by a factor of 2 or more. In Eq. (5) the main contribution to the longitudinal phonon damping comes from the term $\mathbf{K}=0$. The phonon wave vectors range over a region where the ion-electron pseudopotentials given by various authors agree very well (indicated by the solid horizontal line in Fig. 1). For the transverse phonons, on the other hand, the term $\mathbf{K}=0$ vanishes and the $(\mathbf{q}+\mathbf{K})$ vectors of main importance (broken horizontal line in Fig. 1) lie in a region where the pseudopotential is most uncertain. We have also made calculations using Bardeen's model,[7,9] which is expected to give values too small for the pseudopotential at $q\sim 2k_{\mathrm{F}}$. We found significantly longer lifetimes for the transverse phonons whereas the change for the longitudinal ones was smaller. In Kreb's model,[10] the lifetimes are still longer.

Whereas the lifetimes obtained for transverse phonons may be in error by a factor of 2 or more, we feel confident that for the longitudinal phonons the results should be quite good.[11] Therefore, our conclusion is that at least half of the contribution to the damping $1/\tau$ comes from anharmonic effects, and that the damping arising from electron-phonon coupling is by no means negligible.

In principle measurements of particularly the transverse phonon lifetimes after subtraction of the anharmonic contribution should give quite detailed information about the ion-electron interaction in the region where it is most uncertain. However, because of the extreme difficulties in carrying out such measurements, it does not seem feasible at present to pursue such a program.

Concerning the conjecture of Nilsson and Stedman, we stress that the anharmonic contribution should also give longer lifetimes for the transverse phonons than for the longitudinal phonons. At very low temperatures, the transverse phonons are less likely to decay into two or more phonons compared to the longitudinal phonons because of the conservation laws for energy and momentum.[12]

ACKNOWLEDGMENTS

The authors wish to thank Dr. R. Stedman for stimulating discussions and Dr. K. S. Singwi for comments in connection with this problem.

[9] J. Bardeen, Phys. Rev. **52**, 688 (1937).
[10] K. Krebs, Phys. Rev. **138**, A143 (1965).
[11] We have taken for granted that the dielectric function used above is sufficiently accurate for Al.
[12] R. E. Peierls, *Quantum Theory of Solids* (Oxford University Press, London, 1955), pp. 43–45.

PHYSICAL REVIEW VOLUME 159, NUMBER 3 15 JULY 1967

Some Effects of Sample Size on Electrical Transport in Bismuth

ALLEN N. FRIEDMAN*†
IBM Watson Laboratories at Columbia University, New York, New York
(Received 23 November 1966)

The effects of sample size on the electrical conductivity and Hall constant of single crystals of bismuth have been measured between 2 and 10°K. In addition, both components of the conductivity tensor and the Hall tensor component $\rho_{23,1}$, with magnetic field $\parallel$ to the trigonal axis, have been measured between 2 and 77°K in samples which are among the purest reported in the literature. This work was done to determine the proportion of specular reflection of carriers from the electropolished sample surface. The nature of the experimental results, however, has made it impossible to fit observation to any sort of theory with mean free path as the internal parameter. Analysis of these data, however, has shown the following: The scattering is, indeed, partially specular; the scattering time of electrons is highly anisotropic, being of the order of ten times longer in the bisectrix direction than in the binary direction; and both electron and hole mobilities have been shown to have an approximate T^{-2} dependence down to 4.2°K.

I. INTRODUCTION

BISMUTH, a semimetal with electrical properties midway between insulators and metals, has long proven interesting to experimenters. As a result, phenomena such as the de Haas–van Alphen effect and the Azbel-Kaner resonance, which have proven to be valuable tools in the investigation of Fermi surfaces, were first observed in this material. As early as 1939, rather detailed models of the bismuth electron Fermi surface were beginning to appear in the literature[1] and investigations of the galvanomagnetic properties of this material and its alloys date back to 1928.[2]

The conduction electrons, produced by the small overlap of nearly filled 5th Brillouin zone with the 6th, lie within three nearly ellipsoidal regions of momentum space with one major axis of each in each of the binary directions. Directions in the bismuth crystal are specified with respect to three mutually perpendicular direc-

* Present address: Bell Telephone Laboratories, Murray Hill, New Jersey.
† Submitted in partial fulfillment of the requirements for the Ph.D. degree at Columbia University.
[1] D. Schoenberg, Proc. Roy. Soc. (London) **A170**, 341 (1939).
[2] P. Kapitza, Proc. Roy. Soc. **A119**, 358 (1928).

tions usually labeled 1, 2, and 3; 1—one of the three axes of twofold symmetry (binary direction), 3—the axis of threefold symmetry (trigonal direction), 2—an axis perpendicular to 1 and 3 forming a right-hand triad in the order 1-2-3 (bisectrix direction). In the 1, 2, 3 system one ellipsoid can be represented to a good approximation as

$$2m_0E_F{}^e=\alpha_1p_1{}^2+\alpha_2p_2{}^2+\alpha_3p_3{}^2+\alpha_4p_2p_3\,, \qquad (1)$$

the other two being generated by a rotation of $\pm 120°$ about the trigonal axis.

The results of many experiments has yielded a total electron density close to 3.0×10^{17} cm^{-3} and the reciprocal mass-tensor components; $\alpha_1=119, \alpha_2=0.76, \alpha_3=102$, and $\alpha_4=8.6$, and $E_F{}^e=15.4$ meV. We shall not list here the extensive literature on bismuth. More complete information can be obtained from excellent review articles.[3,4]

Since the five electrons per unit cell are just enough to fill one Brillouin zone, the overlap of electrons into the next higher zone leaves an equal number of hole carriers in the lower zone; their Fermi surface can be represented in momentum space as

$$2m_0E_F{}^h=\beta_1(p_1{}^2+p_2{}^2)+\beta_3p_3{}^2. \qquad (2)$$

Here

$$\beta_1=14.7\,,\quad \beta_2=1.09\,,\quad \text{and}\quad E_F{}^h=\sim 11 \text{ meV (Ref. 5)}.$$

A study of the galvanomagnetic properties of bismuth can start with consideration of its crystal symmetry. Group theory has been applied to the $\bar{3}m$ symmetry of bismuth by Okada[6] and Juretschke,[7] who found that, to second order in magnetic field, 12 independent constants characterize this material. If one generalizes Ohm's law as

$$E_i=\sum_j \rho_{ij}(H)J_j \qquad (3)$$

and expands $\rho_{ij}(H)$ for small magnetic fields as

$$\rho_{ij}(H)=\rho_{ij}+\rho_{ij,k}H_k+\rho_{ij,kl}H_kH_l+\cdots\,, \qquad (4)$$

then there are two independent resistivity constants ρ_{ij}, two Hall constants $\rho_{ij,k}$, and eight magnetoresistance constants $\rho_{ij,kl}$. Here the notation of Okada is used with the sum over repeated indices being assumed.

Herring and Vogt[8] (HV) have shown that, when all scattering processes either randomize velocity or cause energy changes small compared to the particle energy, a relaxation time exists. This allows the Boltzmann equation to be solved for the current corresponding to given fields. The second condition is well satisfied for a degenerate Fermi system such as bismuth at low temperatures, since scatterings are confined to the neighborhood of the Fermi surface. The first condition, however, is not met, as will be shown later. HV also show that the relaxation time tensor τ and reciprocal mass tensor α occur in the Boltzmann equation only as their product so that isotropic τ is not required for the validity of the analysis. Zitter[9] has applied the results of HV to the bismuth Fermi-surface model and calculated the conductivity-tensor components in terms of the components of the mobility tensor,

$$\mu=(e/m_0)\tau\alpha. \qquad (5)$$

The quantity of primary interest in this investigation, the carrier mean free path, can now be defined as

$$\lambda=\tau v. \qquad (6)$$

Several types of experiments are sensitive to the size of this quantity. It has been a natural length scale in interpretation of anamolous skin effect and Azbel-Kaner resonance and in the study of galvanomagnetic size effects.

Experiments of the last type have been done on several metals in which, at low temperatures, the carrier mean free path increases to a size at which samples can be made. Andrew[10] studied the size variation or resistivity of mercury wires down to 6-μ diam and tin foils down to 3 μ thick in the liquid-helium temperature range. He found that the mean free path of conduction electrons was shortened by collisions with the walls. MacDonald[11] was able to show that these were indeed boundary effects when he imposed a longitudinal magnetic field on thin wires of sodium. After an initial increase, the resistance of these wires decreased with increasing magnetic field since the carriers were caused to spiral around the field and stay away from the walls longer. Many experiments, then, have been successfully analyzed in terms of mean free path.

Several aspects of the galvanomagnetic size-effect problem have been theoretically investigated. The size effects of a quasifree-electron metal with spherical Fermi surfaces have been studied for several geometries[12–14] using the statistical approach of the Boltzmann equation. The size effect of conductivity in zero magnetic field was calculated by Fuchs[12] for a thin film with a combination of specular and diffuse surface scattering in materials with spherical energy surfaces. The similar problem for a square wire was done by MacDonald and Sarginson[13] and for a round wire by Dingle.[14] A common feature of all these calculations is the limiting behavior of conductivity at large and small thicknesses. For thicknesses much larger than the mean free path the effect of the surfaces is negligible and bulk properties

[3] A. L. Jain and S. H. Koenig, Phys. Rev. **127**, 442 (1962).
[4] W. S. Boyle and G. E. Smith, Progr. Semicond. **7**, 1 (1963).
[5] J. K. Galt, W. A. Yager, F. R. Merritt, and B. B. Cetlin, Phys. Rev. **114**, 1396 (1959).
[6] T. Okada, J. Phys. Soc. Japan **11**, 89 (1956).
[7] H. J. Juretschke, Acta Cryst. **8**, 716 (1955).
[8] C. Herring and E. Vogt, Phys. Rev. **101**, 944 (1956).
[9] R. N. Zitter, Phys. Rev. **127**, 1471 (1962).
[10] E. R. Andrew, Proc. Phys. Soc. (London) **62**, 77 (1949).
[11] D. K. C. MacDonald, Nature **163**, 639 (1949).
[12] K. Fuchs, Proc. Cambridge Phil. Soc. **34**, 100 (1938).
[13] D. K. C. MacDonald and K. Sarginson, Proc. Roy. Soc. (London) **A203**, 223 (1950).
[14] R. B. Dingle, Proc. Roy. Soc. (London) **A201**, 545 (1950).

are observed, while as the thickness approaches zero the conductivity approaches zero. The only exception to the latter is the case of pure specular reflection, in which no size effect is produced.

The case of anisotropic energy surfaces was investigated by Price[15] for pure specular and pure diffuse reflection using the kinetic approach and his "vector mean free path" concept. In this method the trajectory of a single carrier particle is followed until it scatters with loss of memory. In order to determine the relation between carrier states before and after a specular reflection from the boundary it is supposed that the direction of the change in crystal momentum on reflection is normal to the surface. These trajectories are then integrated over the Brillouin zone and averaged over the volume of the sample. The main results of this calculation are that there is a size effect of conductivity for specular reflection; the conductivity becomes size-independent as the thickness is reduced with a value less, in general, than the bulk value. These calculations have been extended by the present author to the case of mixed specular and diffuse reflection. This appears in the Appendix.

II. EXPERIMENTAL PROCEDURES

A. Sample Preparation

The single crystals used in this investigation were prepared by a modification of the Schubinkov[16] method from material obtained from Consolidated Mining and Smelting Company of Canada, Ltd. quoted as 99.9999% pure. They were grown to the shape needed for measurements in a mold made by stacking microscope slides. The seed orientation was done visually making use of the easy cleavage plane and the fact that slip lines seen in this plane are parallel to the binary direction. The sample orientations were checked by Laue diffraction and usually found to be within 1 deg of those desired.

For the size-effect studies the samples, originally 5 mm thick, were electropolished to remove the required material between measurements. The polishing solution was made by mixing 60 ml of concentrated hydrochloric acid with 540 ml of a saturated solution of potassium bromide. With this solution a good polish is obtained using current densities greater than 1 A/cm^2. A current density of 3 A/cm^2 was often used. With these large power inputs, care had to be taken to keep the solution from heating excessively. At current densities less than 1 A/cm^2, preferential etching takes place.

Using a copper etching jig designed to minimize stress on the sample, the samples were reduced in thickness to as little as 0.05 mm. This method was checked over the whole range of thickness by going through the polishing procedure between measurements but passing no current through the etching cell. The data were found to be quite reproducible through such checking, implying that no deformation took place during this procedure.

B. Cryogenic Technique

The cryostat was a Pyrex double Dewar with an extra independently sealed Pyrex tube within the pumpable helium space. This design permitted three methods of temperature control. With a moderate pressure of helium gas in the inner tube the sample could be held at the temperature of the liquid. With less transfer gas, temperatures to 10°K could be held by heating the copper can within which the sample was situated. The introduction of 50 cc of activated charcoal as an adsorptive heat sink permitted measurement to 77°K. Carbon-resistance thermometry was used.

C. The Measurement of Small Signals

The measurement of microvolt and submicrovolt dc signals at low temperatures is made difficult by the long leads and large thermal gradients present in any cryogenic system. One method which has been successfully used is the conversion of the dc signal to ac in the cryostat by means of a superconducting chopper.[17,18,9] The signal can then be put through a transformer and increased by orders of magnitude before being brought out and amplified. However, since the temperature range desired precluded a superconducting circuit, the apparatus was constructed using annealed copper with low thermal emf junctions and appropriate heat sinking to minimize the effects of these gradients.

The sample holder incorporated several design features which were aimed at accomplishing the above. First, all soldered connections outside of the low-temperature zone were made with a Sn-Cd alloy which has a low thermoelectric potential against copper.[19] Secondly, the head was L shaped in order to keep the connector with its solder joints away from cold gas currents. From the connector in the head, No. 22 annealed copper wires are brought down to a copper block situated in a region of near 77°K temperature. Connections to the finer (No. 40) wires which go to the sample were made within the copper block and electrically insulated with thin plastic tubing.

Connection to the measuring circuit was made through a 48-pin hermetically sealed connector. This connector went directly into an aluminum box magnetically shielded with three layers of Co-netic (trade name of Perfection Mica Co.) foil. Outside of this box an effort was made to minimize the effective pickup loop area. Within this box the necessary switching was done

[15] P. J. Price, IBM J. Res. Develop. 4, 152 (1960).
[16] L. Schubinkov and W. J. de Haas, Leiden Commun. No. 207c (unpublished).
[17] I. M. Templeton, J. Sci. Instr. 32, 314 (1955).
[18] A. R. de Vroomen and C. van Baarle, Physica 23, 785 (1957).
[19] B. K. Smith, Electron. Design 7, 52 (1959).

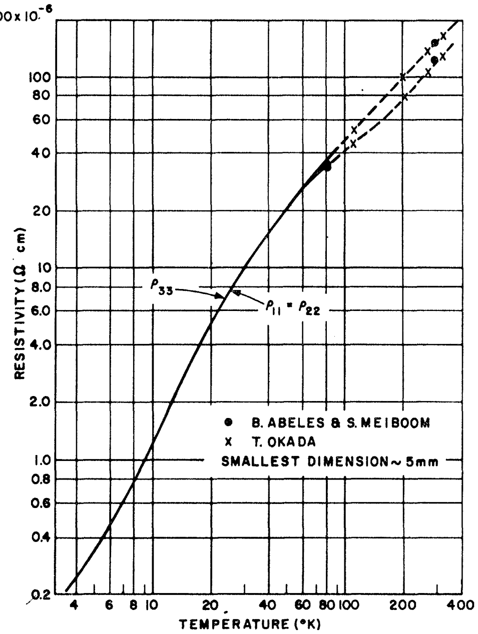

FIG. 1. Resistivity, ρ_{11} and ρ_{33}, 4.2–77°K. Smallest sample dimension ~5 mm.

with a Leeds and Northrup low-thermal multipole switch. Two heavy copper leads went from this box to a Kiethley model 150 microvoltammeter which has a sensitivity of 10^{-6} V full scale on the lowest scale with a 10^6 Ω input impedance. In this instrument the output which drives the meter movement is 10 V full scale on every scale and is available at binding posts in the rear for driving a recorder. The output signal was taken from the recorder output and read on a $\frac{1}{2}\%$ meter. To counter the effects of drifts, measurements were made by reversing the direction of current flow and averaging. Using this method, measurements were made which were reproducible to within 3×10^{-9} V.

D. Low-Field Hall Measurements

The magnetic field for the galvanomagnetic measurements was generated by a Helmholtz coil capable of producing 200 Oe. In order to keep the Hall-effect measurements in the low-field region at the lowest temperature, these measurements were made at ~1.5 Oe. The current was measured on a $\frac{1}{2}\%$ multirange meter.

With the low magnetic fields used here for measurement, the effect of the earth's field must be considered. At the sample the background field was measured to have a vertical component of −0.7 Oe and a horizontal component of 0.5 Oe. By suitable orientation of the sample, the effects of this field could be made negligibly small. The sample was mounted so that the current flowed vertically, and the horizontal component was parallel to the Hall contacts. In addition to this care in orientation, use was made of the fact that the measurements were being done in the linear Hall-effect region. The measuring field was turned on and off. Then the difference in observed voltage was just the part of the voltage which depended on the applied magnetic field (i.e., the Hall voltage). To prevent the self-magnetic field (i.e., the magnetic field generated by the current flow) from affecting the measurements, the sample current was kept below 0.1 A.

III. RESULTS

The results of measurements of bulk resistivity between 4.2 and 77°K are shown in Fig. 1. ρ_{33} was measured in one sample and $\rho_{11}=\rho_{22}$ was measured in three samples with the current in various directions in the trigonal plane. The cross section of these samples was 0.5 cm×1 cm. The solid lines represent a total of 34 measurements at 19 temperatures. The data of Okada[20] and Abeles and Meiboom[21] above 77°K is included for comparison with these results. As can be seen the resistivity ratio between 300 and 4.2°K is ~485 and the resistivity has not yet reached the impurity or boundary-limited region. The resistivities of these samples were measured to be approximately a factor of 2 lower at 2°K but at this temperature boundary scattering and impurity scattering are becoming appreciable.

In Fig. 2 are shown the results of measurements of $R_\perp$ (or $\rho_{13,2}$), the large Hall tensor component, taken with the magnetic field perpendicular to the trigonal axis. Data on five different samples are shown for comparison, also a result of Zitter[9] at 4.2°K.

The measurement of $\rho_{12,3}$ (H parallel to the trigonal axis) presents greater difficulties since the observed signal is much smaller. Also, greater precision is required in the positioning of the contacts to prevent mixing in the larger galvanomagnetic coefficients. These difficulties prevented an accurate determination of this constant. The data allow the placing of outer limits of $\pm2\times10^{-8}$ Ω cm/G over the whole temperature range.

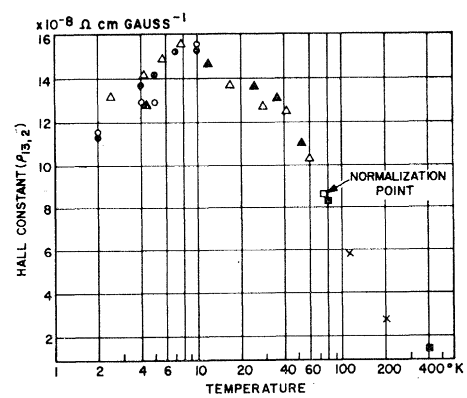

FIG. 2. Large Hall constant, $\rho_{13,2}$.

[20] T. Okada, J. Phys. Soc. Japan 12, 1327 (1957).
[21] G. E. Smith (private communication).

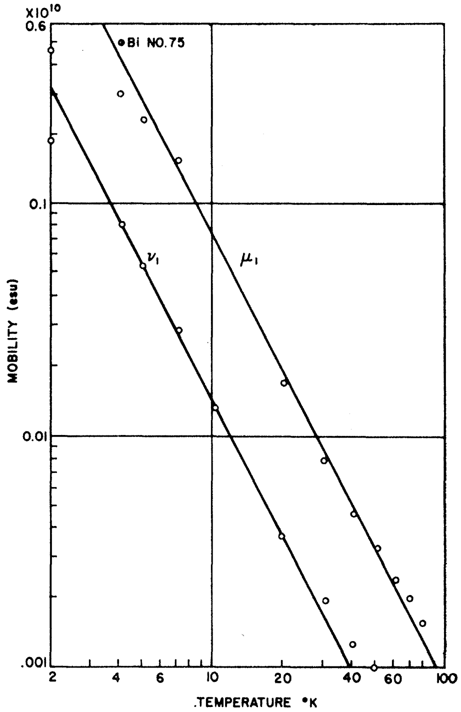

FIG. 3. Temperature variation of electron and hole mobilities for the higher conductivity sample at its greatest thickness.

The principal data taken in this study are measurements of the size dependence of ρ_{11} and $\rho_{13,2}$ taken at five temperatures between 2 and 10°K. These data were taken on two samples, one with $\varphi=0.8°$ (for the $\varphi=0.8°$ and $\varphi=23.8°$ samples, the current direction lies in the trigonal plane and 0.8° and 23.8°, respectively, from a binary axis. The trigonal axis in both samples is perpendicular to the thin direction.) and the other with $\varphi=23.8°$ while the sample thickness was varied from $\sim$0.1 to 4.0 mm. Before taking these data, the samples were cycled between 300 and 77°K several times. Between runs they were stored in liquid nitrogen to prevent any effects of annealing at room temperature.

IV. ANALYSIS

Analysis of the size and temperature dependence of the Hall constant and conductivity allows the separation of the electron and hole contribution and the evaluation of hole and electron mobilities. A detailed analysis of the transport process in the Bi Fermi-surface geometry[9,21] leads to the following expressions for the two components of the Hall tensor in terms of the partial mobilities, μ_i.[21]

$$\rho_{13,2}\equiv R_\perp=\frac{1}{Nec}\frac{\frac{1}{2}(\mu_1+\mu_2)\mu_3-\mu_4{}^2-\nu_1\nu_3}{[\frac{1}{2}(\mu_1+\mu_2)+\nu_1](\mu_3+\nu_3)}, \quad (7)$$

$$\rho_{12,3}\equiv R_{\parallel}=\frac{1}{Nec}\frac{\mu_1\mu_2-\nu_1{}^2}{[\frac{1}{2}(\mu_1+\mu_2)+\nu_1]^2}. \quad (8)$$

Since $\mu_2\ll\mu_1$, $\mu_4{}^2\ll\mu_1\mu_3$, and $\nu_1\nu_3\ll\mu_1\mu_3$ we obtain from (7),

$$R_\perp\approx\frac{1}{Nec}\frac{1}{1+2\nu_1/\mu_1}. \quad (9)$$

In addition,

$$\frac{\sigma_\perp}{\sigma_\perp{}^0}=\frac{\frac{1}{2}\mu_1+\nu_1}{\frac{1}{2}\mu_1{}^0+\nu_1{}^0}=\frac{\frac{1}{2}+\nu_1/\mu_1}{\frac{1}{2}+\nu_1{}^0/\mu_1{}^0}\frac{\mu_1}{\mu_1{}^0} \quad (10)$$

since

$$\sigma_\perp=\tfrac{1}{2}(\mu_1+\mu_2)+\nu_1 \quad (11)$$

and

$$\nu_1/\nu_1{}^0\equiv[(\nu_1/\mu_1)\mu_1]/[(\nu_1{}^0/\mu_1{}^0)\mu_1{}^0]. \quad (12)$$

In Eqs. (7)–(12) $\sigma_\perp$, μ_1, and ν_1 represent the conductivity $\perp$ to the trigonal axis, electron mobility in the binary direction (in that ellipsoid for which that binary direction coincides with a major axis), and the hole mobility $\perp$ to the trigonal axis, respectively. The superscript zero signifies the value for the greatest thickness measured.

The temperature dependence of conductivity (Fig. 1) and Hall constant $R_\perp$ (Fig. 2) can now be used to calculate the temperature dependence of the electron and hole mobilities in the binary direction. This is shown in Fig. 3, which reveals the as yet unexplained $\approx T^{-2}$ dependence of electron and hole mobilities in the low-temperature region. Above 20°K, the mobilities begin to approach the slower dependence characteristic of

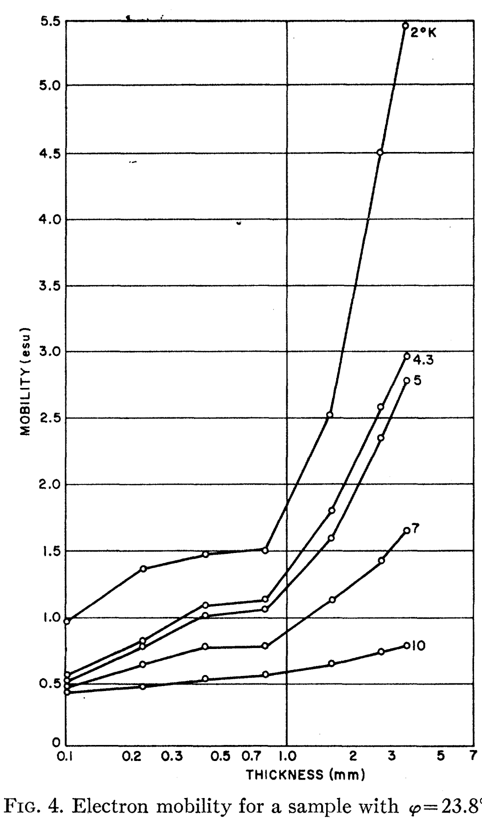

FIG. 4. Electron mobility for a sample with $\varphi=23.8°$.

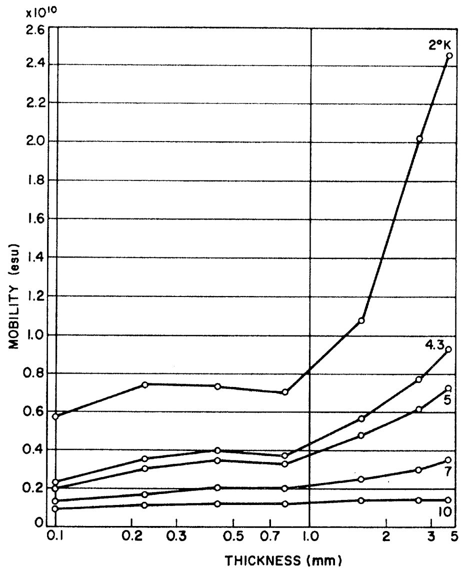

FIG. 5. Hole mobility for a sample with $\varphi=23.8°$.

metals at high temperatures. Included in the analysis is the temperature dependence of the carrier concentration computed by Smith,[21] employing the nonellipsoidal nonparabolic model.

The point marked No. 75 was taken on the highest conductivity sample measured ($\rho_{300°}/\rho_{4.2°}=632$). $R_\perp$ for this sample was not measured. The electron mobility shown was calculated using the T^{-2} hole mobility shown in the figure. The former lies within experimental error, on the T^{-2} extrapolation of the electron-mobility curve. From the foregoing, one infers that ≈ 630 is the largest attainable resistivity ratio between 300 and 4.2°K and that the $\approx T^{-2}$ behavior at low temperatures is characteristic of the material.

The difference between the thickness dependence of $R_\perp$ for the two samples is marked and can be understood in terms of the above analysis, which gives the thickness variation of the hole and electron mobilities. These figures (Figs. 4–7) show that for thickness greater than 0.3 mm the major difference between the two samples lies in the behavior of the electron mobility. The $\varphi=0$ sample (Bi No. 70) shows a much smaller thickness dependence of electron mobility which is understandable since $\frac{2}{3}$ of the bulk-electron conductivity is carried by the electrons in the ellipsoid whose binary axis is in the current direction. These electrons have a much shorter mean free path in the thin direction than those in the other two ellipsoids. In addition, this is the ellipsoid which has no size effect for pure specular scattering.[22] The mean free path of these electrons is also shorter than those of the electrons which contribute 97% of the electron conductivity in the $\varphi=24°$ sample. The similarity between the hole-mobility behavior of the two samples is to be expected because of the rotational symmetry of the hole band about the trigonal axis. (Remember that the two samples differ only by a rotation of 24° about the trigonal axis.) The sudden drop in the hole mobility of sample No. 70 below 0.3 mm is unexplained.

When one tries to fit the observations of R_{11}, the small Hall constant, into this framework, one runs into trouble if τ is assumed isotropic. Putting the mobilities at 2°K into Eq. (8) with $\mu_2=10^{-2}\,\mu_1$ (isotropic τ) one gets $R_{11}=0.17(Nec)^{-1}$. This coefficient, however, was measured to be no greater than $0.1(Nec)^{-1}$. In order to bring the calculated R_{11} within the experimental observation, τ in the bisectrix direction must be 10 times τ in the other major directions. This fact is explained as follows:

Isotropic τ implies that an electron traveling in one direction has an equal probability of being scattered to

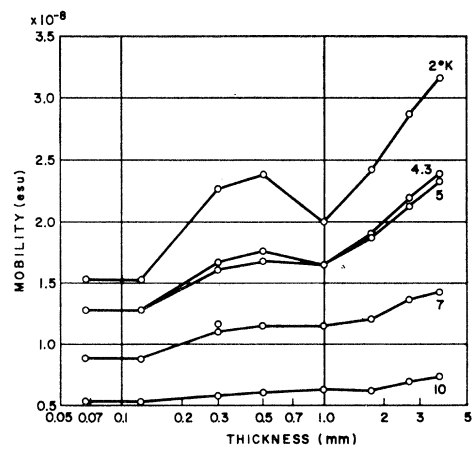

FIG. 6. Electron mobility for a sample with $\varphi=0.8°$.

any point on the Fermi surface. However, in momentum space the Fermi surface is greatly elongated in the high mass, bisectrix direction. For an electron traveling in this direction to be scattered across the Fermi surface a phonon whose energy is approximately 50°K is required. As the temperature goes below 50°K there are fewer 50°K phonons available for large angle scattering. Therefore, the high-mass direction electrons retain more memory of their state before scattering. This effect is expressible in terms of a longer mean free time. This is the same effect as was observed by Koenig, Brown, and Schilliger in n-type germanium[23] and it is, indeed, the effect which leads to the T^{-5} variation of the electrical conductivity of metals at low temperatures. In the low-mass directions the solid angle available to the electrons for scattering is much greater than that available to the high-mass electron down to quite low temperatures. Since the mass ratio is ~ 100, the temperature of a phonon which will scatter an electron across the Fermi surface in this direction is 5.4°K. Zitter's[9] observation of isotropic τ at 4°K is presumably due to

[22] A. N. Friedman and S. H. Koenig, IBM J. Res. Develop. **4**, 158 (1960).

[23] S. H. Koenig, R. D. Brown, III, and W. Schillinger, Phys. Rev. **128**, 1668 (1962).

the fact that his samples were in the residual resistance region. (S-wave scattering would be expected from such centers as neutral impurities, localized defects, and well-shielded ionized impurities which are small compared to the carrier wavelength.)

Now that the electron and hole size effects have been separated, the problem of characterizing the carrier surface scattering can be considered. Preliminary results of this investigation[22] indicated that the surface scattering was mostly specular. (The specular nature of the scattering is thought to be due to the large de Broglie wavelength of both electrons and holes.) These results were compared to Price's[15] model of specular scattering by carriers in ellipsoidal energy bands.

Independent evidence of specular reflection was obtained in the following experiment: A $\varphi=0$ sample was electropolished well down into the size-effect region. One side was then sand blasted with an S. S. White dental cutter in order to change the character of the carrier reflection from the surface. This side was then electropolished again and the cycle repeated three more times. Each time the sand blasting resulted in an increase of resistivity and the subsequent electropolishing resulted in a decrease of resistivity.

One possible explanation of these results is the production of a high-resistivity layer at the sand-blasted surface. For a zeroth approximation consider that no current flows through this layer at 4.2° but full current flows at 77°. (The resistivity of this layer can be 10

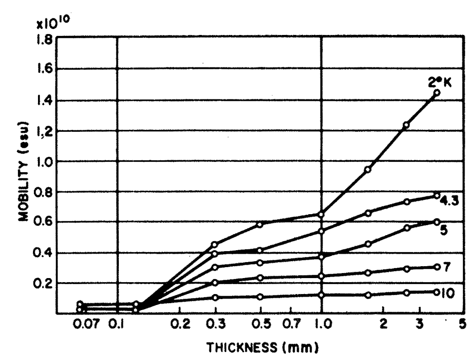

FIG. 7. Hole mobility for a sample with $\varphi=0.8°$.

times the resistivity of the remainder at 4.2°K and still be only 25% greater at 77°K.) Since the sample thickness is calculated from the 77°K potential measurement, it will be correct. If this thickness is used to calculate the 4.2°K resistivity, the latter will be high since the sample is effectively thinner. When the high-resistivity layer is electropolished away, the true resistivity will be calculated. Thus, the resistivity ratio will be higher after the polish.

To eliminate this possibility, another sample was prepared and the air pressure in the cutting unit was reduced in order to minimize the amount of material which had to be removed by the polish. In this case, with constant sample current, the potential difference between the resistance measuring electrodes was reduced after the polish showing that the resistivity increase was not a result of the above effect but that the sand blast had, indeed, caused an increase in the proportion of diffuse reflection at the surface.

The sand blasting also caused an increase in $R_\perp$. Analysis of the changes in $R_\perp$ and σ_1 of the first sample along the lines developed above shows that both the electron and hole conductivities are effected by the sand blast and both carriers are, at least in part, specularly scattered from an electropolished surface. The change of electron mobility was $\approx 5\%$ while the change of hole mobility was $\approx 10\%$.

The task remaining is the fitting of Figs. 4–7 to the size-effect theories and extraction of the dependence of mean free path and proportion of specular surface scattering on temperature. Many characteristics of the mobility behavior seem to be reasonable products of size-effect considerations. However, other characteristics make it impossible to fit these data to any surface-scattering theory.

Before one tries to put in numbers the curves look quite reasonable. The proportional change in resistivity with thickness increases at lower temperatures, as one would expect, as the mean free path increases and boundary scattering becomes more important. As was mentioned above, the size dependence of the hole mobility is similar for the two samples, as it should be because of the rotational symmetry of the hole band around the trigonal axis, while the electron mobilities have a markedly different behavior. In addition, the mobility curves of the lower mobility (shorter mean free path) sample have smaller slope at 3.5 mm thickness.

The most obvious disagreement between these data and size-effect theory is that, even though the mobility, thus the mean free path, changes by as much as a factor of 17.5 between 2 and 10°K, the curves of each set seems to inflect at the same thickness. However, the

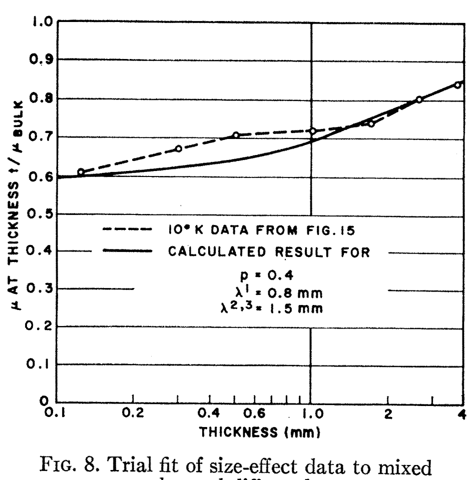

FIG. 8. Trial fit of size-effect data to mixed specular and diffuse theory.

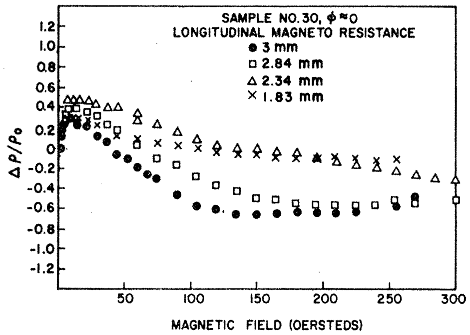

FIG. 9. Longitudinal magnetoresistance at 4.2°K.

only theoretical parameter is the ratio between thickness and mean free path, so that the curve shapes should scale with thickness.

A more detailed comparison between theory and experiment shows that the change of mobility with thickness is much more rapid in the region of major change than that predicted by any of the pre-existing theories or the one developed in the Appendix. Even more disturbing than this is the magnitude of the mean free paths calculated from the measured mobilities and the known Fermi velocities. As can be seen in Table I, these mean free paths are much smaller than those one would deduce from the shape of the data plots. This is illustrated in Fig. 8, which is an attempt to get a best fit of these data to the mixed-scattering theory. One can see that mean free paths many times longer than those calculated from mobilities were required to get any sort of fit to the data. The fit finally obtained did not reproduce the structure of the data curve. In particular the data curve has zero slope at 3.5 mm which, in theory, occurs at a thickness of several times the mean free path representing a size effect near zero. As can be seen, in order to fit the rest of the curve a size effect of 12% at 3.5 mm had to be assumed.

The question remains at this point whether the inability to fit the data with theory is due to a deficiency of theory or to experimental difficulty. In particular, is the wide discrepancy between the observed and calculated mean free paths due to an underestimate of the true mobilities calculated from conductivity, on the one hand, or an overestimate of the mean free path deduced from the mobility versus thickness data on the other? The former can be eliminated quickly since the calculated τ are of the same order as those observed by Galt[3] in cyclotron resonance and are a reasonable extrapolation from the τ obtained by Mase, von Molnar, and Lawson[24] at 20°K. The latter involves the consideration of another mechanism to explain the observed mobility variation.

TABLE I. Mean free paths (mfp) of Bi No. 70.

Sample	Temp °K	Electron mfp in binary direction λ_1^e (mm)	Hole mfp in binary direction λ_1^h (mm)
Bi No. 70	2	0.67	0.64
$\varphi=0.8°$	10	0.11	0.076

One mechanism which was extensively investigated is the possibility that strains were introduced into the sample during each thickness reduction, thus providing an apparent size effect. Three facts argue against this mechanism. Several times during the course of measurement of these samples and others the entire etching procedure was carried through without any current being passed through the electropolishing cell. No change in resistivity was observed due to this procedure over the range of sample thicknesses used. In addition, accidental processes of this sort would presumably lead to much more erratic behavior than that observed. Finally, one might expect such resistivity increases to obey Matthiessen's rule. That is to say, the increase in resistivity caused by each damage cycle should be independent of temperature. This was not found.

Other mechanisms such as phonon drag have been explored and discarded, leaving the problem unresolved. Another discrepancy of a similar sort has been found in bismuth and antimony. Figure 9 shows a measurement of the longitudinal magnetoresistance (i.e., magnetic field parallel to the current direction) in bismuth at 4.2°K. The same initial increase then decrease of resistivity with applied field has been observed by Babiskin[25] in bismuth and by Steele[26] in antimony. The initial increase is normal magnetoresistance and the subsequent decrease is believed to be due to the diameter of the carrier orbits becoming smaller than the distance between scattering surfaces. It was found, however, that in both the present experiment and the previous ones the orbit diameter at the resistivity maximum was much smaller than the sample dimension. In this case the explanation presented was the presence of internal slip planes which were the relevant scattering surfaces. This mechanism, however, has not been successful in resolving the problem at hand.

V. CONCLUSIONS

The mean-free-path concept has been a useful tool in the understanding of transport phenomena and should be directly applicable to the case at hand. However, the data here presented cannot be explained by any known theory which includes the mean free path as the internal parameter. The fact that the partial mobilities scale in temperature with sample thickness rather than the ratio of thickness to mean free path implies that, in this case, mean free path is not the internal "yardstick." This thickness scaling takes place while the mobility (and presumably the mean free path)

[24] S. Mase, S. von Molnar, and A. W. Lawson, Phys. Rev. **127**, 1030 (1962).
[25] J. Babiskin, Phys. Rev. **107**, 981 (1957).
[26] M. C. Steele, Phys. Rev. **97**, 1720 (1955).

vary by greater than an order of magnitude between 2 and 10°K.

Several other interesting things have been established. Firstly, the scattering time of electrons has been shown to be highly anisotropic, being of the order of ten times longer in the bisectrix direction than in the binary direction. Secondly, both electron and hole mobilities have been shown to have an approximate T^{-2} dependence down to 4.2°K. Thirdly, it has been established that scattering from an electropolished bismuth surface is partially specular.

ACKNOWLEDGMENTS

The author is grateful to Professor S. H. Koenig for his stimulation and advice during this work and for the interest and comments of Dr. P. J. Price during the preparation of the Appendix. He is also indebted to W. Schillinger and many other associates at this laboratory for their technical advice during his initiation into experimental physics.

APPENDIX

Price[15] has used his "vector mean free path" concept to treat the specular reflection of carriers in ellipsoidal energy bands. This technique allows the calculation of conductivity without calculation of the distribution of carriers. The main assumptions of the model used are: (1) "The direction of the change in crystal momentum on reflection is normal to the surface." (2) The magnitude of the change is given by energy conservation. This was shown pictorially in Fig. 1 of Ref. 15, which is reproduced here, Fig. 10. In the figure, **p** and **p**′ are the points in momentum space corresponding to the carrier state before and after reflection, respectively, the particle velocity being normal to the constant-energy surface.

The model used here to extend Price's calculation to include diffuse reflection is similar to that used by Fuchs[12] in his treatment of this problem for spherical energy bands. It is assumed that each time a carrier is scattered from a surface there is a probability $(1-p)$ that the carrier will be diffusely reflected. That is to say, there is a probability $(1-p)$ that the expectation value of the carrier velocity will be zero after reflection.

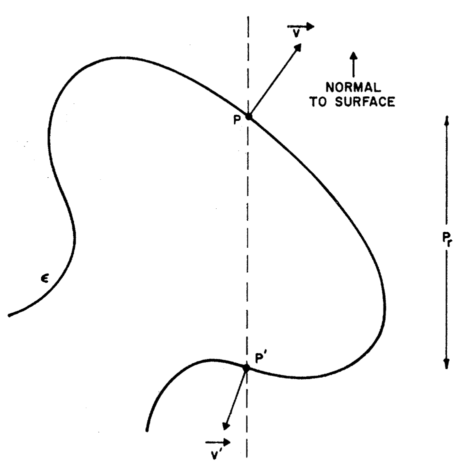

FIG. 10. Change of crystal momentum upon reflection of a carrier in the neighborhood of an ellipsoidal band edge.

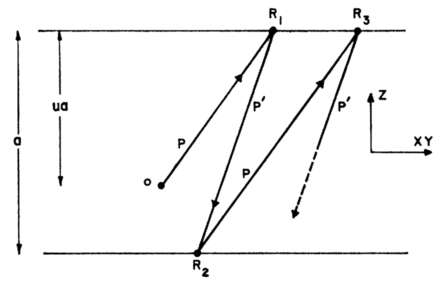

FIG. 11. Path of a carrier from start of observation to next scattering.

Following Price's notation, the vector mean free path is given by

$$\mathbf{l}=\int_0^\infty \mathbf{V}'|t)dt. \tag{A1}$$

Here $\mathbf{V}|t)$ stands for the expectation of the velocity $\mathbf{V}$ of a carrier after an interval t since it was in a definite specified state (of which $\mathbf{l}$ is then a function). The components of the conductivity tensor are then given by

$$A\boldsymbol{\sigma}=e^2/h^3kTIf_0(1-f_0)(A\mathbf{l})\mathbf{v}. \tag{A2}$$

Here

$$I\equiv\int d^3p \tag{A3}$$

over the Brillouin zone,

$$A\equiv(V)^{-1}\int d^3r \tag{A4}$$

and $f_0(\epsilon)$ is the Fermi-distribution function.

A carrier in the neighborhood of an ellipsoidal band edge, such as that shown in Fig. 10, will follow a path as shown in Fig. 11 until diffusely scattered at a surface or in the bulk. The initial state, then, is one with spatial position $z=ua$ and momentum **p**. If the bulk scattering can be characterized by a relaxation time $\tau(\epsilon)$, then the probability that a carrier will be scattered at or later than t_1 is $\exp(-t_1/\tau)$. The mean free path of a carrier initially in the given state is, summing over the segments of the path,

$$\mathbf{l}=\int_0^{u\tau\Phi}\mathbf{V}e^{-t/\tau}dt+p\int_{u\tau\Phi}^{u\tau\Phi+\tau\Phi'}\mathbf{V}'e^{-t/\tau}dt+p^2\int\cdots. \tag{A5}$$

Here

$$\Phi\equiv a/\tau|V_z| \tag{A6}$$

and

$$\Phi' \equiv a/\tau|V_z'|. \tag{A7}$$

Therefore,

$$\mathbf{l} = \tau\mathbf{V} + \frac{\{\tau\mathbf{V}(p^2e^{-\Phi'}-1)+\tau\mathbf{V}'p(1-e^{-\Phi'})\}}{1-p^2e^{-(\Phi+\Phi')}}e^{-u\Phi}. \tag{A8}$$

Now the volume average of $\mathbf{l}$ reduces to

$$A\mathbf{l} = \int_0^1 \mathbf{l}du. \tag{A9}$$

Then

$$A\mathbf{l} = \tau\mathbf{V} + \frac{1-e^{-\Phi}}{\Phi} \times \frac{\{\tau\mathbf{V}(p^2e^{-\Phi'}-1)+\tau\mathbf{V}'p(1-e^{-\Phi'})\}}{1-p^2e^{-(\Phi+\Phi')}}. \tag{A10}$$

To apply this in the neighborhood of an ellipsoidal band edge we note that the carrier energy is

$$\epsilon = \tfrac{1}{2}\mathbf{p}\cdot\boldsymbol{\gamma}\cdot\mathbf{p}, \tag{A11}$$

where $\boldsymbol{\gamma}$ is the inverse mass tensor. The carrier velocity is

$$\mathbf{V} = \partial\epsilon/\partial\mathbf{p} = \boldsymbol{\gamma}\cdot\mathbf{p}. \tag{A12}$$

Then

$$2(\epsilon'-\epsilon) = (\mathbf{p}'-\mathbf{p})\cdot(\mathbf{V}'+\mathbf{V}). \tag{A13}$$

Remembering the two conditions imposed on the momentum, which can be written

$$\mathbf{p}'-\mathbf{p} = P_r\mathbf{k} \tag{A14}$$

and

$$\epsilon(\mathbf{p}')-\epsilon(\mathbf{p}) = 0, \tag{A15}$$

we have

$$0 = P_r\mathbf{k}\cdot(\mathbf{V}'+\mathbf{V}). \tag{A16}$$

Therefore, the z compound of velocity is reversed upon specular reflection, or

$$V_z' = -V_z. \tag{A17}$$

Combining (A14) and (A17) with the expression

$$\mathbf{V}'-\mathbf{V} = \boldsymbol{\gamma}\cdot(\mathbf{p}'-\mathbf{p}), \tag{A18}$$

we get

$$p_z'-p_z = -2v_z/\gamma_{zz}, \tag{A19}$$

$$v_x'-v_x = -2(\gamma_{xz}/\gamma_{zz})v_z, \tag{A20}$$

and

$$v_y'-v_y = -2(\gamma_{yz}/\gamma_{zz})v_z. \tag{A21}$$

From (A17) we see that $\Phi'=\Phi$, and (A10) reads

$$A\mathbf{l} = \tau\mathbf{V} + \frac{1-e^{-\Phi}}{\Phi} \times \frac{\{\tau\mathbf{V}(p^2e^{-\Phi}-1)+\tau\mathbf{V}'p(1-e^{-\Phi})\}}{1-p^2e^{-2\Phi}}. \tag{A22}$$

If the x direction is taken as the direction of current flow, (A20) can be used to evaluate σ_x. Putting (A20) into (A22) we get

$$Al_x = \tau v_x + \frac{1-e^{-\Phi}}{\Phi} \times \frac{\{\tau v_x(p^2e^{-\Phi}-1)+\tau(v_x-2(\gamma_{xz}/\gamma_{zz})v_z)p(1-e^{-\Phi})\}}{1-p^2e^{-2\Phi}} \tag{A23}$$

$$= \tau v_x - \Xi(\Phi)\tau v_x - \Psi(\Phi)2\tau(\gamma_{xz}/\gamma_{zz})v_z, \tag{A24}$$

where

$$\Xi(\Phi) \equiv (1-p)(1-e^{-\Phi})/\Phi(1+pe^{-\Phi}) \tag{A25}$$

and

$$\Psi(\Phi) \equiv (p/\Phi)(1-e^{-\Phi})^2/(1-p^2e^{-2\Phi}). \tag{A26}$$

It is shown in Price's paper that

$$\sigma_{pq} \propto \langle Al_pv_q\rangle_{\text{av}}; \tag{A27}$$

therefore

$$\sigma_{xx} \propto \langle\tau v_x^2\rangle_{\text{av}} - \langle\tau\Xi(\Phi)v_x^2\rangle_{\text{av}} - \langle 2\tau(\gamma_{xz}/\gamma_{zz})\Psi(\Phi)v_zv_x\rangle_{\text{av}}. \tag{A28}$$

The ellipsoidal geometry can now be used to transform the average over an energy surface to integrals over v_z. We expand

$$v_Pv_q = (\gamma_{px}\gamma_{qx}p_x^2+\gamma_{px}\gamma_{qy}p_xp_y+\cdots) + (\gamma_{px}\gamma_{qz}p_xp_z+\cdots)+\gamma_{pz}\gamma_{qz}p_z^2 = A+B\lambda^2, \tag{A29}$$

where

$$\lambda = v_z(2\gamma_{zz}\epsilon)^{-1/2} \tag{A30}$$

when $p_x=p_y=0$, $p_z=\pm(2\epsilon/\gamma_{zz})^{1/2}$ and $\lambda=1$. Therefore, from (A29),

$$\begin{aligned} v_pv_q &= \gamma_{pz}\gamma_{qz}2\epsilon/\gamma_{zz} \\ &= A+B. \end{aligned} \tag{A31}$$

Since the system of ellipsoids $\epsilon=\frac{1}{2}\mathbf{p}\cdot\mathbf{p}$ is obtained from the system of spheres $\epsilon=\text{const}\times p^2$ by a linear transformation in $\mathbf{p}$ space, the average $\langle g(v_z)\rangle$ of any function g of electron state may be expressed as an integral over v_z with a constant weighting factor;

$$\langle g(v_z)\rangle_{\text{av}} = \frac{1}{2}\int_{-1}^{1} g(v_z)d\lambda. \tag{A32}$$

Therefore,

$$\begin{aligned} \langle v_pv_q\rangle &= \frac{1}{2}\int_{-1}^{1}(A+B\lambda^2)d\lambda \\ &= A+\tfrac{1}{3}B \\ &= \tfrac{2}{3}\epsilon\gamma. \end{aligned} \tag{A33}$$

Using (A31) and (A32) we evaluate A and B and find

$$v_p v_q = \tfrac{2}{3}\langle v_p v_q\rangle_{\mathrm{av}} \times\left\{\left(1-\frac{\gamma_{pz}\gamma_{qz}}{\gamma_{zz}\gamma_{pq}}\right)-\left(1-3\frac{\gamma_{pz}\gamma_{qz}}{\gamma_{zz}\gamma_{pq}}\right)\lambda^2\right\}. \quad \text{(A34)}$$

v_q must also be replaced by the appropriate average over each ellipse of constant v_z. This average must be proportional to v_z. Therefore,

$$v_q \to v_z\langle v_q v_z\rangle_{\mathrm{av}}/(v_z^2). \quad \text{(A35)}$$

The first term of (A28) is the bulk conductivity. The second term is evaluated using (A32) and (A34).

$$\langle\tau\Xi(\Phi)v_x^2\rangle_{\mathrm{av}} = \tau(1-p)k_z^{-1}\epsilon\gamma_{xx}\int_0^1 \frac{1-\exp(-k_z/\lambda)}{1+p\exp(-k_2/\lambda)} \times[(1-\omega_{xx})-\lambda^2(1-3\omega_{xx})]\lambda d\lambda, \quad \text{(A36)}$$

where

$$\omega_{pq}\equiv\gamma_{pz}\gamma_{qz}/(\gamma_{pq}\gamma_{zz}) \text{ and } k_z\equiv a[\tau(2\gamma_{zz}\epsilon)^{1/2}]^{-1}. \quad \text{(A37)}$$

The third term is evaluated using (A32) and (A35).

$$\langle\Psi v_z v_x\rangle_{\mathrm{av}} = 2pk_z^{-1}\epsilon\gamma_{qz} \times\int_0^1 \frac{[1-\exp(-k_z/\lambda)]^2}{1-p^2\exp(-2k_z/\lambda)}\lambda^3 d\lambda. \quad \text{(A38)}$$

In summary, then,

$$\begin{aligned}\sigma_{xx} = (l^2/h^3kT)\tfrac{2}{3}\tau\epsilon\gamma_{xx} \times\Bigg[&1-\tfrac{3}{2}(1-p)k_z^{-1}(1-\omega_{xx})\int_0^1 \frac{1-\exp(-k_z/\lambda)}{1+p\exp(1-k_z/\lambda)}\lambda d\lambda \\ &+\tfrac{3}{2}(1-p)k_z^{-1}(1-3\omega_{xx})\int_0^1 \frac{1-\exp(-k_z/\lambda)}{1+p\exp(-k_z/\lambda)}\lambda^3 d\lambda \\ &-6pk_z^{-1}\omega_{xx}\int_0^1 \frac{[1-\exp(-k_z/\lambda)]^2}{1-p^2\exp(-2k_z/\lambda)}\lambda^3 d\lambda\Bigg], \quad \text{(A39)}\end{aligned}$$

where $(e^2/h^3kT)\frac{2}{3}\tau\epsilon\gamma_{xx}=\sigma_{xx}^0$, the bulk value. This result behaves properly in the several possible limits. In the limit $p=1$, (A39) reduces to the Price result for specular reflection [Eq. (21) of Ref. 25]. In the limit $p=0$ it reduces to the Price result for diffuse reflection [Eq. (26) of Ref. 13], which in turn reduces to the Fuchs[12] result for spherical energy surfaces, $\omega_{pq}=0$. Since the third integral of the above result is not an analytic function of p at $p=1$, the whole expression must be expanded in powers of p and the behavior of the conductivity checked as the sample thickness is reduced to zero, $k_z\to 0$.

For the purpose of carrying out this calculation, let

$$X\equiv\int_0^1 \frac{1-\exp(-k_z/\lambda)}{1+p\exp(-k_z/\lambda)}\lambda d\lambda, \quad \text{(A40)}$$

$$Y\equiv\int_0^1 \frac{1-\exp(-k_z/\lambda)}{1+p\exp(-k_z/\lambda)}\lambda^3 d\lambda, \quad \text{(A41)}$$

and

$$Z\equiv\int_0^1 \frac{[1-\exp(-k_z/\lambda)]^2}{1-p^2\exp(-2k_z/\lambda)}\lambda^3 d\lambda. \quad \text{(A42)}$$

Expanding in powers of p, we obtain

$$X=\left\{\int_0^1 \lambda d\lambda-\int_0^1 a\lambda d\lambda\right\}-p\left\{\int_0^1 a\lambda d\lambda-\int_0^1 a^2\lambda d\lambda\right\} +p^2\left\{\int_0^1 a^2\lambda d\lambda-\int_0^1 a^3\lambda d\lambda\right\}-p^3\cdots, \quad \text{(A43)}$$

$$Y=\left\{\int_0^1 \lambda d\lambda-\int_0^1 a\lambda d\lambda\right\}-p\left\{\int_0^1 a\lambda^3 d\lambda-\int_0^1 a^2\lambda^3 d\lambda\right\} +p^2\left\{\int_0^1 a^2\lambda^3 d\lambda-\int_0^1 a^3\lambda^3 d\lambda\right\}-p^3\cdots, \quad \text{(A44)}$$

and

$$Z=\left\{\int_0^1 \lambda^3 d\lambda-2\int_0^1 a\lambda^3 d\lambda+\int_0^1 a^2\lambda^3 d\lambda\right\} +p^2\left\{\int_0^1 a^2\lambda^3 d\lambda-2\int_0^1 a^3\lambda^3 d\lambda+\int_0^1 a^4\lambda^3 d\lambda\right\}+\cdots, \quad \text{(A45)}$$

where

$$a\equiv\exp(-k_z/\lambda). \quad \text{(A46)}$$

In the integrals involved in (A43)–(A45) the substitution $s=1/\lambda$ is made and the resulting expressions integrated by parts until reduced to $\int_{hk_z}^{\infty} e^{-u}du/u$, which in the limit $k_z\to 0$ approaches

$$\ln(1/k_z)+A+O(k_z)+\cdots. \quad \text{(A47)}$$

When the results of these partial integrations are substituted into (A43)–(A45) and then into (A39) it is found that the coefficient of each power of p approaches const$\times k_z\ln(1/k_z)$ as $k_z\to 0$, which is the required behavior.