PHYSICAL REVIEW VOLUME 185, NUMBER 3 15 SEPTEMBER 1969

Coulomb Screening in Intrinsic Medium-Gap Semiconductors and the Electrical Conductivity of Silicon at Elevated Temperatures

L. C. BURTON AND A. H. MADJID
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802
(Received 25 February 1969)

As the temperature of medium-gap intrinsic semiconductors is raised and the concentration of charge carriers increases beyond 10^{17} cm^{-3}, the average distance between the charges will have decreased sufficiently for the electrostatic interaction energy between each charge and its neighbors to be an appreciable fraction of the width of the forbidden energy region. From this point onwards, and through the onset of degeneracy, the temperature dependence of the concentration and of the width of the energy gap will depend appreciably upon this energy. It is the object of this paper to give an account of this phenomenon; to compute the temperature variation of the concentration, of the energy gap, and of the Herring screening length; and to compare these results to experimental values deduced from the measurement of the electrical conductivity of silicon from 500°K up to about 50° away from the melting point. An estimate of the effective conductivity mobility for the high-temperature range is also presented.

I. INTRODUCTION

IN treating transport phenomena in medium-gap semiconductors such as silicon, it is usual to allow for the onset of degeneracy by changing from Maxwell-Boltzmann to Fermi-Dirac statistics. Although this approach can be correct formally, it is limited and ignores important processes occurring within the semiconductor as its temperature is raised and the electron and hole gases start to become degenerate.

These processes pertain to the interaction of charge carriers. If their concentration, for example, is raised from 10^{15} cm^{-3} to approach 10^{23} cm^{-3}, the average distance between charges drops from about 1000 Å to several angstroms, and the change is from a virtually noninteracting to a very strongly interacting assembly of charges.

Of course, if the concentration reaches 10^{19} cm^{-3}, and the average distance between carriers has decreased sufficiently for particles to begin to feel the spin structure of the neighbor gas, the use of Fermi-Dirac statistics will become mandatory. But earlier than that, for a concentration of 10^{17} cm^{-3}, when the use of Maxwell-Boltzmann statistics still remains legitimate, the charges will begin to interact by long range, classical, Coulomb forces.

This interaction energy will alter the electrochemical potential of the particle gases, and the forbidden energy gap, consequently, will start to depend on the concentration of carriers once the interaction energy exceeds an appreciable fraction of the width of the forbidden energy region. The onset of degeneracy will then cause the collapse of the gap to proceed at a slower rate as the classical interaction forces are gradually being swamped by the more powerful short-range quantum-mechanical forces between the carriers. The point, thus, is that degeneracy is not only characterized by a change in statistics, but additionally, by a decrease in the width of the forbidden energy region; a point that is sometimes ignored when treating transport properties of intrinsic medium-gap semiconductors at high temperatures.

Herring[1] was the first to point out that electrostatic interactions between charge carriers will give rise to a decrease in the energy gap, and to calculate this decrease by applying the Debye-Hückel theory for electrolytes to the electron and hole gases within intrinsic semiconductors; and Morin and Maita[1,2] used Herring's suggestions to explain a discrepancy between their calculated and their experimentally determined data pertaining to the temperature variation of the charge carrier concentration in intrinsic silicon and Ge. The characteristic screening length associated with this phenomenon is therefore referred to as the "Herring length" throughout this paper.

Herring screening in intrinsic semiconductors has as yet not received the attention that the importance of this effect deserves. It is the object of this paper to attempt an experimental verification of the effect by studying the electrical conductivity of silicon from about 50° away from the melting point down to 500°K. Such conductivity data (that are free of the criticism that sample contamination from the ambient during high-temperature cycling is serious) have in turn not as yet been reported in the literature over the temperature range stated.

The specific procedure will be to calculate the temperature dependence of the charge carrier concentration using the concept of screening and the pertinent property parameters of silicon, and to compare this result to the temperature variation of the charge carriers deduced from the measurement of the conductivity of various samples of silicon.

II. THEORY

In this section we will first formulate the expression for the temperature dependence of the width of the forbidden energy region including the screening term. Next the screening length will be discussed for non-, moderate, and complete degeneracy. These results will

[1] F. J. Morin and J. P. Maita, Phys. Rev. **94**, 1525 (1954).
[2] F. J. Morin and J. P. Maita, Phys. Rev. **96**, 28 (1954).

then be used to compute the temperature dependence of the charge-carrier concentration of intrinsic silicon.

A. Forbidden Energy Gap

The true temperature-dependent forbidden energy gap $E_G(T)$ is, for our purposes, most conveniently expressed as

$$E_G(T)=E_G{}^0+(dE_G/dT)^L T-\frac{e^2}{4\pi\epsilon_0\kappa}\frac{1}{\lambda_H}. \qquad (1)$$

Here $E_G{}^0$ is arrived at by linearly extrapolating from the intermediate temperature range, where $E_G(T)$ varies linearly with temperature, to $T=0°$K. This is shown in Fig. 1 which is based on data published by MacFarlane *et al.*[3] and which shows $E_G{}^0$ to be 1.205 eV.

Here $(dE_G/dT)^L$ is the temperature derivative of the gap for intermediate temperatures, where the gap variation is linear. It is taken as -4.0×10^{-4} (eV/°K) from Burstein and Egli[4] using conductivity data, from Bardeen and Shockley[5] using the temperature dependence of the spectral sensitivity of a p-n junction photocell, and from Fan, Shepherd, and Spitzer's[6] and MacFarlane and Roberts'[3] results. This term is attributed to lattice dilatation with temperature, and to a band-broadening contribution arising from temperature-dependent electron-lattice interaction terms.[7]

The last terms finally, represents the decrease of the forbidden energy region that is caused by the interaction

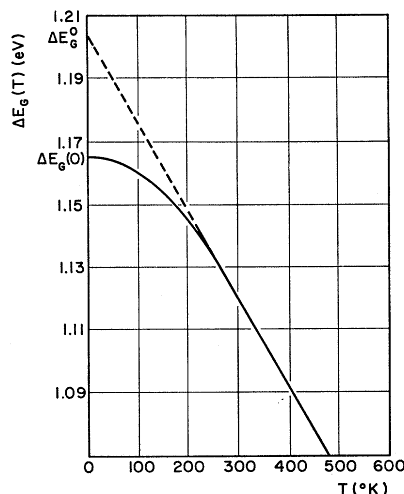

FIG. 1. The temperature dependence of the width of the forbidden energy region after the data of MacFarlane *et al.* (Ref. 3).

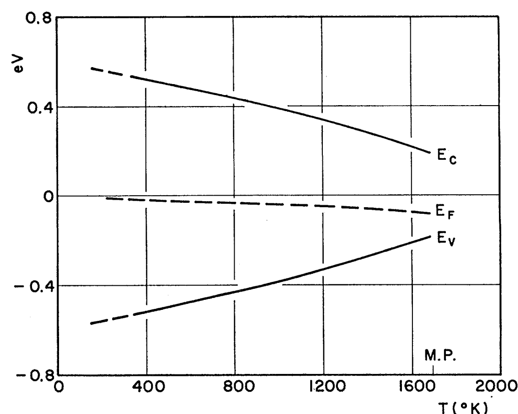

FIG. 2. The computed temperature variation of the forbidden energy region for silicon. The departure from linearity in the high-temperature region is caused by screening effects.

energy of the charge carriers. Here e is the absolute value of the electronic charge, κ is the dielectric constant for silicon taken as 11.7[8,9] and λ_H is the characteristic Herring screening length, which is the subject of Sec. II B.

Once $\lambda_H(T)$ is determined, $E_G(T)$ can be computed, and these results are shown in Fig. 2.

It is seen that from about 300 (Fig. 1) to about 700°K, $E_G(T)$ remains linear with $(dE_G/dT)^L$ predominating. From 700 to about 1300°K, the forbidden energy region collapses more rapidly owing to the additional effect of the (as yet nondegenerate) screening term. And above 1300° up to melting, the gap variation slows down because of saturation effects caused by the onset of degeneracy.

B. Characteristic Screening Length

The expression for the characteristic screening length follows from the Debye-Hückel theory, and is

$$\lambda_H=\left[\frac{\epsilon_0\kappa kT}{e^2(n+p)}\right]^{1/2}. \qquad (2)$$

Here ϵ_0 is the permittivity of free space, k is Boltzmann's constant, T is the absolute temperature, and n and p are the concentrations of the electrons and the holes, respectively.

The problem of screening in a charge-carrier gas of arbitrary degeneracy has been treated by Dingle[10] and by Conwell and Levinger[11] with the result that λ_H may be calculated for the case of intrinsic silicon from

$$\lambda_H=\left[\frac{8\pi e^2(2\pi kT)^{1/2}}{\epsilon_0\kappa h^3}M_c m_{dn}{}^{3/2}\mathbf{F}_{-1/2}(\eta)\right]^{-1/2}. \qquad (3)$$

[3] G. G. MacFarlane, T. P. McLean, J. E. Quarington, and V. Roberts, Phys. Rev. **111**, 1245 (1958).
[4] E. Burstein and P. H. Egli, *Advances in Electronics and Electron Physics* (Academic Press Inc., New York, 1955), Vol. 7.
[5] J. Bardeen and W. Shockley, Phys. Rev. **80**, 72 (1950).
[6] H. Y. Fan, M. L. Shepherd, and W. Spitzer, in *Photoconductivity Conference*, edited by R. G. Breckenridge, B. R. Russell, and E. E. Hahn (John Wiley & Sons, Inc., New York, 1956).
[7] T. S. Moss, *Optical Properties of Semiconductors* (Butterworths Scientific Publications Ltd., London, 1959).
[8] C. D. Salzberg and J. J. Villa, J. Opt. Soc. Am. **47**, 244 (1957).
[9] W. C. Dunlap and R. L. Watters, Phys. Rev. **92**, 1396 (1953).
[10] R. B. Dingle, Phil. Mag. **46**, 831 (1955).
[11] E. M. Conwell and B. W. Levinger, in *Proceedings of the Conference on the Physics of Semiconductors, Exeter* (The Institute of Physics and the Physical Society, London, 1962).

h is Planck's constant, $M_c=6$ allows for the 6 valleys in the conduction band of silicon, m_{dn} is the density-of-states effective mass of the conduction band electrons, and $\mathsf{F}_{-1/2}(\eta)$ is the Fermi-Dirac function defined and tabulated by Dingle.[10] In this expression $\eta=(E_F-E_c)/kT$ with E_F the Fermi energy and E_c the energy of the conduction band edge.

The two limiting forms,

$$\lambda_H^{\text{nondegen.}}=\left[\frac{\epsilon_0\kappa kT}{2e^2n}\right]^{1/2} \tag{4}$$

and

$$\lambda_H^{\text{compl. degen.}}=\frac{h(\epsilon_0\kappa)^{1/2}}{(4\pi e^2M_c{}^{2/3}m_{dn})^{1/2}(6/\pi n)^{1/6}}, \tag{5}$$

are also given by Dingle. But the electron and hole gases never reach complete degeneracy up to the melting point of silicon. $\lambda_H^{\text{nondegen.}}$ is valid only up to about 1300°K. It was therefore necessary to compute λ_H for the partially degenerate case. Using the approximation given by Blakemore,[12]

$$\mathsf{F}_{1/2}(\eta)\approx\frac{e^\eta}{1+0.27e^\eta}, \tag{6}$$

which is valid to $\pm 3\%$ for $\eta<+1.3$, the screening length was computed by first solving for $\eta(T)$. [This involved noting that for intrinsic conditions the concentration of electrons and holes is equal, giving a relation between $E_G(T)$ and η; eliminating $E_G(T)$ using (1); eliminating λ_H by using (3); and by plotting the resulting equation in η and T to determine the temperature dependence $\eta(T)$.] The results are shown in Fig. 3, which specifies λ_H for silicon together with the limiting forms (4) and (5).

C. Concentration of Charge Carriers

Having arrived at the temperature dependence of $\eta(T)$, $E_G(T)$, and $\lambda_H(T)$, the computation of the temperature variation of the concentration of the charge carriers is straight forward. Up to 700°K the usual expression,

$$n(T)^{\text{comp.}}=2\left(\frac{2\pi kT}{h^2}\right)^{3/2}[m_{dp}m_{dn}M_c{}^{2/3}]^{3/4}\times\exp-\left\{\left[E_G{}^0+\left(\frac{dE_G}{dT}\right)^L\right]\Big/2kT\right\}, \tag{7}$$

suffices. From 700 to 1300°K, $n(T)^{\text{comp.}}$ follows from

$$n(T)^{\text{comp.}}=2\left(\frac{2\pi kT}{h^2}\right)^{3/2}[m_{dp}m_{dn}M_c{}^{2/3}]^{3/4}\times\exp-\left\{\left[E_G{}^0+\left(\frac{dE_G}{dT}\right)^L-\left(\frac{e^2\epsilon_0kT}{2\kappa n(T)}\right)^{1/2}\right]\Big/2kT\right\}. \tag{8}$$

This equation remains based on classical statistics but now includes screening. Above 1300° up to the melting point, allowance for the onset of degeneracy must be made, and $n(T)$ must be determined from

$$n(T)^{\text{comp.}}=2\left(\frac{2\pi kT}{h^2}\right)^{3/2}[m_{dn}{}^2M_c{}^{2/3}]^{3/4}\mathsf{F}_{1/2}(\eta). \tag{9}$$

The results, with $m_{dn}=0.33m_0$ and $m_{dp}=0.58m_0$,[13] are shown in Fig. 4, where $n(T)^{\text{comp.}}$ is shown together with a curve that is based on Eq. (7) (screening effect not

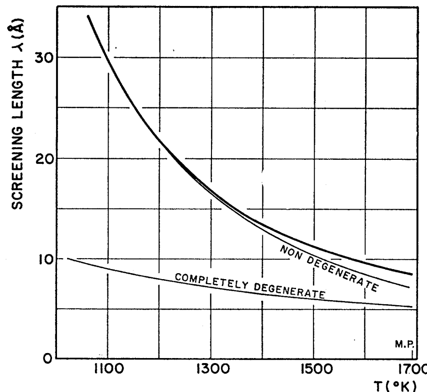

FIG. 3. The temperature variation of the Herring screening length. The limiting cases of λ_H, for nondegeneracy and complete degeneracy are also shown.

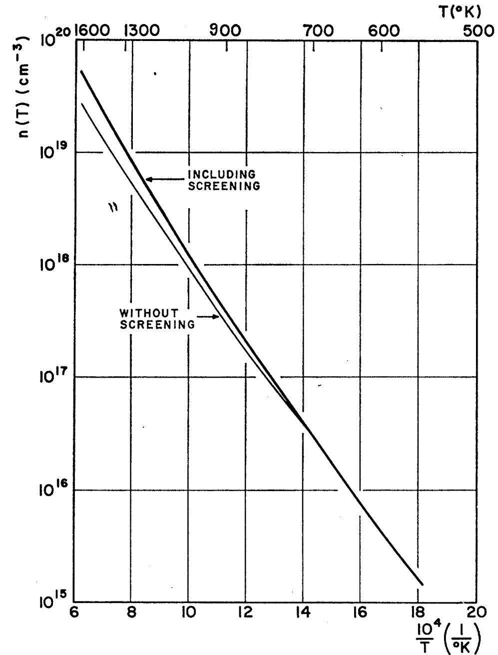

FIG. 4. The comparison of the computed temperature variation of the concentration of electrons and holes in silicon including and excluding screening effects.

[12] J. S. Blakemore, Proc. Phys. Soc. (London) A65, 460 (1952).

[13] B. Lax and J. G. Mavroides, Phys. Rev. 100, 1650 (1955).

included). It is seen that both curves differ appreciably in the high-temperature range where screening effects may not be ignored. It is this curve $n(T)^{\text{comp.}}$ that we ultimately wish to compare to the results deduced from the measurement of conductivity.

III. EXPERIMENT

The conductivity was measured by conventional four-probe dc and also ac techniques. The current through the sample when using the dc method, and a superimposed dc current when using the ac method, served to maintain the sample at the required temperature. In this way only the sample proper and the parts of the four tantalum probes that were in contact with the sample assumed operating temperatures during the runs. This assured that the contamination, arising from the source that a hot ambient certainly represents, remained minimal. The temperature was measured with an optical pyrometer (using the emissivity values published by Allen[14]) for high temperatures, and using W-5% Re–W-26% Re thermocouples for temperatures below the pyrometer range.

The conductivity samples consisted of wafers cut from single crystal ingots of zone-refined silicon. The doping range of the crystals used is summarized in Table I.

Although samples of all doping should, in principle, exhibit the same intrinsic characteristic, we have chosen to measure over the entire doping range. To begin with, samples of low doping remain intrinsic over a large temperature range and are therefore most suitable for our purpose. Yet, samples of low impurity content are easily contaminated; and it is therefore well to have the assurance of the same results for samples of moderate doping. For silicon of heavy doping, finally, the possibility of anomalous secondary effects could not have been *a priori* excluded. All elements used as dopants for silicon, such as boron and phosphorus, occupy substitutional sites in the silicon host lattice; but there will always exist a certain mismatch between the ionic radii of the host and the impurity atoms. An excessive aggregate of such perturbations, such as will occur in both n- and p-type silicon of extreme doping, may (but need not) give rise to secondary effects not characteristic of specimens of greater purity. (An example within this context is that a silicon lattice that is strained by extreme doping will relax along the slip planes. This effect is easily observable under a microscope.) Our results showed, however, that these cautions were unnecessary, since all samples examined showed identical intrinsic characteristics within the limits of error of the measurement.

TABLE I. Tabulation of the resistivity at room temperature for the n- and the p-type samples used in this experiment.

Sample number	Type	Room-temperature resistivity (Ω cm)
5	n	0.05
9	n	0.05
7	n	2
10	n	2
8	n	100
11	n	100
12	p	0.4
13	p	5

[14] F. G. Allen, J. Appl. Phys. **28**, 1510 (1957).

The most crucial experimental entity was the vacuum system within which the silicon was heated by passing current through the sample. Workers in the field of silicon surface technology know that to heat-cycle silicon within such systems is a most demanding task of vacuum technology.[15–17] The adversity rests in the fact that when silicon is heated during the normal course of vacuum processing within systems that contain (or even only recently contained) the slightest glass component, a high-conductivity p-type layer invariably develops on its surface. Busch, Schade, Gobbi, and Marmier,[16] using nuclear activation techniques, verified that the impurity involved was boron. Depending upon the particular experimental set up and upon the heating schedule that the silicon is subjected to, the concentration of boron atoms at the surface may range between 10^{17} and 10^{21} cm^{-3}. Layer thicknesses do not exceed the

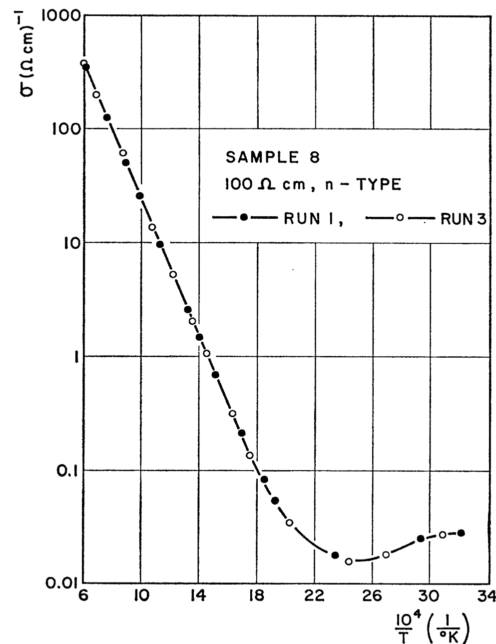

FIG. 5. The first and the third conductivity run for a 100-Ω cm, n-type silicon sample. The coincidence of these two plots in the extrinsic region suggests that the ambient in which the samples were heated did not contaminate the sample.

[15] F. G. Allen, T. M. Buck, and J. T. Law, J. Appl. Phys. **31**, 979 (1960).

[16] G. Busch, G. H. Schade, A. Gobbi, and P. Marmier, J. Appl. Phys. **127**, 149 (1962).

[17] G. Busch and A. H. Madjid, Physik Kondensierten Materie **4**, 131 (1965).

order of 10 μm, however, even if heating is prolonged to many hours, because of the slow diffusion of boron atoms in silicon host crystals.

Layers such as these will effectively short out the bulk of all but the most extremely boron doped p-type samples, and they will therefore introduce grave errors when measuring the conductivity of thus contaminated silicon samples.

To avoid difficulties such as these we have developed a vacuum system in which silicon samples could be heat-cycled in the 10^{-9}–10^{-10} Torr range without contaminating them above the limit of resolution of the present measurement. The most sensitive test of this contention is to heat-cycle a moderately doped n-type sample within the enclosure and to monitor the extrinsic conductivity as a function of the cycling. The results of such a test are shown in Fig. 5. It is seen that the conductivity versus reciprocal temperature plots for the first and the third cycle coincide within the resolution of the measurement, and that the system is therefore virtually noncontaminating.

This apparatus will be described in greater detail elsewhere,[18] but it consisted, briefly, of an ion-pumped, sapphire ported, water-cooled stainless-steel system incorporating Al_2O_3 feed throughs.

The results of the conductivity measurements for the n- and p-type samples are shown in Fig. 6. The appearance of the data is straightforward, and no anomalies are evident except for a seeming slight oscillation of the data points in the intrinsic region around a straight line, with the amplitude of the oscillations decaying towards higher temperatures.

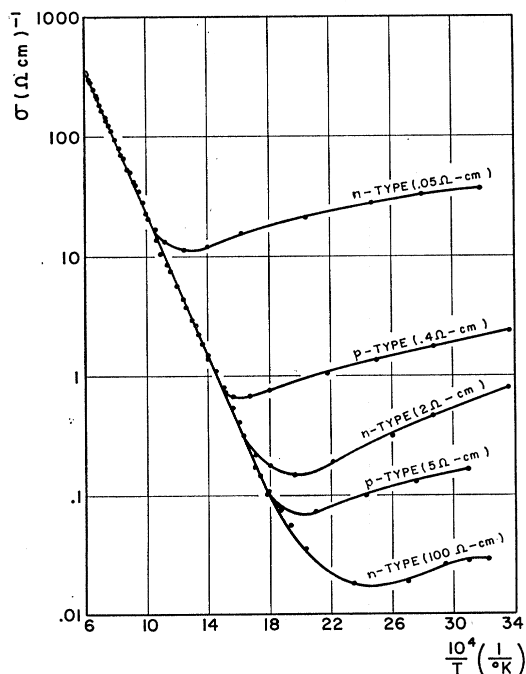

FIG. 6. The results of the conductivity measurements for n- and p-type silicon samples for various doping.

[18] L. Burton, S. Slack, and A. H. Madjid (unpublished).

TABLE II. Tabulation of the lattice conductivity mobilities for electrons, $\mu_n{}^c$, and for holes, $\mu_p{}^c$, and the effective conductivity $\bar{\mu}_c$ after Morin and Maita (Ref. 2) and Ludwig and Watters (Ref. 19).

$\frac{(\text{cm}^2)}{(\text{volt sec})}$	Morin and Maita[a]	Ludwig and Watters[b]
$\mu_n{}^c$	$4\times10^9T^{-2.6}$	$2.1\times10^9T^{-2.5}$
$\mu_p{}^c$	$2.5\times10^8T^{-2.3}$	$2.3\times10^9T^{-2.7}$
$\bar{\mu}_c$	$\cong1.5\times10^9T^{-2.4}$	$\cong3.2\times10^9T^{-2.6}$

[a] Reference 2. [b] Reference 19.

These oscillations are not apparent in the plots shown because of the smallness of the effect and because of the scale factor of the graphs. The behavior did not depend on doping and it thus seemed to be a property of the silicon host lattice. We are at present trying to establish this phenomenon unambiguously and quantitatively and will make it the subject of a separate communication.

IV. COMPARISON BETWEEN THEORY AND EXPERIMENT

A. Concentration of Charge Carriers

The object is to compare the computed temperature dependence of the concentration $n(T)^{\text{comp.}}$ with the experimental results using

$$n(T)^{\text{expt}}=\frac{\sigma}{e(\mu_n{}^c+\mu_p{}^c)} \tag{10}$$

with σ the intrinsic conductivity of silicon, $\mu_n{}^c$ the conductivity mobility of electrons, and $\mu_p{}^c$ the conductivity mobility of holes. ($\mu_p{}^c$ actually is the weighted average over the light- and the heavy-hole mobilities.) It is convenient within this context to define an effective conductivity mobility,

$$\bar{\mu}_c=\mu_n{}^c+\mu_p{}^c, \tag{11}$$

which was determined from the results of Morin and Maita[2] and of Ludwig and Watters.[19] These results are summarized in Table II.

The use of these two sets of mobility values did not lead to an appreciable difference as far as the determination of $n(T)^{\text{expt}}$ was concerned; and we have therefore used the values published by Morin and Maita,[2] which gave a slightly better fit to the computed data, to display the comparison between $n(T)^{\text{comp.}}$ and $n(T)^{\text{expt}}$ in Fig. 7.

B. Discussion of Parameters Used and of Results

The agreement between the computed concentration and the one deduced from conductivity data is surprisingly good, considering the paucity of the data that is available describing the properties of silicon in the high-temperature range. The agreement is, at any rate,

[19] G. W. Ludwig and R. L. Watters, Phys. Rev. **101**, 1699 (1956).

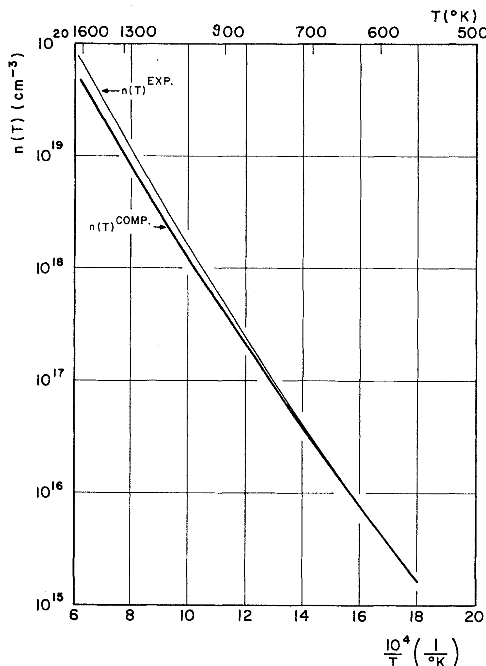

Fig. 7. The comparison of the computed temperature dependence of the concentration of the charge carriers in silicon with that deduced from electrical conductivity data.

much better than it would have been if screening terms had been ignored. (See Fig. 4.) But the two curves do not coincide, and their slopes also differ by some moderate amount. It is therefore necessary to examine the parameters used to obtain the two results displayed in Fig. 7, all of which were extrapolated from low or moderate temperatures to about 1600°K.

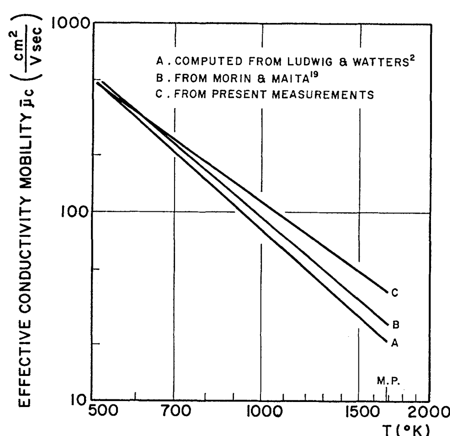

Fig. 8. The temperature variation of the effective conductivity mobility obtained by using the data of Morin and Maita (Ref. 2), Ludwig and Watters (Ref. 19), and that deduced from the present measurement of the conductivity and the computed temperature dependence of the concentration.

There are, to begin with, the values of the effective masses and of $(dE_G/dT)^L$. It would have been conservative to increase the low-temperature values of the effective masses used by 10–20%[20,21]; and it would not have been unreasonable to expect that $(dE_G/dT)^L$ changes by about 5% in a thousand degrees. These moderate changes in these two pre-exponential parameters would have been sufficient to bring the curves of Fig. 7 into coincidence.

This still leaves the disparity in the slopes to deal with. The difficulty here must probably be apportioned to the value of the effective mobility $\bar{\mu}_c$ used to obtain $n(T)^{\text{expt}}$. The extrapolation involved was over many hundred degrees and it is not surprising that this procedure, dictated by necessity but nevertheless arbitrary, leads to some disagreement. Lattice mobilities were used to compute $\bar{\mu}_c$ because the corrections resulting from electron-hole scattering terms would have been minor when compared to the uncertainty represented by the extrapolation mentioned.

It is interesting to note within this context that the prediction that, owing to intervalley scattering, the temperature dependence of the mobility would revert to $\mu \propto T^{-1.5}$, is not born out by our results.

Table III. Effective mobility values $\bar{\mu}_c = \mu_n^c + \mu_p^c$ computed from Morin and Maita's (Ref. 2) and Ludwig and Watters's (Ref. 19) data and those determined from present measurements.

$\frac{(\text{cm}^2)}{(\text{volt sec})}$	Computed from Morin and Maita[a]	Computed from Ludwig and Watters[b]	Determined from present measurements
$\bar{\mu}_c$	$\cong 1.5\times 10^9 T^{-2.4}$	$\cong 3.2\times 10^9 T^{-2.5}$	$2.13\times 10^8 T^{-2.1}$

[a] Reference 2. [b] Reference 19.

This is most easily seen by inverting our argument by considering $n(T)^{\text{comp.}}$ as being true and to compute a mobility value for high temperatures from

$$\bar{\mu}_c = \frac{\sigma^{\text{meas.}}}{en(T)^{\text{comp.}}}. \tag{12}$$

The results are shown in Fig. 8 and are summarized in Table III.

ACKNOWLEDGMENTS

The authors wish to express their appreciation to the National Aeronautics and Space Administration, Washington, D. C., for their support of this work; to H. Henry and T. Sauer for their help with shop problems; and to Miss F. Fogle and Miss S. Wallace for preparing the manuscript.

[20] E. M. Conwell, Proc. I.R.E. **46**, 1281 (1958).

[21] Yu. I. Ukhanov and Yu. V. Mal'tsev, Fiz. Tverd. Tela **5**, 2926 (1964) [English transl.: Soviet Phys.—Solid State **5**, 2144 (1964)].