of E_H over a wide range of current is very small compared to its variation with H, then the variation of E_H with H constant current or constant carrier density will be the same.

For $\Delta E/E_B\ll 1$, Eq. (38) reduces to

$$\xi(\mu H/c)^2=2(\Delta E/E_B). \tag{39}$$

Figure 21 is a plot of $(\Delta E/E_B)$ vs H^2 for both longitudinal and transverse fields. The point where $(\mu H/c)=\frac{1}{4}$ (where deviations from a straight line should become observable) is indicated. It is seen that the longitudinal and transverse effects are essentially equal in magnitude.

The values for ξ_T and ξ_L (the subscripts representing transverse and longitudinal magnetic field, respectively) are related to the mobility anisotropy K and a parameter N by the expressions [cf. Eqs. (5.30) and (5.31) of Brooks[14]]

$$\begin{aligned}\xi_L&=N[2(K-1)^2(2K+1)/3K(K+2)^2],\\ \xi_T&=N[(2K+1)^2/3K(K+2)]-1.\end{aligned} \tag{40}$$

N is a combination of averages of different powers of the scattering relaxation time over a distribution function related to the momentum distribution.[14] Both N and the value for K derived by using Eq. (40) are very sensitive to the exact form of the distribution function and the energy dependence of the scattering time, so that, though Eq. (6) was useful for estimating electron mean energies, it is not adequate for reliably computing the magnetoresistance anisotropy.

A better procedure is to obtain an experimental value for $\xi_T=\xi_L=\xi$ by fitting the data of Fig. 21 to Eq. (38). The result is $\xi=0.65$, $K=4.2$, for acoustic phonon scattering at breakdown. The value of K is much less than the value for acoustic phonon scattering for thermal electrons at higher temperatures[64] (~ 17), indicating that the mean scattering time for electrons in the "heavy" direction is several times longer than for the "light" direction, rather than being roughly equal (the mass ratio is 19.7). This is nothing more than the "for-

[64] See C. Herring, T. H. Geballe, and J. E. Kunzler, reference 6.

ward scattering effect" postulated and discussed in Sec. III B. That discussion showed that the phonon needed to scatter a "typical breakdown electron" across the long diameter of the energy ellipsoid must have energy $\sim 5kT$; then roughly, the anisotropy K should be reduced from 17 by a factor ~ 5, in qualitative agreement with the experimental results.

C. Anisotropy and Valley Repopulation

(i) General Considerations

If the electric field is applied, not along a ⟨100⟩ direction as has been the case for the data discussed so far, but along an arbitrary direction with respect to the crystallographic axes of the sample, the electron distribution and therefore the carrier density in each of the four ⟨111⟩ valleys will, in general, be different. Breakdown will occur in the "hottest" valley first, and in the other valleys only when the field is further increased. If a mechanism for intervalley scattering exists that involves dissipation, such as emission of a phonon, the carriers will scatter preferentially out of the hottest valley in order to degrade the energy; a "repopulation" of the valleys will occur.

As a first approximation, it may be assumed that the criterion for breakdown [cf. Eq. (15) and the associated discussion] is valid for the hottest valley by itself. (Implicit in this is the assumption that an electron produced by impact ionization emerge in the same valley as the colliding electron.) That is, breakdown occurs in a valley when the distribution function attains a particular mean energy, a particular value of the field parameter p. Since the loss rate from the distribution depends only on the distribution and is independent of how the distribution is established [see Eq. (11)], the entire variation of breakdown field with direction is contained in the expression [cf. Eq. (10)].

$$\mathbf{E}_B \cdot \boldsymbol{\mu}_h \cdot \mathbf{E}_B = \text{const.} \tag{41}$$

Here $\boldsymbol{\mu}_h$ refer to the mobility tensor for the hottest valley; its principle directions are the same as those of the mass tensor for this valley.

The magnitude of the current density throughout the non-Ohmic region will depend on the amount of repopulation of the valleys; all the valleys will, of course, contribute to the measured current. The longitudinal current *per electron* $j^{(i)}$ in the ith valley is given by

$$j^{(i)} = \frac{\mathbf{E}\cdot\mathbf{j}^{(i)}}{E} = e\frac{\mathbf{E}\cdot\boldsymbol{\mu}^{(i)}\cdot\mathbf{E}}{E}. \tag{42}$$

The quantity $\boldsymbol{\mu}^{(i)}$, a function of $p^{(i)}$, is the mobility tensor for the ith valley. For every value of E, each $j^{(i)}(E)$ may be obtained (within a normalization factor that cancels in what follows) by first finding the field $E_{100}{}^{(i)}$ and the associated current $j_{100}{}^{(i)}$ for a ⟨100⟩ oriented sample that corresponds to the same power input *per carrier* as for the ith valley. With the aid of Eq. (42), $j^{(i)}$ is readily obtained. If the orientation of the electric field is such that the four valleys are divided into but two classes, H for "hot" and C for "cold," then by summing the contributions of the current per electron, using a parameter to represent the relative repopulations of the hot and cold classes, it is possible to experimentally determine the repopulation as a function of E. One need only compare the experimental current variation with that computed from the ⟨100⟩ orientation data by the procedure implied above.[65] More specifically, the total current j_T in either the ⟨110⟩ or ⟨111⟩ valleys is given by

$$j_T = j_{100}{}^H(E_{100}{}^H/E)(1-f)a_H + j_{100}{}^C(E_{100}{}^C/E)(1+f)a_C, \tag{43}$$

where

$a_H = \frac{3}{4}$, $a_C = \frac{1}{4}$ for the ⟨111⟩ sample,

$a_H = a_C = \frac{1}{2}$ for the ⟨110⟩ sample.

The quantity f represents the fraction of carriers transferred from the hot class of valley. The value of j_T so obtained may be compared with the measured values of current density for the ⟨110⟩ and ⟨111⟩ samples to obtain the variation of f with E.

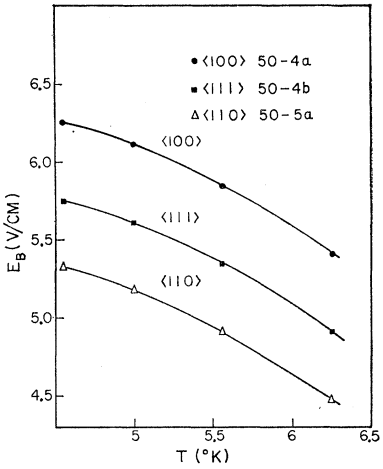


FIG. 22. Variation of the breakdown field with temperature for three samples of different orientation cut from the same single crystal.

(ii) Experimental Results and Discussion

Figure 22 shows the variation of the breakdown field with temperature for three samples cut from the same crystal with the same ultrasonic die. The low-temperature Ohmic mobility and resistivity are shown in Table III. The crystal was zone leveled, 10 Ω cm and Sb doped, corresponding to $\sim 3\times 10^{14}$ donors per cm^3; the relatively high doping level was chosen to assure a homogeneous distribution of $N_D - N_A$ throughout the crystal. That this was the case is clear from the data in Table III.

For a ⟨110⟩ oriented field, the valleys divide in pairs;

[65] The possibility of using this procedure was appreciated as a result of a suggestion made by P. J. Price in another connection.

TABLE III. Pertinent data for samples cut from crystal No. 50 used for anisotropy studies.

Sample	Orientation	(N_D-N_A) (10^{14} cm^{-3})	ρ(77°K) (Ω-cm)	μ_H(77°K) (cm^2/V sec)	ρ(5°K) (10^6 Ω-cm)	μ_H(5°K) (10^5 cm^2/V sec)	E_B (V/cm)
50–4a	⟨100⟩	1.66	1.14	31 500	3.35	2.72	6.11
50–4b	⟨111⟩	1.62	1.15	32 000	3.27	2.77	5.61
50–5a	⟨110⟩	1.72	1.12	30 600	2.90	2.96	5.18

two hot and two cold. The hot ones have the easy mobility direction along the field so that one would expect the breakdown field to be minimum. For a ⟨111⟩ oriented field, three valleys are equally hot and the forth (with its major axis or hard direction along ⟨111⟩) is "very cold." The relation among the breakdown fields for the three samples is readily shown from Eq. (41) to be

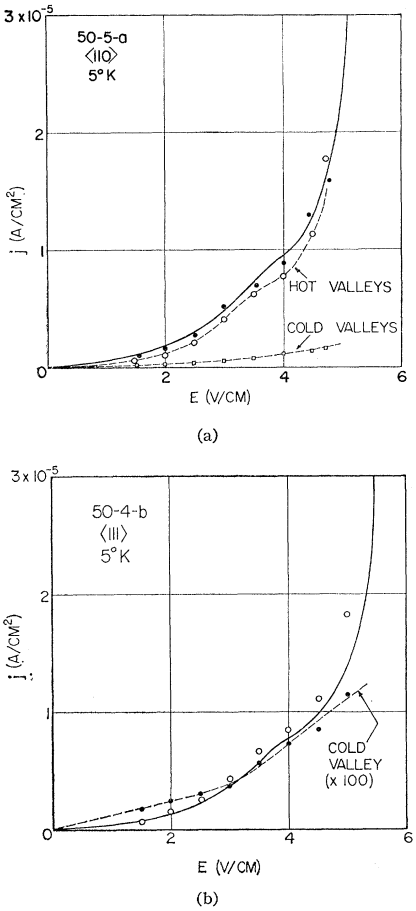


FIG. 23. The solid curve in each plot shows the variation of current density with electric field for the respective samples. (a) The dotted curves show the current density in each of the two classes of valleys, computed as described in the text from the solid curve, assuming *no* repopulation of the valleys. The black circles are the sum of the contributions. (b) The open circles are the sum of the contribution from both classes of valleys, again assuming no repopulation. For this ⟨111⟩ case, however, the "cold" valley contributes very little of the total current. For both samples, there appears to be *no* repopulation from "hot" to "cold" valleys.

$$KE_{110}^2=E_{111}^2(1+8K)/9=E_{100}^2(1+2K)/3, \quad (44)$$
$$K=\mu_\perp/\mu_{||}.$$

Similar results for the energy gain vs orientation of the field are to be found in a paper by Gold.[66]

The appropriate value of K to use in Eq. (44) is not known, but it can be roughly estimated. However, so long as K is large, Eq. (44) is quite insensitive to its exact value. For pure lattice scattering,[64] $K\simeq 17$. The "forward-scattering effect" discussed earlier considerably reduces this value. On the other hand, neutral impurity scattering, which is roughly as frequent at breakdown for these rather impure samples as is lattice scattering, is isotropic,[67] corresponding to $K=m_\perp/m_{||}=19.7$. A reasonable value for K for the three samples might then be ~ 10. Values for the ratio of breakdown field, computed using Eq. (44) for $K=10$ and $K=\infty$, are compared with the data from Fig. 22 in Table IV. The excellent agreement would indicate that impact ionization does produce carriers in the same valley as the colliding electron. This is somewhat surprising since the bound electron is in a ground state that is a linear combination of wave functions from each band edge minimum. It may be that the matrix element for impact ionization may be mainly between that part of the bound electron wave function that belongs to the same valley as that of the incoming electron. As a result, the impact-ionization cross section obtained from a simple hydrogenic model would be lowered by a factor $\sim\frac{1}{4}$; this may, in fact, be part of the reason that the experimental value ($\sim 2\times 10^{-2}\sigma_0$, see Sec. IV C) is so much less than the analogous quantity for atomic hydrogen[68] ($\sim\sigma_0$).

In Fig. 23(a) the measured current-density–electric-field variation for the ⟨110⟩ sample is plotted, as well as

TABLE IV. Ratios of the breakdown field for differently oriented samples cut from the same crystal. The experimental values are compared with a theoretical value computed for two values of the usual anisotropy parameter $K=\mu_\perp/\mu_{||}$.

	$K=10$	$K=\infty$	Experimental	T
$E_{110}:E_{111}:E_{100}$	1:1.06:1.20	1:1.065:1.23	1:1.08:1.18	5°K
			1:1.10:1.19	6.25°K

[66] L. Gold, Phys. Rev. **104**, 1580 (1956).
[67] R. A. Laff and H. Y. Fan, Phys. Rev. **112**, 317 (1958).
[68] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Oxford University Press, New York, 1949), 2nd ed., p. 245.

the separate contributions to j_T of the hot and cold valleys, computed as per Eqs. (42), (43) and the associated discussion, *assuming* $f=0$ and $K=\infty$. Values of j_T are shown for representative values of E. [For $K=10$, the results are altered from those shown by an amount too small to be resolved on the scale of Fig. 23(a).] The relative power input to the hot and cold classes of valleys is $[(K+2)/3K]\simeq\frac{1}{3}$, so that the electron temperatures (for $E\gtrsim 1$ V/cm) are in the ratio $\sim 2:1$ [$T_E\propto(\text{power})^{\frac{2}{3}}$; cf. Eqs. (7)–(9)]. Yet given the temperature differential between the pairs of valleys and the associated carrier density difference, *no repopulation* has occurred, except possibly very near breakdown. Here, however, any uncertainty in the entire procedure is magnified since $j_{100}{}^H$ and $j_{100}{}^C$ vary significantly for a small change in $E_{100}{}^{H,C}$.

Similar results are shown in Fig. 23(b) for the ⟨111⟩ sample. The contribution to j_T of the cold valley is, however, negligible; it is shown here magnified by a factor 100. Again, for an even greater electron "temperature" ratio between the two classes of valley ($\sim 4:1$) and a greater carrier density differential than for the ⟨110⟩ sample, there appears to be *no repopulation*, except possibly very close to breakdown. These results are in contrast to the situation at ~ 80°K, where large repopulations occur.[69]

The absence of repopulation, except possibly near breakdown, which at first would appear to be quite unexpected, is, in fact, readily explicable if the main mechanism for intervalley scattering is the giant trap, ionized impurity mechanism discussed by Weinreich[59] (cf. Sec. IV A, ii, above). There are three types of intervalley (I–V) scattering that may occur; that due to phonon interactions, to ionized impurity scattering and to neutral impurity scattering.[70] Phonon scattering by longitudinal phonons requires a phonon with energy ~ 300°K; there can, of course, be no induced intervalley scattering by such phonons at 5°K. Spontaneous emission would require the hot electron to have at least the phonon energy; the hot electrons for the most part are not that hot. Intervalley scattering by transverse phonons (~ 100°K) is strictly forbidden by symmetry, between points at the centers of the zone boundary. The selection rule is somewhat weakened because the hot electrons, when in the heavy mass direction, are $\sim 5\%$ in towards the zone center. However, spontaneous emission would be the only scattering mode; the carriers that scattered would lose essentially all their energy so that the density of final states would be very small. The total intervalley scattering rate by phonons is then expected to be quite small.

Since ionized impurity (I–V) scattering increase very rapidly as the temperature is decreased, while neutral impurity I–V scattering is approximately independent of T, ionized I–V scattering should dominate.[59] We can write for, R_{ij}, the rate of I–V scattering from valley i to j.

$$R_{ij}=n_iB_{ij}N_A, \quad n_i\ll N_A. \tag{45}$$

If the I–V mechanism is a capture in an excited impurity orbit and subsequent reemission in another valley before cascading to the ground state occurs, then

$$B_{ij}\propto B_T{}^{(i)}, \tag{46}$$

and from Eq. (15)

$$R_{ij}\propto B_T{}^{(i)}/[B_T{}^{(i)}N_A-A_I{}^{(i)}(N_D-N_A)]. \tag{47}$$

So long as $A_I{}^{(i)}(N_D-N_A)\ll B_T{}^{(i)}N_A$, $R_{ij}=$ constant (for fixed T) independent of the pair of valleys considered and the carrier density and electron distribution in each. Thus, *no repopulation should occur* ($f=0$) until impact ionization becomes significant, at which point from Eq. (47), $R_{HC}>R_{CH}$. This is consistent with the data near breakdown in Figs. 23(a), (b) and indicates, as conjectured earlier (Sec. IV C ii), that A_I only becomes significant quite near breakdown at the lower-lattice temperatures.

D. The "Bump"

At fields ~ 3–4 V/cm, for the more pure samples, there is a "bump" in the current density-electric field characteristics. It is most clearly seen when the characteristics are plotted on a linear scale [cf. Figs. 23(a), (b)]. Several experimental facts are significant: The bump occurs for a range of temperatures such that the current at the bump, and, therefore, the carrier density, varies by many orders of magnitude; the bump occurs for all orientations measured in As, Sb, and p-type[57] germanium. These facts suggest that the bump arises from a mobility effect; it would appear that beyond ~ 3–4 V/cm the electron (or hole) distribution no longer heats at the same rate with increasing electric field, but rather more slowly. This would of course slow down the rate of increase of carrier concentration and produce the "bump" observed.

Presumably the field has been reached that corresp .ds to the point where in the simple isotropic theory equipartition for the phonons involved in a scattering no longer holds, i.e., where the distribution function goes from the regime where[3] $\mu\propto E^{-\frac{1}{2}}$ to that where $\mu\propto E^{-0.8}$. Of course, for the low temperature, anisotropic case the situation is more complex. The "forward-scattering effect" is really a stepwise transition between these two regions. Experimentally, the mobility does begin to decrease more rapidly with E at the bump, so that the rate of power input to the distribution no longer increases with E as rapidly. It remains for a proper solution of the Boltzmann equation, that includes anisotropy and is valid for high E and low T, to be made to put these ideas on a firmer basis.

---

[69] W. Sasaki, M. Shibuya, K. Mizuguchi, and G. M. Hatoyama, J. Phys. Chem. Solids **8**, 250 (1959); S. H. Koenig, Proc. Phys. Soc. (London) **83**, 959 (1959).

[70] These, and the following points, are discussed in detail by Weinreich *et al.*, reference 59.

E. Remarks on the Activation Energy of P, As, and Sb as Donors

In order to grow germanium crystals with, say, 10^{14} donors/cm^3, the doping must be done in several steps. First a "master alloy" must be made, a small amount of which is used to dope "pure" germanium to make a secondary alloy with perhaps 10^{16}–10^{17} donors/cm^3. This new, dilute alloy is further used for producing still more dilute samples. If the desire is for very lightly Sb-doped germanium, this iterative procedure must be carefully controlled. The distribution coefficient of Sb is an order of magnitude less than that of As, and two orders of magnitude less than that of P[26], so that each successive recrystallization will concentrate (relative to the Sb) any stray P or As that may have been in the original Sb, or that may have come from the crystal growing apparatus if it were once contaminated with (highly volatile) phosphorus or arsenic. This is presumably how samples 45–2 and 45–10, nominally antimony doped, came to contain arsenic. (P is used rather infrequently in doping germanium so that it is a less likely accidental contaminant. We therefore presume As to be the contaminant. The properties of P as a dopant are, however, sufficiently close to those of As so as not to alter the discussions in the earlier parts of the paper if P rather than As were present.)

By the above reasoning, one would expect an As- or P-doped sample to be relatively free of Sb. However, since ϵ_1 for Sb is only $\sim$70% of that for As or P, the thermal ionization rate for Sb at 4°K is $\sim 10^4$ greater than for the other donors. One might then at first expect that a small admixture of Sb could significantly affect the measured activation energy of an As- or P-doped sample. That this is *not* the case is shown in Appendix A. There it is shown that if $N_{Sb} > N_A$, the sample will behave below $\sim$7°K as though it were Sb doped; if $N_{Sb} < N_A$, the activation energy measured will be that appropriate to As or P.

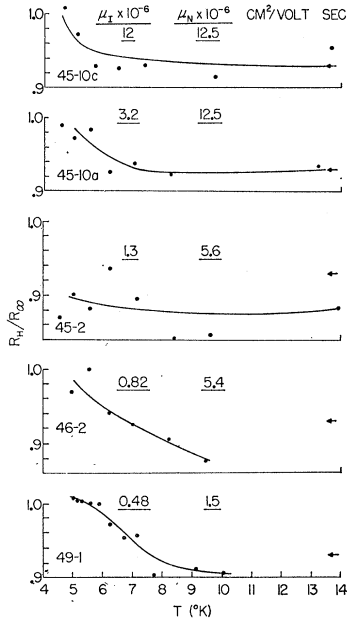


FIG. 24. Temperature variation of (R_H/R_∞), the ratio of the Hall constant measured in the linear Hall effect range to that at "infinite" magnetic field for several samples of differing purity. "Infinite" field is not so large as to affect the donor binding energy. The arrows at the right indicate the value for pure acoustic phonon scattering, which would be approached by all the samples listed at $\sim$80°K.

The "acknowledged" thermal activation energies for P-, As-, and Sb-doped samples are those reported by Geballe and Morin (GM)[23]. These (early) data were interpreted before the ground-state splitting of the donor states was appreciated, so that, in principle, the data should be corrected. Moreover the assumption of a temperature-independent Hall to drift mobility ratio was made, though GM gave arguments to show this assumption to be reasonable (cf. following section). In Table I we report the value of ϵ_1 for an As- and a P-doped sample; these agree with those of GM within 1%, as one would expect since the singlet-triplet splitting for As and P donors[27] is $\sim$40°K, well above the temperature range where most data for ϵ_1 was obtained. The value for activation energy reported by GM for Sb donors, however, should be a measure more of ϵ_2 rather than ϵ_1. This is so since the singlet-triplet splitting is $\sim$5°K; over their temperature range, considering the larger statistical weight of the triplet state, the donors would be for the most part activated from the triplet level. Their result of 9.7×10^{-3} eV, so considered, is also in excellent agreement with those reported here.

There would then appear to be little uncertainty in the thermal activation energy of P, As, and Sb donors. However the optically measured values for As and P donors are different[25] from the thermal values and the difference is real. For As the optical value is greater by 10%, for P, by 7% and for Sb, essentially no difference. The Franck-Condon effect must be operative, the more so the deeper the penetration of the singlet ground-state wave function into the central impurity cell. The magnitude, experimentally, is, in fact, linearly proportional to the singlet-triplet splitting. Qualitatively this is as it ought to be. The ground-state splitting is a measure of the change of the polarization energy of the lattice when the donor electron is sufficiently close to the impurity so that the lattice dielectric constant no longer screens the Coulomb field. But it is just this energy that contributes to the difference between the optical and thermal "activation" energy, so long as the optical absorption occurs in a time comparable to, or shorter than, the dielectric relaxation time. The absorption time τ_A is of the order of the time taken for a photon to transverse the ground-state orbit: $\tau_A \sim 60 \times 10^{-8}$ cm$/7 \times 10^9$ cm sec$^{-1} \sim 10^{-16}$ sec. This frequency, in the ultraviolet and corresponding to an energy $\sim$4 eV, is sufficiently high so that dielectric relaxation

effects should be present. (The dielectric constant of germanium is, of course, entirely electronic.)

F. R_H/R_∞

The ratio of R_H, the Hall constant measured in the linear Hall effect range, to R_∞, the Hall constant measured for $(\mu H/c)\gg 1$, is a number usually of the order of unity, but whose exact value depends on the mass and scattering anisotropy as well as on the energy dependence of the scattering time.[71,72] For spherical energy surfaces and idealized acoustic phonon scattering the ratio is[14] $3\pi/8$; for Coulomb scattering, 1.93; for neutral impurity scattering, 1.0. If the mass anisotropy of germanium is included, but the scattering time assumed isotropic, these three numbers get multiplied by a factor 0.78 (Herring's B).[71] For the more complicated case of anisotropic scattering, so long as the energy dependence is the same for all directions, the B factor is altered in a simple manner.[72] For mixed scattering though, no tabulated computations of R_H/R_∞ that take into account the anisotropy exist. However, there exist experimental results for mixed ionic and acoustic scattering[73]; they are quite similar to what one would compute assuming isotropic scattering.[74] The significant point is that the number that would be either $3\pi/8$ or 1.93 at either extreme goes through a broad minimum $\sim$1.07 when the ionic and phonon contributions to the mobility are about equal.

Figure 24 shows the variation of R_H/R_∞ as a function of temperature for several samples. The approximate relative contribution of phonon, ionic and neutral impurity scattering at 5°K is also given. The general behavior can be understood, for the most part. For the purest samples (the two upper curves), R_H/R_∞ above $\sim$7°K has the value appropriate to acoustic phonon scattering (0.93), (cf. Fig. 2). At lower temperatures the curves turn upward. This is somewhat surprising as one would expect a small admixture of ionic scattering to do the converse. We conjecture that perhaps the "forward scattering effect" (it was suggested earlier that this might affect the Ohmic mobility at the lowest temperature) is reducing the mobility anisotropy and thereby increasing R_H/R_∞. Sample 45–2, which in the range 4–20°K has at least as much ionic scattering as phonon scattering, has a value of R_H/R_∞ lower than the acoustic phonon value by $\sim$7%, consistent with the data of Moss and Walton[73] taken at higher temperatures. The two lower curves are for samples which at 5°K have ionic scattering rates several times the acoustic rate. It would then be expected that R_H/R_∞ would be high at 5°K, higher for 49–1 than for 46–2.

As T is increased, R_H/R_∞ should decrease, going through a minima which is <0.93 (the phonon value), and finally return to the phonon value. This behavior is what is observed.

ACKNOWLEDGMENTS

We first must thank David Hartmann of the General Electric Development Laboratories, Syracuse, New York, for having grown crystal 45 for us. Without this crystal, the bulk of the measurements, reported here, would not have been possible. We thank Ted Geballe at Bell Laboratories, Murray Hill, for supplying the bulk of the other crystals used in the experiments reported here; G. Ascarelli for the loan of a crystal; John J. Hall for making the strain measurements on sample 45.

Over the years we have had many useful discussions with many people; particularly valuable were discussions with P. J. Price and M. Lax.

APPENDIX A

Consider the case of a semiconductor at low temperature that contains N_a donors with a set of energy levels $\mathcal{E}_{ai}$ and N_b donors (of a different type) with a set of levels $\mathcal{E}_{bj}$. We have in mind that N_a represents the Sb concentration and N_b the As or P concentration in an n-Ge sample, so that $\mathcal{E}_{a1}, \mathcal{E}_{b1}>0$ ($i, j=1$ represents the respective donor ground states). The fraction of neutral impurities of either type is then[7]

$$\begin{aligned}(N_a-N_a^*)/N_a&=(xZ+1)^{-1},\\ (N_b-N_b^*)/N_b&=(yZ+1)^{-1},\end{aligned}\tag{A1}$$

where

$$\begin{aligned}Z&=\exp(-\epsilon_F/kT),\\ x^{-1}&=\textstyle\sum_i g_{ai}\exp(-\mathcal{E}_{ai}/kT)\\ &=[\exp(-\mathcal{E}_{a2}/kT)]\{2[\exp(\delta\epsilon_a/kT)+3]\\ &\qquad+\textstyle\sum_{i>2} g_{ai}\exp[(-\mathcal{E}_{ai}+\mathcal{E}_{a2})/kT]\},\\ y^{-1}&=\textstyle\sum_j g_{bj}\exp(-\mathcal{E}_{bj}/kT)\\ &=[\exp(-\mathcal{E}_{a2}/kT)]\textstyle\sum_j g_{bj}\exp[(-\mathcal{E}_{bj}+\mathcal{E}_{a2})/kT].\end{aligned}\tag{A2}$$

Here ϵ_F is the Fermi energy, N_a^* and N_b^* the respective density of ionized donors of each type, and $\delta\epsilon_a$ the (positive) singlet-triplet splitting of the type a donor. In writing the second equality for x^{-1}, the fact that the conduction band minima are at the Brillouin zone boundary in germanium was used. The set of equations (A1) may be solved for the Fermi energy in terms of known or measurable parameters, if in addition the charge neutrality condition

$$n+N_A=N_a^*+N_b^*\tag{A3}$$

is invoked. One obtains readily

---

[71] See C. Herring, reference 15.
[72] C. Herring and E. Vogt, Phys. Rev. **101**, 944 (1956).
[73] T. S. Moss and A. K. Watson, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961), p. 338.
[74] V. A. Johnson and K. Lark-Horovitz, Phys. Rev. **82**, 977 (1951).

$$Z=\frac{\{-(N_b-N_A-n)-(N_a-N_A-n)(x/y)+[((N_b-N_A-n)+(N_a-N_A-n)(x/y))^2+4(N_A+n)(N-N_A-n)(x/y)]^{1/2}\}}{2x(N-N_A-n)}, \quad \text{(A4)}$$

where $N=N_a+N_b$.

Since the Fermi level is always well below the conduction band edge [this information may be extracted from Eq. (A4)], one may write for the conduction band carrier density

$$n=(2v/h^3)(2\pi m^*kT)^{3/2}\exp[-(\epsilon_C-\epsilon_F)/kT]$$
$$=N_C\exp(-\epsilon_2/kT)\exp[(\epsilon_F-\mathcal{E}_{a2})/kT]; \quad \text{(A5)}$$
$$\epsilon_2=\epsilon_C-\mathcal{E}_{a2}; \quad N_C=(2v/h^3)(2\pi m^*kT)^{3/2}.$$

Combining Eqs. (A2) and (A5) yields

$$(v/h)^3(2\pi m^*k)^{3/2}\exp(-\epsilon_2/kT)=nZxT^{-3/2}\{\exp(\delta\epsilon_a/kT)$$
$$+3+\sum_{i>2}(g_{ai}/2)\exp[(-\mathcal{E}_{ai}+\mathcal{E}_{a2})/kT]\}=\chi_2, \quad \text{(A6)}$$

which is the quantity discussed in Sec. II C and plotted in Fig. 5. The product Zx is readily computed using Eq. (A4). To obtain (x/y) accurately requires accurate knowledge mainly of the single-triplet splitting for Sb and As or P donors and the activation energy for arsenic or phosphorus donors. For the summations over the higher excited donor states the theoretical values computed in the effective mass approximation are entirely adequate.

We consider now the case where As or P is the majority dopant and ask whether the presence of a small amount of Sb can affect the activation energy measured in the region of 4°K. At this temperature

$$(x/y)\simeq\exp[(-\mathcal{E}_{b1}+\mathcal{E}_{a1})/kT]\simeq 10^4, \text{ and } n\ll N_A. \quad \text{(A7)}$$

One must then distinguish three cases for Eq. (A4). Either $N_a=N_A$, or it is greater or less. If $N_a=N_A$, then since (x/y) is so large, realistically only the terms in (x/y), Eq. (A4) need be considered. One obtains for $N_a>N_A$:

$$Z=N_A/x(N_a-N_A),$$
$$n=[N_C(N_a-N_A)/N_A]\exp(-\epsilon_a/kT), \quad \text{(A8)}$$
$$\epsilon_a=\epsilon_C-\mathcal{E}_{a1},$$

for $N_a<N_A$:

$$Z=(N_A-N_a)/y(N-N_A),$$
$$n=[N_C(N-N_A)/(N_A-N_a)]\exp(-\epsilon_b/kT),$$
$$\epsilon_b=\epsilon_C-\mathcal{E}_{b1},$$

for $N_a=N_A$:

$$Z=(N_bN_A/xy)^{1/2}/N_b,$$
$$n=[N_C(N_b/N_A)^{1/2}]\exp(-\epsilon/kT),$$
$$\epsilon=(\epsilon_a+\epsilon_b)/2.$$

These results are what one would expect. If the Sb donors are not completely compensated (i.e., $N_a>N_A$), both n and the Fermi level are independent of the amount of As or P present. If the Sb donors are exactly compensated (an idealized situation essentially impossible to achieve in practice, unless possibly by controlled radiation damage), the activation energy is the mean of the activation energies for the empty donors and the full ones. If some of the As or P donors are also compensated (i.e., $N_a<N_A$), the Fermi level is locked near the As or P ground state, and the activation energy that is measured is that appropriate to As or P. The value for the acceptor concentration obtained by using Eq. 2 to reduce the data, if it were incorrectly assumed that no Sb were present ($N_a=0$), would however be too *small* by exactly the amount N_a.

APPENDIX B

The drift mobility was determined by applying a rapidly rising (<1 nsec) voltage pulse E to the Ge sample and determining the initial change in current J_0. This current change is proportional to the initial carrier concentration n_0 (since the number of carriers cannot change appreciably in a time $\approx$1 nsec), the electric field E and the drift mobility μ appropriate to that field (since the distribution function can change in a time $\lesssim$1 nsec). The drift mobility is given by $\mu=J_0/n_0eE$; one need only vary E and determine J_0 to obtain the variation of drift mobility with electric field. The experimental difficulties involve (1) the sensitivity, noise, and rise time of the equipment, and (2) the capacitive feedthrough of the Ge sample and apparatus.

The experimental apparatus used was essentially that shown in Fig. 12; no bias was used, and the integrator, amplifier, and oscilloscope were replaced with a sampling unit and XY plotter. The voltage pulses (10-nsec duration, 200 per sec with a rise time less than 0.5 nsec) were produced by discharging a coaxial cable through a mercury relay.[56] The current waveform was viewed with a Tektronix type N sampling unit, appropriately altered to drive an XY recorder; the Y input to the recorder was the dc output of a boxcar-integrator, driven by the unblanking pulse, which samples the output of the N unit. The X input of the recorder was driven by a linear 10-sec sawtooth which also swept the N unit. The frequency response of the system was measured up to 2 kMc/sec and is approximately that of an ideal sampling system that uses a sinusoidal sampling pulse. This is also consistent with the measured response to a rectangular pulse.

The time variation of current is shown for a typical

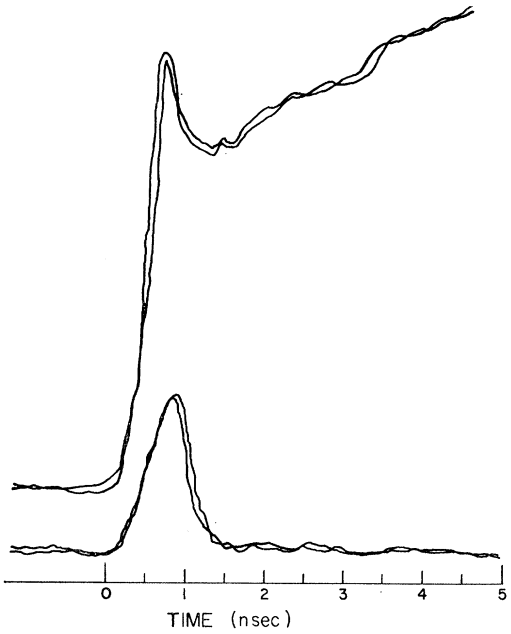


FIG. 25. The upper curve shows the variation of current as a function of time for sample 45–10 at 8.3°K after application of a 33 V/cm pulse. The initial change of current (~0.5 nsec) involves only the change in the distribution function due to the field. The subsequent slower rise is the onset of the increase of the carrier concentration, describable by the rate equation discussed throughout the text. The overshoot after the initial rise is due to capacitative feedthrough of the sample. This feedthrough is shown in the lower trace, which was obtained by reducing the temperature so that the current rise for the first several nsecs would be negligible on the scale shown.

situation ($E=33$ V/cm, $T=8.33$°K) in the upper curve of Fig. 25. The lower curve shows the capacitive feedthrough, obtained by lowering the temperature to 4.2°K, thereby reducing n_0 and, hence, the conduction current by a factor of $\sim 10^{-5}$. If the capacitive feedthrough is subtracted from the upper curve (assuming superposition applies), the conduction current is obtained, Fig. 26. This current is seen to rise to a value J_0 within the rise time of the system, and then increase more slowly from that point as the carrier concentration increases. J_0 is then obtained by extrapolating this slowly rising curve backward to a time t_0. (If a ramp or slowly rising exponential is viewed by this sampling system, the response for times greater than a rise time will be directly proportional to the input but delayed in time by $t_0=0.86$ of a rise time or 0.50 nsec.) This shift in time was significant in correctly determining J_0, especially at high temperatures and high fields, where the carrier concentration increases rapidly after the pulse.

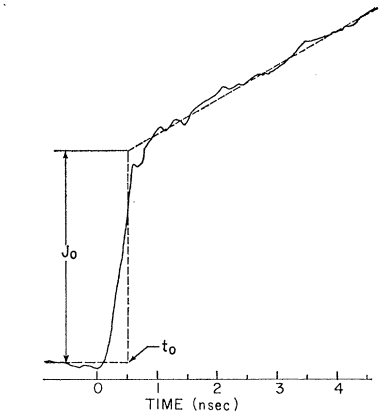


FIG. 26. The difference of the two curves of Fig. 25 is shown here as well as the procedure used for extrapolating backward in time to obtain the correct value for J_0, the initial increase in current. The ramp must be extrapolated back to a time $t_0=0.5$ nsec, which is the time lag that this sampling system has to a ramp voltage applied at $t=0$.

The temperature range over which measurements can be made is limited at low temperatures by the signal intensity which decreases exponentially with temperature (as n_0), ultimately to be lost in the temperature-independent capacitive feedthrough. At high temperatures, the limitation is the decreasing sample impedance and the decreasing time constant for the increase of carrier concentration.

The drift mobility is normalized in the Ohmic region to that measured by dc techniques in the "infinite" magnetic field limit.

APPENDIX C

We want to solve for the time variation of n in the Ascarelli and Brown experiment,[4] and from that find the intercept at $t=0$ of the extrapolation of the behavior at long times. Let n_i be the (large) carrier density at $t=0$ and n_T the thermal equilibrium value that is approached as $t\rightarrow\infty$. Then from Eq. (25),

$$\begin{aligned} dN/dt &= a-b(N+n_T)-c(N+n_T)^2-d(N+n_T)^3 \\ &= -N(b+2n_Tc+3n_T^2d) \\ &\quad -N^2(c+3n_Td)-dN^3, \end{aligned} \tag{C1}$$

since

$$a-bn_T-cn_T^2-dn_T^3=0.$$

Here $N=n-n_T$. Throughout the decay $N\gg n_T$; $n_T \sim 10^5/\text{cm}^3$ is in fact too small to be observed both in the pulse experiments of AB and those reported here. The coefficients b, c, and d have their thermal equilibrium value during the decay; with their definitions, Eq. (25), and the conditions of thermal equilibrium, we may write

$$\begin{aligned} b &\simeq B_TN_A, \\ c &\simeq B_IN_A+B_T, \\ d &= B_I. \end{aligned} \tag{C2}$$

Since $n_T\ll N_A$, Eq. (C1) may finally be rewritten as

$$\begin{aligned} dN/dt &= -B_TN_AN-(B_IN_A+B_T)N^2-B_IN^3 \\ &= -B_TN_AN(1+B_IN/B_T)(1+N/N_A). \end{aligned} \tag{C3}$$

The solution of this equation, obtained by the method

of partial fractions, is

$$\frac{N}{N_i}\left[\frac{1+(B_IN)/B_T}{1+(B_IN_i)/B_T}\right]^{-(B_IN_A)/(B_IN_A-B_T)} \times\left[\frac{1+(N/N_A)}{1+(N_i/N_A)}\right]^{B_T/(B_IN_A-B_T)}=\exp(-B_TN_At), \quad \text{(C4)}$$

where $N_i=n_i-n_T$. For $t\rightarrow\infty$, $N\rightarrow 0$, and an exponential law is obtained, which when extrapolated back to $t=0$ gives for the intercept with the ordinate:

$$N_i\left[1+\frac{B_IN_i}{B_T}\right]^{-B_IN_A/(B_IN_A-B_T)} \times\left[1+\frac{N_i}{N_A}\right]^{B_T/(B_IN_A-B_T)}. \quad \text{(C5)}$$

If n_T is ignored with respect to N_i, expression (C5) becomes Eq. (20) in the text.

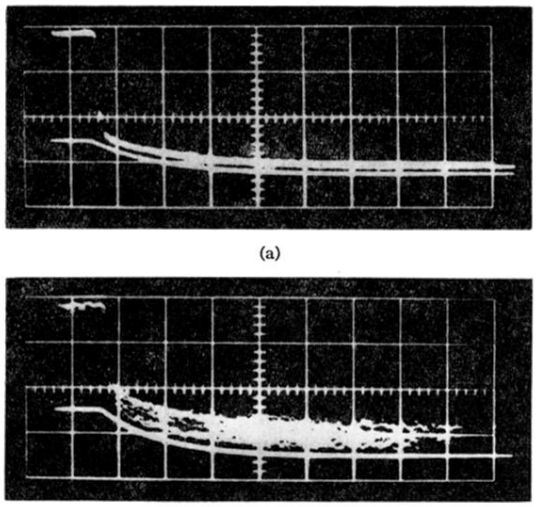
APPENDIX D

After the bulk of the present paper had been written, work by Ascarelli and Rodiriguez[75,76] (AR) appeared in print, the main point of which was to compute in a manner somewhat different from Lax,[51] a value for B_T for thermal electrons. The basic physics, as best we can tell, is qualitatively the same; the numerical results for thermal electrons are similar. These are included in Fig. 16.

As AR compute B_T for thermal electrons only, they are able to make liberal use of the principle of detailed balance to considerably simplify the computations. However, as this precludes obtaining results as an explicit function of electron energy, further comparisons with our experimental data cannot be made.

Ascarelli and Rodriguez also reiterate[76] an incorrect remark made earlier[4] concerning results reported by Koenig and Brown[77] on the relative photosensitivity of As- and Sb-doped germanium to optical radiation emitted when electrons in a hot distribution ($\sim$100°K) recombine with ionized As or Sb donors. The typical photon emitted will in this case have an energy $\sim$ twice the As or Sb binding energy and will, therefore, ionize an As or Sb donor with essentially equal probability. It should, then, be obvious that the efficiency of detection, i.e., the fractional change in conductivity per incident photon, is given simply by the ratio of the rate of carrier generation by photons to the thermal ionization rate. Since the latter is much less at 4°K for As donors than for Sb donors, the photoefficiency for As-doped Ge is the greater. However AR and AB state that the reverse should be the case, apparently implicitly assuming,[78] for some reason, that the electrons are not hot when recombining.

[75] G. Ascarelli and S. Rodriguez, J. Phys. Chem. Soc. **22**, 57 (1961).
[76] G. Ascarelli and S. Rodriguez, Phys. Rev. **124**, 1321 (1961).
[77] S. H. Koenig and R. D. Brown, III, Phys. Rev. Letters **4**, 170 (1960).
[78] S. Rodriguez (private communication).



(a)

(b)

FIG. 13. The upper pair of traces shows: (1) the time variation of the current through sample 45–10*a* when the voltage across the sample is reduced, by means of a fast falling pulse, from the breakdown value to a value of $\sim\frac{1}{4}$ this amount; and (2) the voltage pulse integrated by the variable time constant *RC* network. By superimposing both curves, using a greater gain setting, the time constant for decay of the current after the initial decrease can be readily measured. These data were taken at 6.35°K. The lower pair of curves shows similar data taken at 5.03°K. Here the measurements are obscured by noise, presumably associated with intrinsic thermal instabilities of the sample (cf. footnote 44). These curves illustrate the worst conditions under which data must at times be taken. The time scale for both sets of curves is 200 nsec/cm; the bias settings, 60 μA. The situation in the lower traces can be somewhat improved by using a larger value of bias.