PHYSICAL REVIEW VOLUME 98, NUMBER 4 MAY 15, 1955

Seebeck Effect in Silicon

T. H. Geballe and G. W. Hull
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received February 3, 1955)

The Seebeck effect has been measured from liquid hydrogen temperatures into the intrinsic range for a series of single-crystal silicon samples in which varying concentrations of donor and acceptor atoms have been incorporated. Large values of Seebeck voltage believed to be caused by the type of phonon-electron coupling previously postulated as occurring in germanium have been found. This effect is found to be dependent upon charge carrier concentration, and upon sample dimension below 100°K. A low-temperature reversal of the sign of the Seebeck voltage is observed for large carrier concentrations. It is understandable in terms of impurity band conduction. The behavior above room temperature into the intrinsic range is found to be consistent with electrical conductivity data.

INTRODUCTION

THE previously reported investigations of the Seebeck effect in germanium[1,2] have proved to be understandable in terms of the band structure of germanium and have also indicated an interesting type of interaction between mobile charge carriers and phonons. We have therefore been led to extend this type of study to silicon, where considerable knowledge of the band structure has recently been obtained from cyclotron resonance experiments.[3] Good single crystals of silicon are now available in which a type of mobile charge carrier-phonon interaction similar to that found in germanium should be likely to exist. There have been no recent investigations reported in the literature other than on a few samples of polycrystalline material.[4,5]

EXPERIMENTAL PROCEDURE

We are indebted to Morin and Maita for making available the same set of silicon samples containing boron and arsenic used in their investigation of the electrical properties of silicon.[6] These samples were measured in the apparatus previously described.[1] The following method was adopted in order to make low-impedance ohmic electrical contacts as well as good thermal contacts between the thermocouple blocks and the sample. First a small region of high electrical conductivity was produced on the surface. This was accomplished by alloying a small amount of gold, containing a few percent antimony (for n-type) or aluminum (for p-type), to the surface of the sample directly under the thermocouple. A thin layer of silver paste was fired on to this region at 550°C for 10 minutes. This provided a surface which could be wet easily with solder but was omitted for high-resistivity samples where heat treatment might affect the carrier concentration. Occasional poor thermal contacts were noted by observing the thermal impedance across the copper-silicon junctions as described in the previous paper. In all cases, Hall and electrical conductivity measurements were made at various temperatures and were in good agreement with the published results of Morin and Maita.[6] Seebeck voltages obtained with helium or hydrogen gas in the container disagreed by several percent with the results obtained in high vacuum. This indicates a poorer thermal contact at the junctions in silicon than in germanium where no such disagreement was found. This made it necessary to make all measurements in high vacuum.

The very high thermal conductivity of the purer silicon samples in the low-temperature range made it appear likely that end effects in the flow of heat could complicate the measurement of the temperature difference between the junctions. We were led therefore to measure a few samples with a geometry designed to minimize end effects. This was accomplished by measuring some long (about 10 cm) filaments of rectangular cross section. The thermocouples were secured in the central portion away from boundaries. The heater was constructed from No. 40 manganin wire wrapped directly onto the filament. Short lengths of No. 28 copper wire were wrapped and soldered around the filament at each thermocouple-silicon junction and used to minimize temperature gradients along the thermocouple leads near the junction. The region between the thermocouples was electrically probed at room temperature and in the samples used was uniform to 20 percent. The sample was mounted in a conventional apparatus consisting of a radiation shield inside an evacuated container. The bottom of the sample was soldered to the radiation shield which was in thermal contact with the bath. The lead wires were brought to the temperature of the radiation shield and were sufficiently long to insure a negligible heat leak along the wires and thus making thermal conductivity measurements possible.[7] Lead wires for electrical conductivity measurements were attached at the filament ends and in some

---

[1] T. H. Geballe and G. W. Hull, Phys. Rev. 94, 1134 (1954).

[2] H. P. R. Frederikse, Phys. Rev. 92, 248 (1953).

[3] Dexter, Lax, Kip, and Dresselhaus, Phys. Rev. 96, 222 (1954); R. N. Dexter and B. Lax, Phys. Rev. 96, 223 (1954).

[4] A. E. Middleton and W. W. Scanlon, Phys. Rev. 92, 219 (1953).

[5] J. Savornin and F. Savornin, Compt. rend 235, 465 (1952); 236, 898 (1953); B. I. Boltaks, J. Tech. Phys. (U.S.S.R.) 20, 3 (1950).

[6] F. J. Morin and J. P. Maita, Phys. Rev. 96, 28 (1954).

[7] These results have been summarized: G. W. Hull and T. H. Geballe, Phys. Rev. 96, 846 (1954).

TABLE I. Dimensions and composition of samples.

Sample No.	Area, cm^2	Distance between thermocouples, cm	Donor concentration, cm^{-3}	Acceptor concentration, cm^{-3}
		n-type		
131 bridge[a]	0.086×0.16	1.1	2.75×10^{14} arsenic	1.0×10^{14}
537 filament	0.47×0.47	3.8	3.7×10^{14} phosphorous	9×10^{13}
566 filament	0.12×0.47	3.8	similar to 537	
565 filament	0.078×0.087	1.7	similar to 537	
130 bridge[a]	0.093×0.16	1.2	2.6×10^{15} arsenic	5×10^{14}
129 bridge[a]	0.084×0.16	1.2	2.2×10^{16} arsenic	1.5×10^{15}
139 bridge[a]	0.12×0.16	1.2	1.3×10^{17} arsenic	2.2×10^{16}
563 bridge[b]	0.12×0.16	1.1	1.1×10^{18} arsenic	1.0×10^{18} boron
126 bridge[a]	0.085×0.16	1.2	2.2×10^{18} arsenic	
140 bridge[a]	0.12×0.16	1.2	2.7×10^{19} arsenic	
		p-type		
127 bridge[a]	0.12×0.16	1.2	2.2×10^{14}	9.2×10^{14} boron
552 filament	0.50×0.50	4.3	1.9×10^{14}	1.0×10^{15} boron
559 filament	0.25×0.18	2.75	similar to 552	
117 bridge[a]	0.085×0.16	1.1	2.3×10^{15}	2.6×10^{16} boron
119 bridge[a]	0.086×0.16	1.2	4.9×10^{15}	2.0×10^{17} boron
141 bridge[a]	0.12×0.16	1.1		1.0×10^{18} boron
125 bridge[a]	0.084×0.16	1.0		1.5×10^{19} boron
		Measured in intrinsic range		
554 (b)	0.040×0.67	1.5	similar to 537	
562 filament	0.078×0.25	5.0	similar to 537	
561 filament	0.058×0.25	5.0	similar to 552	
561 (a)			filament 561, second series of measurements	

[a] Concentration data for these samples taken from reference 6.
[b] Concentration data for all samples except 563 estimated from curve fitting of Hall data as described in reference 5. Data for sample 563 estimated from growth conditions as indicated in text.

cases Hall leads were attached midway between the thermocouples. Samples measured in this apparatus are referred to as "filaments" and in the previous apparatus as "bridges." Bridges and filaments were lightly sandblasted with 180 mesh silicon carbide powder on all surfaces.

EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we shall consider the experimental data obtained for the 20 samples characterized in Table I. The data for samples measured in the extrinsic range are presented in Figs. 1 and 2. The QT product, where Q is the thermoelectric power, is plotted as a function of the absolute temperature T for *n*-type samples in Fig. 1, and for *p*-type samples in Fig. 2. It can be seen that there are certain common types of behavior and also some strikingly different types; these are discussed in detail in the following sections.

1. Dependence of Q upon Carrier Concentration near Room Temperature

It can be seen from Figs. 1 and 2 and Table I that the samples with the highest value of QT around room temperature are those with the smallest concentration difference between the donor and acceptor atoms, that is, those with the smallest density of mobile charge carriers. This is the same type of behavior observed for germanium and is to be expected for any model in which the thermoelectric energy is transported by mobile charge carriers. In fact for any two samples in Fig. 1, or any two samples in Fig. 2, other than the degenerate samples No. 125 and 140, $(Q_2-Q_1)=\log(n_1/n_2)$ near room temperature, where n is the charge carrier density. This probably means that the activity coefficients of the mobile carriers and the kinetic energy transported by them [ΔE_T of Eq. (1), reference 1] are constant over the concentration range 10^{-6} to 10^{-2} mole percent.

2. The Contribution Due to Phonon-Electron Scattering

It is evident from the marked rise in Q at low temperatures that the phonon-electron coupling mechanism invoked by Herring[8] and Frederikse[2] to explain the germanium data must also be present in silicon. Using Eqs. (4) and (5) of the previous paper,[1] assuming a density of states mass, $N_v^{\frac{2}{3}}m^*/m$, equal to unity we have calculated values of Q_p. Q_p is the contribution to Q arising from the anisotropic phonon-electron scattering. The values are plotted for the nondegenerate *n*-type samples in Fig. 3 and for the nondegenerate *p*-type samples in Fig. 4. Values for samples 127 and 130 have been omitted. Inspection of curves for those samples in Figs. 2 and 1, respectively, shows a discontinuity at low temperatures. In both cases, the discontinuity was observed upon warming from liquid hydrogen to liquid nitrogen temperatures and was accompanied by an

---

[8] C. Herring, Phys. Rev. 96, 1163 (1954).

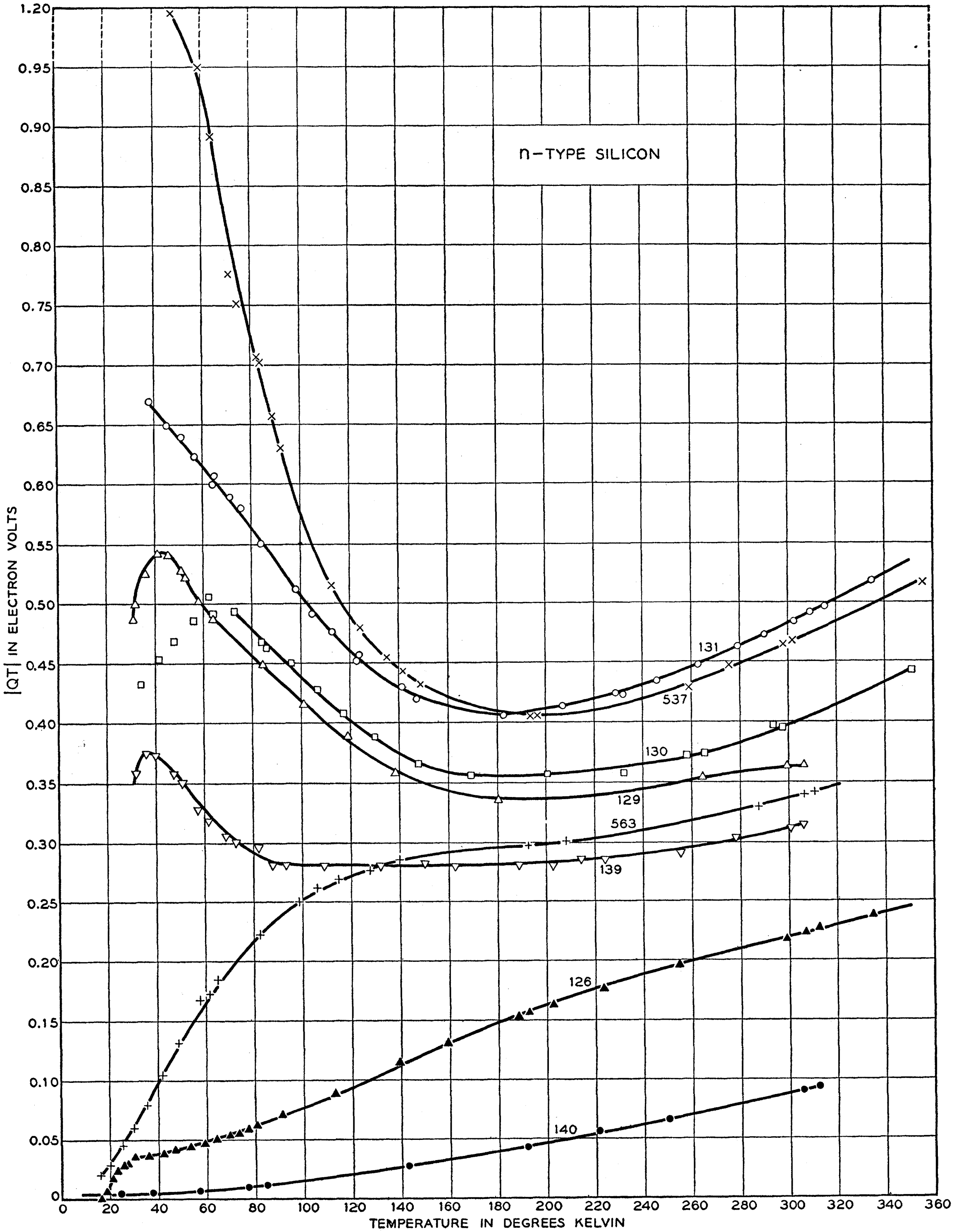

FIG. 1. Measured values of the QT product for eight n-type samples listed in Table I.

increased thermal gradient across the thermocouple-silicon junction. This was determined by the differential thermocouples shown in Fig. 1 of the previous paper. The other samples showed no such effects and therefore they are the ones considered here.

The pure n- and p-type filaments of large dimension

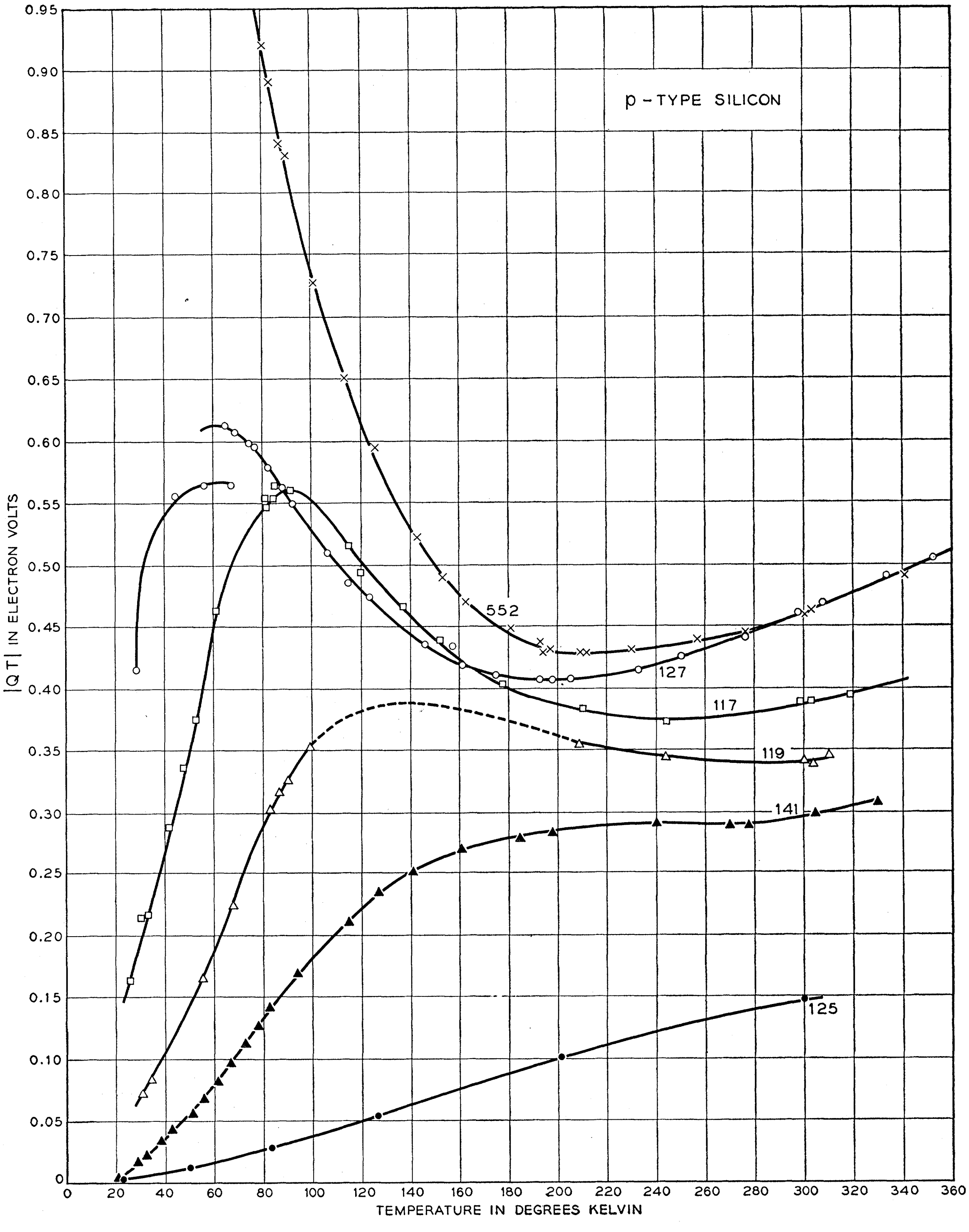

FIG. 2. Measured values of the QT product for six p-type samples listed in Table I.

(No. 537 and No. 552) have a value of Q_p proportional to $T^{-2.3}$ over a wide range of temperature. This is shown in Figs. 3 and 4. This is the general type of behavior predicted by Herring (reference 8, Sec. V) for the ideal case where the carrier concentration is small, where there is no boundary scattering and where the temperature is well below the Debye temperature. However, the magnitude of the exponent, 2.3, is much lower than the theoretical expectation for the ideal case which is from 3.5 to 4.0. The effect of the finite size of the sample and of the saturation effect, both of which are discussed below, act so as to decrease the magnitude of the exponent at low temperatures. In addition the expected behavior as the Debye temperature ($\sim$650°K) is approached is also such as to reduce the magnitude of the exponent. The combination of these effects may be influential in causing the magnitude of the observed exponent to be less than the ideal value over the whole temperature range.

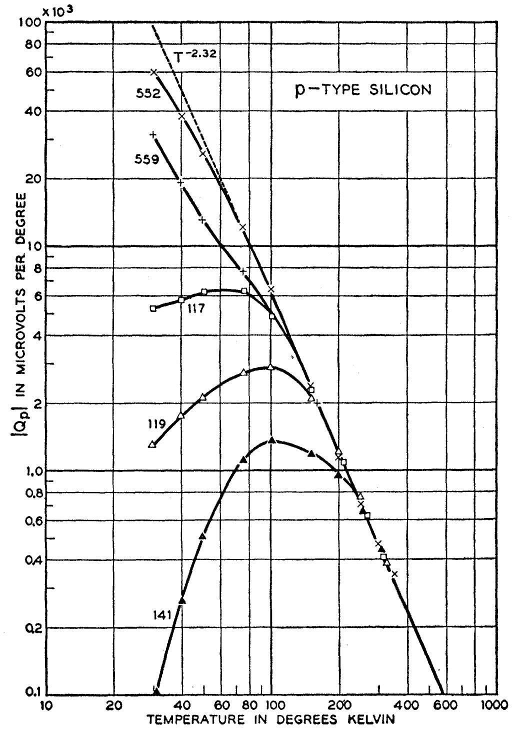

FIG. 4. Dependence of Q_p upon temperature for *p*-type samples which differ in size or acceptor concentration.

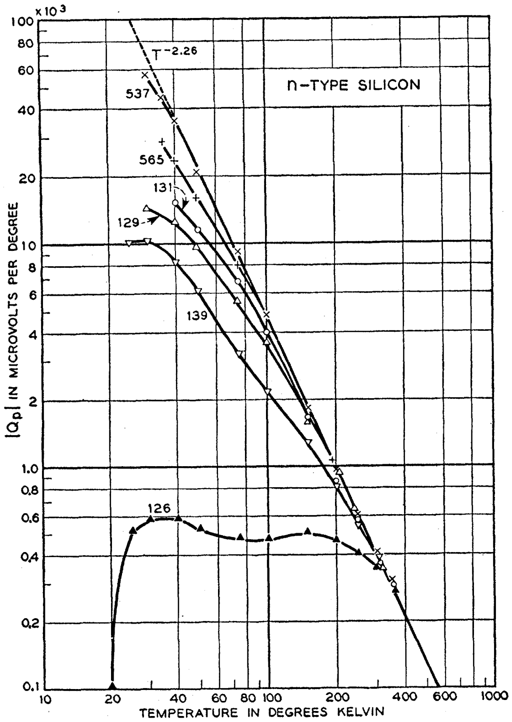

FIG. 3. Dependence of Q_p upon temperature for *n*-type samples which differ in size or donor concentration.

The saturation effect in reference 8, Sec. 7, becomes important when the concentration of carriers is sufficiently great to perturb the phonon system away from the steady state distribution it would have in the absence of the carriers and acts so as to decrease Q_p. There is evidence from thermal conductivity data that the bridge results are consistently low below 100°K where the thermal conductivity becomes high. This is probably due to nonuniform heat flow near the ends of the bridges where the thermocouple blocks are soldered. For this reason, we choose to compare bridge samples only with other bridge samples when analyzing for the saturation effects, but hesitate, for example, to infer anything by comparing bridge 131 with filament 565. The strikingly low values of Q_p for samples 126 and 141 in Figs. 3 and 4 are partly ascribed to the saturation effect. However, the rapid drop off of Q_p for these samples at the low temperature end is believed due to "impurity band" conduction and is discussed further below. The saturation effect is clearly evident in the *n*-type samples, 126, 139 and 129, even above 100°K. At lower temperatures, the carrier concentration falls off rapidly as given in Fig. 6, reference 6. This prevents the values of Q_p from saturating out more rapidly except for the case of sample 126 where the carrier concentration falls off much less rapidly with temperature. The *p*-type samples 141, 119, and 117 saturate in a qualitatively different manner and at not such high temperatures. This is not surprising. It is evident that the perturbation of the phonon system must be dependent upon the inertial mass of the charge carriers. Indeed the detailed analysis of Herring indicates that the *p*-type samples which have the larger average inertial mass[2] and thus a smaller thermal wavelength should saturate to a lesser extent for a given concentration. If the spin-orbit splitting reported in reference 2 for *p*-type samples is comparable with thermal energy the analysis for the phonon contribution will be more complicated than our simple one.

The effect of the finite size of the sample can be seen in Fig. 5. The *n*-type sample No. 537 was reduced by approximately a factor of 4 in one cross-sectional dimension to make sample 566; sample 565 was prepared from an adjacent slice with each cross-sectional dimension reduced by a factor of 6. It can be seen from the figure that this caused a reduction of 18 percent and 32 percent respectively in the measured values of Q at 55°K. In terms of Q_p the reduction is almost the same since the electronic contribution, Q_e, amounts to only

5 percent of the measured value. At 30°K, the 36-fold reduction in the cross section of sample 565 has caused a 44 percent reduction in the measured value of Q.

For the p-type sample No. 552 a reduction of the cross-sectional area by a factor of 2×2.8 caused a decrease of 10 percent in Q at 55°. The lower-temperature values shown as dashed in the figure are uncertain to about ± 10 percent because of an unaccountably low measured value of the thermal conductivity of sample 559. In all cases, the electrical conductivity of the reduced dimension samples checked that of the original sample.

The persistence of the size effect to an absolute temperature about 5 times that at which size effects are important in thermal conductivity measurements is similar to the case of germanium. Again this is evidence that the thermoelectric energy transported by the phonon system which is given by the Q_pT product resides largely in long-wavelength phonons. This is a basic assumption in Herring's theory.

3. Measurements at High Concentration of Donors and Acceptors[9]

It is evident from the mobility data of Morin and Maita, Figs. 8 and 9, reference 6, that there must be a different type of mechanism responsible for the low mobilities of samples 126 and 141. They suggest that impurity band conduction by low-mobility carriers may be occurring. We are able to offer further evidence that this indeed is the case.

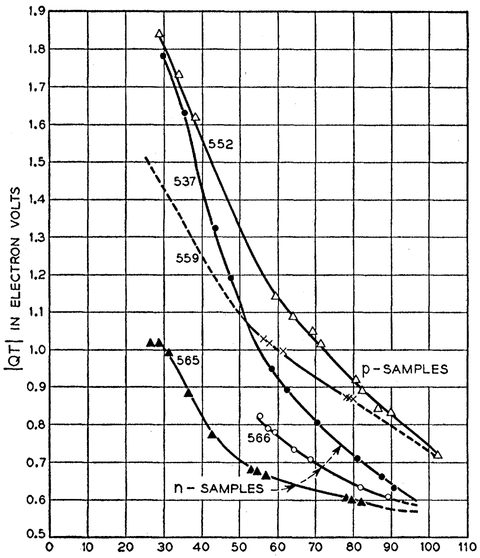

FIG. 5. Dependence of measured value of QT upon sample dimension for n-type and p-type samples. At higher temperatures the curves merge to within 1 percent of that shown for sample 537 in Fig. 1 or sample 552 in Fig. 2.

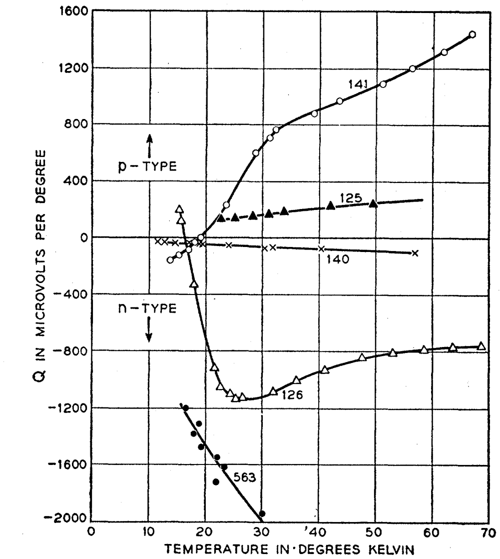

FIG. 6. Low-temperature results for samples with large donor or acceptor concentrations showing how the reversal of sign of the nondegenerate samples 126 and 141 differs from the degenerate samples 125 and 140.

By impurity band conduction is roughly meant the charge transfer between donor or acceptor atoms interacting directly.[10] In sample 126, there is approximately 1 arsenic atom per 2×10^4 silicon atoms and in sample 141 there is approximately one boron atom per 5×10^4 silicon atoms. Therefore, assuming a random distribution of donor and acceptor atoms, charge carriers must interact at the order of 10 to 20 lattice distances. The energy level associated with the isolated donor or acceptor atom is broadened into an energy band. This is the so-called impurity band. For sample 126 there are only 2×10^{15} electrons per cc in the conduction band at 25°K. In this sample the donor atoms greatly outnumber the acceptor atoms; that is, the sample is not highly compensated. This means that below 25°K most of the electrons are attached to arsenic atoms and are interacting to form the impurity band. If the impurity band is practically full, and if it does not overlap the conduction band, it should act like the valence band of a semiconductor and give rise to a positive Seebeck effect. This indeed we believe is the qualitative explanation of the reversal of sign shown in Fig. 6. The Seebeck voltage goes from negative to positive at 16°K. It becomes more positive as the temperature is lowered, perhaps roughly according to the equation

$$Q=(Q_c\sigma_c+Q_i\sigma_i)/\sigma, \tag{1}$$

where σ is conductivity and the subscripts refer to conduction and impurity bands. This equation follows from the crude assumption that the impurity and

---

[9] The results given below were presented in part at the Seattle meeting of the Physical Society [Phys. Rev. **96**, 847(A) (1954)].

[10] This mechanism has been discussed recently by many authors, see footnote 11 of reference 8. See also C. Hung and J. R. Gliessman, Phys. Rev. **96**, 1226 (1954).

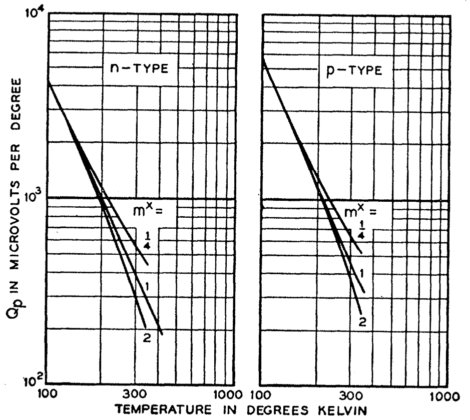

FIG. 7. Dependence of the calculated value of Q_p of the purer samples upon choice of mass parameter used in the density of states equation.

conduction bands are acting in parallel. Q_i must be essentially constant below 25°K. This is because the small percentage of compensation by the acceptor atoms present is now large with respect to the number of electrons in the conduction band, n_c; hence the number of holes in the impurity band is almost constant. Since σ_c varies directly as n_c while Q_c varies as $-\log n_c + dQ_p/dT$, σ_c is probably the dominant term in Eq. (1) for values of Q near 0. A value of Q_i would enable one to estimate the energy difference between the conduction band and the almost filled impurity band by a plot of $\log(Q_c - Q)$ against $1/T$. We are currently constructing apparatus that will extend the measurements to lower temperatures in order to provide an experimental estimate of Q_i.

The aforementioned model can equally well explain the reversal of the Seebeck voltage of the normally *p*-type sample No. 141 at 19°K.

In order to test this model further, a crystal was grown with both arsenic and boron present. From this crystal, sample 563 was prepared. From the growth conditions of the sample, and taking the distribution coefficient of boron in silicon as 1,[11] it is estimated that the concentration of boron in sample 563 is 1.0×10^{18}. The room-temperature Hall effect gives the concentration of arsenic as 1.25×10^{17} atoms per cc greater than the boron concentration. Thus sample 563 is about 90 percent compensated. At low temperatures, therefore, the impurity band must be only about 10 percent full. On the basis of the sample model described above, Q should not change sign. The low-temperature experimental points taken for the sample are also shown in Fig. 6. The general nature of the curve can be seen in Fig. 1. Because of the difficulty of measuring a small voltage over a high-impedance circuit, the low-temperature points scatter; however, they appear to extrapolate to 0 at $T = 0$. This is the limit required by the third law. More conclusive evidence must await the extension of the measurements to lower temperatures.

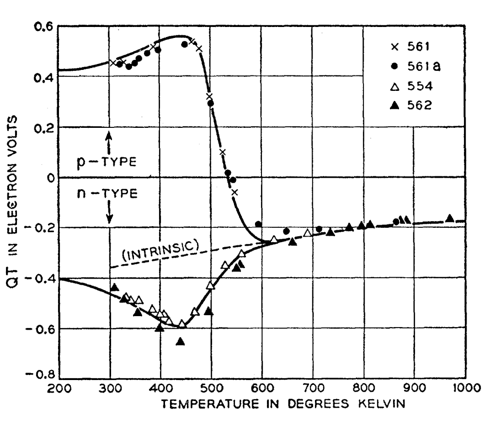

FIG. 8. Calculated values (curves) and measured values (points) for *n*-type and *p*-type samples measured into the intrinsic range.

4. The Effective Mass

The effective mass which enters into the density of states equation [Eq. (2), reference 1] can be determined by a comparison of charge carrier density with the electronic contribution to the thermoelectric power, Q_e. This is discussed in the previous paper.[1] In the present case it is difficult to obtain an accurate value of Q_e at room temperature because of the uncertainty in Q_p. Q_p in silicon does not extrapolate to a small value as it does in germanium. An uncertainty of 90 μv is equivalent to a factor of two in the mass parameter. We have chosen a mass parameter of unity since this minimizes the curvature in the $\log Q_p$ *versus* $\log T$ curves near room temperature. It can be seen from Fig. 7 that changing the mass parameter results in curvature. In this figure curves are given for Q_p of samples 537 and 552 calculated for a mass parameter of $\frac{1}{4}$, 1, and 2. It is seen that Herring's theory which predicts $Q_p \propto T^{-n}$ at temperatures well below the Debye temperature agrees with a choice of mass parameter close to unity. A value near unity is in agreement with the results of Dexter, Lax, Kip, and Dresselhaus[2] for *n*-type silicon which give a value $N_v^{\frac{2}{3}} m^*/m = 6^{\frac{2}{3}}(0.26) = 0.86$. A comparison of the *p*-type results is difficult due to spin-orbit splitting. The results in the following section, above room temperature, where Q_p is smaller are not accurate enough to use in this type of analysis.

5. Measurements in the Intrinsic Range

The samples measured in the intrinsic region listed in Table I were prepared from samples 537 and 552. The measurements were made in a conventional manner in a vertical furnace using platinum-platinum rhodium thermocouples. The Seebeck voltage was read across the Pt leads. Due to the high thermal conductivity of the silicon, it was necessary to reduce the cross section

---

[11] N. B. Hannay (personal communication).

of the samples between the thermocouples in order to avoid radial thermal discontinuities. The points shown in Fig. 8 are the measured values. The curves are calculated from the equation

$$QT=[(Q_+n_+\mu_++Q_-n_-\mu_-)T_++Q_{(\mathrm{Pt})}T]/(n_+\mu_++n_-\mu_-) \quad (2)$$

Q_+ and Q_- are the values either the holes or electrons would have in the absence of the other. The values include a phonon contribution estimated from Figs. 3 and 4. $Q_{(\mathrm{Pt})}$, the absolute thermoelectric power of platinum, is used in order to have the calculated value apply to the silicon-platinum couple. It is important only in the purely intrinsic range where it amounts to something less than a 10 percent correction. The concentrations of holes and electrons, n_+ and n_-, and their mobilities u_+ and μ_-, have been taken from the data of Morin and Maita.[6] Values of $|n_+-n_-|$ are determined from resistivity measurements in the saturation region taken concurrently with the thermoelectric readings using the thermocouples as potential probes. These are combined with the values of $(n_+n_-)^{+\frac{1}{2}}$ given in Fig. 13 of Morin and Maita. The values of μ_+ and μ_- are taken from the extrapolated values given in their Fig. 12. These represent a major uncertainty. The reasonable agreement shown in Fig. 7 is significant in demonstrating the overall consistency of the data; however, it seems doubtful that the choice of different mass parameters to improve the fit would be meaningful.

ACKNOWLEDGMENTS

We are greatly indebted to C. Herring and F. J. Morin for many useful comments and discussions concerning the ideas presented here. We also wish to acknowledge the important contribution of E. Buehler and E. Kolb in preparing the single crystals of silicon from which our samples were prepared.

PHYSICAL REVIEW VOLUME 98, NUMBER 4 MAY 15, 1955

Plasma Oscillations in Metals

E. N. Adams
Institute for Nuclear Studies, University of Chicago, Chicago, Illinois

(Received July 26, 1954; revised manuscript received February 7, 1955)

The self-consistent field treatment of the electronic wave functions of a solid is used as a basis for a theory of plasma oscillations in a metal. Two procedures are used to calculate the plasma frequency, both being based on the Bohm and Pines method of introducing the plasma oscillators. It is found that band-to-band virtual transitions have a strong effect on the plasma oscillations in a metal. It is concluded that the effective mass treatments are not adequate to treat the electron-plasma interaction, except perhaps in certain semiconductors.

INTRODUCTION

In the last few years, several papers have appeared which discuss the theory of the formation of an electron plasma[1–5] in metals. In this paper we would like to treat an aspect of that problem which we believe has been given insufficient attention, *viz.*, the effect of the energy band structure on the plasma frequency.

The effect of the lattice potential on the plasma frequency has heretofore to our knowledge been treated only by Wolff,[5] who limited his explicit consideration to matrix elements of the electron-electron interaction leading to in-band scattering processes. Wolff estimated for a particular case the effects of interband scattering processes on the plasma frequency and found that they caused a large frequency shift. We will show in the following that for metals band-to-band processes are in general quite important, so that Wolff's quantitative results are probably limited in applicability to semiconductors.

Although previous work has sought to establish the dispersion relation for plasma waves, we have been content to calculate the plasma frequency for very long wavelength plasma waves. Our results indicate that this frequency is strongly affected by band-band virtual transitions, the effects of which are difficult to calculate accurately. We have not concerned ourselves, therefore, with extending the dispersion relation since we feel that this cannot be done correctly in a treatment which does not make an adequate treatment of band-band transitions.

In carrying out our calculations we have chosen to use the method of Bohm and Pines (BP)[2,3] for introducing the generalized coordinates of the plasma modes. We made this choice because we feel that their method is basically the simplest we know and leads most directly to the results we want. However, we suspect that certain aspects of their treatment of the auxiliary conditions are incorrect, and believe the problem of the auxiliary con-

---

[1] D. Bohm and E. P. Gross, Phys. Rev. 75, 1851, 1854 (1949).

[2] D. Bohm and D. Pines, Phys. Rev. 85, 338 (1952).

[3] D. Bohm and D. Pines, Phys. Rev. 92, 609 (1953).

[4] D. Pines, Phys. Rev. 92, 626 (1953).

[5] P. Wolff, Phys. Rev. 92, 18 (1953).