THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 96, NO. 5 DECEMBER 1, 1954

Theory of the Thermoelectric Power of Semiconductors

CONYERS HERRING *
Institute for Advanced study, Princeton, New Jersey and Bell Telephone Laboratories, Murray Hill, New Jersey
(Received June 29, 1954)

The usual theory of thermoelectric power fails to account for the marked rise in this quantity which has been found recently for some semiconductors as the temperature is lowered below room temperature. This paper develops the recently suggested explanation that the thermoelectric power Q is the sum of the usual electronic term Q_e, resulting from the spontaneous tendency of the charge carriers to diffuse from hot to cold, and a phonon term Q_p, resulting from the drag on the charge carriers exerted by the phonons streaming from hot to cold in thermal conduction. An equivalent description of the term Q_p can be given in terms of the contribution to the Peltier heat flux in an isothermal specimen, due to phonons dragged along by the electric current. As a prelude to the discussion of Q_p, the existing theory of Q_e is first subjected to some refinements required by recent developments in semiconductor theory. The theory of Q_p is then formulated in a simple but general way by making use of an approximate proportionality between heat current and crystal momentum in the phonon system. Using recently-derived results on the probability that a very low-frequency phonon will be scattered by other phonons, an explicit expression for $Q_p(T)$ is derived, which should be valid in the range of moderately low temperatures and low carrier concentrations. At lower temperatures, but still far above the range where the thermal conductivity is appreciably size-dependent, Q_p is dominated by the scattering of phonons from the boundaries of the specimen; the theory of this effect is worked out in detail. Although Q_p is independent of carrier concentration when the latter is low, a considerable decrease is predicted at high carrier concentrations, or at very low temperatures, because of a saturation effect. The effect of Fermi degeneracy on all these phenomena is discussed. Available data on p germanium show all these effects and can be fitted by the theory. The comparison indicates that a large proportion of the low-temperature lattice scattering of holes in p germanium is by shear modes. Although n germanium seems less suited for quantitative comparison, it, also, shows all the predicted effects.

I. INTRODUCTION

THE thermoelectric powers of a number of single crystal specimens of germanium have recently been measured by Frederikse[1] and by Geballe and Hull[2] over the range from room temperature to that of liquid hydrogen. Both workers find, for specimens of high to moderate resistivity, a marked rise in the numerical value of the thermoelectric power at low temperatures, a rise quite incomprehensible in the framework of the usual theory[3–7] of thermoelectric power. A similar effect has been observed by Mansfield and Salam[8] for single crystals of molybdenite. The explanation, suggested independently by Frederikse[1] and the author,[9] seems to lie in the role played by the thermal vibrations of the lattice. In the presence of a temperature gadient these are not isotropic; in fact, the phenomenon of thermal conduction implies that the waves travel preferentially from hot to cold. The scattering of charge carriers by the phonons is therefore not random, but such as to push the carriers more often toward the cold end than the reverse. Zero current results only when the cold end acquires enough of an excess of carriers so that their electrostatic field counterbalances the combined effect of this phonon drag and the normal tendency of carriers to diffuse from hot to cold.[10] Thus the observed thermo-

* Present address: Bell Telephone Laboratories, Murray Hill, New Jersey.

[1] H. P. R. Frederikse, Phys. Rev. **91**, 491 (1953); **92**, 248 (1953).

[2] T. H. Geballe, Phys. Rev. **92**, 857 (1953); T. H. Geballe and G. W. Hull, Phys. Rev. **94**, 1134 (1954).

[3] C. Wagner, Z. physik. Chem. **B22**, 195 (1933).

[4] R. H. Fowler, Proc. Roy. Soc. (London) **A140**, 505 (1933).

[5] V. A. Johnson and K. Lark-Horovitz, Phys. Rev. **69**, 259 (1946); **92**, 226 (1953).

[6] J. Yamashita, J. Phys. Soc. Japan **4**, 310 (1949); G. Lautz, Z. Naturforsch. **8a**, 361 (1952); O. Madelung and H. Welker, Z. angew. Physik **5**, 12 (1953).

[7] R. W. Wright, Proc. Phys. Soc. (London) **A64**, 984 (1951).

[8] R. Mansfield and S. A. Salam, Proc. Phys. Soc. (London) **B66**, 377 (1953).

[9] C. Herring, Phys. Rev. **92**, 857 (1953).

[10] While this manuscript was being prepared, Dr. H. P. R. Frederikse and Dr. R. H. Parmenter independently called to my attention two papers of L. Gurevich, J. Phys. (U.S.S.R.) **9**, 477 (1945); **10**, 67 (1946), which develop the theory of this phonon drag effect in a metal. Gurevich does not discuss the effect for semiconductors, although he mentions that it should be present.

electric power Q can be expressed as the sum of an electronic part Q_e, given by previous theories, and a phonon part Q_p.

The explanation just given has been formulated in terms of the differential Seebeck voltage Q; it may therefore be called the Q approach. A different, though necessarily equivalent, approach is through the Peltier heat Π, which is related thermodynamically to Q. The high Q values at low temperatures mean that the energy flux accompanying a given electric current in an isothermal semiconductor is much greater than that which the electronic charge carriers of the usual theory would transport. The proposed explanation is that in the presence of a current the scattering of the charge carriers by the lattice vibrations tends to increase the amplitudes of lattice waves traveling in the same direction as the carriers and to decrease the amplitudes of waves moving in the opposite direction. This results in a net transport of energy by the waves in the direction of motion of the carriers. This approach to the phenomenon may be called the Π approach. Since the Kelvin relation, Eq. (1) below, must apply in all cases, these two approaches must give equivalent results, and this can be confirmed in detail using specific models such as the one of Secs. III to V below. However, in this paper we shall use only the Π approach. Frederikse[1] has treated the subject from the standpoint of the Q approach.

A few qualitative features of the suggested interpretation are obvious at once, and their agreement with observation gives one confidence to explore the proposed effect in detail. The mutual drag of charge carriers and phonons should obviously increase the magnitude of Q, or Π, regardless of sign, i.e., for either n- or p-type material. Moreover, as long as one sign of carrier predominates over the other the effect on Q or Π should clearly be independent of the concentration n of carriers, at least if n is sufficiently low. Finally, the effect should be more pronounced at low than at high temperatures since at high temperatures the lattice vibrations are more rapidly restored to randomness by the effects of anharmonicity of the interatomic forces, i.e., by phonon-phonon collisions. All these predictions agree with observation.

The object of the present paper is to develop the theory of the phonon effect quantitatively, and to see what basic information on semiconductors can be derived from a comparison of the results with experiment. Since the attack on this problem has to be based on the theoretical concepts available in the present-day theory of semiconductors, it is pertinent to begin by examining the reliability of these concepts. A general survey of the relation of theory and experiment in regard to electronic transport shows that for impurity semiconductors one needs to distinguish three or four regions in the range of possible impurity concentrations. When the impurity (and dislocation) concentration is very low, nearly all the charge carriers move as isolated entities in regions of perfect crystalline material, and their motion can be rigorously described in terms of the concepts of band theory, at least if the mobility is high enough to justify treating the interation with the lattice vibrations as a small perturabation on the motion of the carriers. As the impurity concentration is increased it becomes necessary to take account of impurity scattering and related effects. While this can be done with present-day theory, the available treatment is not altogether satisfactory. At still higher impurity concentrations, available theory becomes completely unreliable for quantitative predictions This is because of what can be crudely described as a jumping of charge carriers from one impurity atom to the next, an "impurity band" conduction which minges and competes with normal conduction.[11] The onset of Fermi degeneracy normally occurs in this range. When the impurity and carrier concentrations get very igh, available theoretical concepts start to work a little better again, as one can treat the carriers as a degenerate Fermi gas and treat the impurities as merely sources of scattering. However, in this region present theory uses many unjustified simplifications. Moreover, it is by no means clear that one should use here the same effective mass parameters, etc., as in the region of very ow impurity concentration.

The exposition given in the following sections reflects the different positions of theory in the several regions just described, and the success achieved by the present theory in these various regions turns out, as expected, to be in proportion to the reliability just estimated for present concepts in these same regions. No attempt is made to give a detailed discussion of thermoelectric power in the region of impurity band conduction, although it will be clear that both the electronic and phonon contributions will usually have the same order of magnitude in this range as in the neighboring regions of lower and higher impurity concentration. Section II reviews and modernizes the theory of the electronic contribution Q_e for low to moderate impurity concentrations. Section III is devoted to a formulation of the phonon effect problem in terms of relaxation times and rates of transfer of lattice momentum, a formulation applicable in principle to all ranges of impurity concentration. Sections IV and V give the detailed solution for small concentrations, neglecting the effects of scattering of phonons from the boundaries of the specimen. The latter effects, which make Q_p size-dependent at low temperatures, are taken up in Sec. VI. Section VII discusses, for concentrations below the impurity band range, a saturation effect which decreases the phonon contribution when the carrier concentration becomes sufficiently high or the

[11] K. Shifrin, J. Phys. (U.S.S.R.) **8**, 242 (1944); H. M. James and A. S. Ginzbarg, Phys. Rev. **77**, 749 (1950); J. Phys. Chem. **57**, 840 (1953); C. Erginsoy, Phys. Rev. **80**, 1104 (1950); **88**, 893 (1952); C. S. Hung and J. R. Gliessman, Phys. Rev. **79**, 726 (1950) C. S. Hung, Phys. Rev. **79**, 727 (1950); G. W. Castellan and F. Seitz, *Semiconducting Materials*, H. K. Henisch, Editor, (Butterworths Scientific Publications, London, 1951), p. 8; W. Baltensperger, Phil. Mag. **44**, 1355 (1953).

temperature sufficiently low. The region of Fermi degeneracy is treated briefly in Sec. VIII, where both the electronic and phonon contributions are discussed. All detailed comparisons of the various parts of the theory with experiment are relegated to Sec. IX, as many of the significant conclusions of this paper come from a comparison of different regions of temperature and concentration.

II. ELECTRONIC THEORY *VERSUS* EXPERIMENT

Let Q be the absolute thermoelectric power of any substance and let Π be its absolute Peltier coefficient, defined as the energy flux per unit current when the energy zero is taken at the Fermi level, or more properly, the electrochemical potential of the electrons. At any absolute temperature T these quantities are known to obey the relation

$$Q=\Pi/T, \tag{1}$$

first obtained (in difference form) by Lord Kelvin[12] in 1854. The relation in Eq. (1) is now known to be a consequence of Onsager's principle of microscopic reversibility.[13] For the sake of later applications it is worth emphasizing that the derivation of (1) remains valid for the apparent Q and Π of a rod-shaped specimen, even when the rod is of such small diameter that boundary effects cause this apparent Q and Π to differ from the values which would be measured on a very large specimen of identical material.

As Wagner[3] and others have pointed out, Eq. (1) provides an easy way of calculating the thermoelectric power of any model of a metal or semiconductor for which the solution of the Boltzmann equation for isothermal electrical conduction is known: one need merely calculate Π from this solution. The familiar theory of thermoelectric power, derived from the Lorentz-Sommerfeld theory of conduction in metals, considers only transport by electronic carriers. For a nondegenerate semiconductor with a single sign of carrier predominating this gives the electronic contribution

$$e\Pi_e=\epsilon_F-\epsilon_b-\Delta\epsilon_T, \tag{2}$$

where ϵ_F is the Fermi level, ϵ_b is the energy of the edge of the band (valence or conduction) in which the carriers move, and $\Delta\epsilon_T$, the average energy of the transported electrons relative to the band edge, is a quantity of the order of kT and therefore usually rather smaller than $\epsilon_F-\epsilon_b$. From this alone it is fairly obvious that no reasonable assumptions regarding band structure and scattering processes will suffice to bring (2) into agreement with the sharp rise of Q at low temperatures which has been found experimentally for germanium. Nevertheless, we shall discuss the terms of (2) in some detail, partly to reinforce this point, and partly because it will be necessary later on to have as accurate an expression as possible for the electronic contribution to Q.

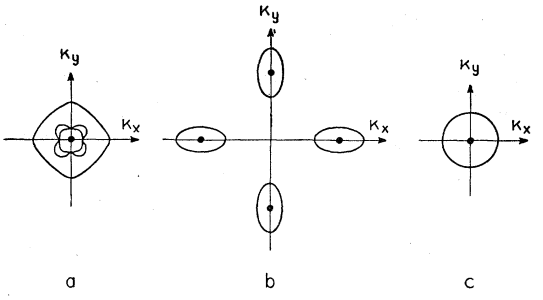

FIG. 1. Schematic forms of the locus in $\mathbf{K}$ space of states with a given value of energy, near the band edge value, for three possible cases: *a*, single-valley, many-sheeted case; *b*, many-valley, single-sheeted case; *c*, single-valley, single-sheeted case, the model most often assumed in the literature. The band edge points $\mathbf{K}_i$, at which the energy is an extremum, are identified by heavy dots.

Consider first the term $\epsilon_F-\epsilon_b$ in Eq. (2). In order to evaluate this from measurements of Hall coefficient or density of carriers one must make some assumption about the effective mass of the carriers. Now it is known[14] that the electronic energy band structures of silicon and germanium are more complicated than has usually been assumed in semiconductor theory, although the exact structures have not yet been established conclusively. Some of the models which are likely to occur for germanium, silicon, and similar semiconductors are illustrated in Fig. 1. One possibility is that the band edge is a triply degenerate state with wave vector $\mathbf{K}=0$, slightly split by spin-orbit coupling. For this case the locus in $\mathbf{K}$-space of states of a single energy near the band edge consists of three concentric surfaces of (in general) complicated shapes [Fig. 1(a)]. Another possibility, illustrated in Fig. 1(b) is that the band edge consists of a number of states with different wave vectors $\mathbf{K}_\nu\neq 0$, symmetrically related to one another; the locus of a given energy consists of a number of identical ellipsoids centered around the points $\mathbf{K}_\nu$. A third possibility, not shown in the picture, is that there are a number of such band edge points with a twofold degeneracy at each $\mathbf{K}_\nu$, slightly lifted by spin-orbit coupling; the energy locus is a family of identical pairs of complex surfaces. If one attempts to use, for these substances, formulas of the conventional model [Fig. 1(c)] which contain effective mass as a parameter, it turns out that different formuals require different effective masses, since they involve different types of averages over the parameters describing the multivalued nonspherical energy surfaces. For example, we may define a "density-of-states mass" $m^{(N)}$ by

$$N(\epsilon)=(4\pi/h^3)(2m^{(N)})^{\frac{3}{2}}|\epsilon-\epsilon_b|^{\frac{1}{2}},$$

[12] W. Thomson, *Mathematical and Physical Papers* (Cambridge University Press, London, 1882), Vol. I, p. 232.

[13] L. Onsager, Phys. Rev. **37**, 405 (1931); **38**, 2265 (1931); H. B. Callen, Phys. Rev. **73**, 1349 (1948). For a detailed discussion of the Kelvin relations and full bibliography see S. R. de Groot, *Thermodynamics of Irreversible Processes* (Interscience Publishers, Inc., New York, 1951), Chap. VIII.

[14] W. Shockley, Phys. Rev. **78**, 173 (1950); F. Herman and J. Callaway, Phys. Rev. **89**, 518 (1953).

where $N(\epsilon)d\epsilon$ is the number of levels per unit volume in the range ϵ to $\epsilon+d\epsilon$. This $m^{(N)}$ equals the inertial effective mass m^* for the conventional model; if each surface of constant ϵ is a single ellipsoid, $m^{(N)}$ equals the geometric mean of the three principal effective masses m_i^*; if each surface of constant ϵ is a set of N_v similar ellipsoids centered on the cubic axes of the Brillouin zone, as in Fig. 1(b), $N(\epsilon)$ is N_v times as great as for one ellipsoid, so $m^{(N)}$ must be $N_v^{\frac{2}{3}}$ times the geometric mean, etc. A different type of effective mass measurement comes from the inertial effect on the high frequency dielectric constant, measured by Benedict and Shockley.[15] The inertial mass $m^{(I)}$ which this effect measures may be defined as the average, over all the carriers, of the ratio of force applied to rate of change of crystal momentum. For ellipsoidal energy surfaces, whether single or multiple, $m^{(I)-1}$ is the arithmetic mean of the curvatures m_i^{*-1} of the plots of ϵ_K against wave number K in the principal directions. We therefore have $m^{(N)}>m^{(I)}$ for any multiellipsoid model, both because of the relation of arithmetic and geometric means and because of the multiplicity of the ellipsoids. In fact, it can be proved that $m^{(N)}>m^{(I)}$ whatever the shape or multiplicity of the constant energy surfaces near the band edge, e.g., the nonellipsoidal shapes suggested by Shockley.[14] However, Eq. (3), and with it the concept of $m^{(N)}$, is only useful for a many-sheeted model like Fig. 1(a) if the spin-orbit splitting is $\gg kT$ or $\ll kT$. If spin-orbit splittings of the order of kT are present, $m^{(N)}$ as we have defined it may vary considerably with energy over the thermal range.

In terms of $m^{(N)}$ the usual formula[16] gives

$$|\epsilon_F-\epsilon_b|/kT=\ln[2(2\pi m^{(N)}kT)^{\frac{3}{2}}/h^3n], \tag{3}$$

where n is the density of carriers. Combining this with (1) and (2), we have for the electronic contribution Q_e to the thermoelectric power, in microvolts per degree,

$$Q_e=\mp 86.2\left[\ln\frac{4.70\times10^{15}}{n}+\tfrac{3}{2}\ln\frac{m^{(N)}}{m}+\frac{|\Delta\epsilon_T|}{kT}+\tfrac{3}{2}\ln T\right], \tag{4}$$

where n is in cm^{-3} and where the upper sign is for n type, the lower for p. This is the familiar formula;[5,6] we have merely elucidated its applicability to complex band structures. Specifically, Eq. (4) applies to all such except many-sheeted cases where the spin-orbit splitting is of the order of kT, and even for these Eq. (4) may be used, though less conveniently, by letting $m^{(N)}$ be a temperature-dependent quantity defined by Eq. (3) and measuring $\Delta\epsilon_T$ from the lowest band-edge state.

Consider now how $\Delta\epsilon_T/kT$ is determined. We shall make the common assumption, briefly discussed at the start of Appendix A, that the processes responsible for scattering the charge carriers can be described by a relaxation time $\tau_e(\mathbf{K})$; we allow τ_e to depend on position $\mathbf{K}$ in wave-number space. Then the first-order perturbation $f^{(1)}$ of the electron distribution function $f(\mathbf{K})$ by and electric field $\mathbf{E}$ obeys

$$f^{(1)}\propto\tau_e\mathbf{E}\cdot\nabla_K f^{(0)}\propto\tau_e\mathbf{E}\cdot\mathbf{v}_K f^{(0)}, \tag{5}$$

where $f^{(0)}$ is the unperturbed distribution function for the electrons or holes and $\mathbf{v}_K$ is the group velocity of a carrier in the $\mathbf{K}$th state. From (5) it follows that, for a cubic crystal

$$\Delta\epsilon_T=\langle v^2(\epsilon-\epsilon_b)\tau_e\rangle/\langle v^2\tau_e\rangle, \tag{6}$$

where the angular brackets mean averages over $f^{(0)}$. Note that Eqs. (5) and (6) involve no assumption regarding the dependence of ϵ on $\mathbf{K}$, except that of cubic symmetry; they are therefore valid for all the types of degenerate or "many-valley" bands, even in the presence of spin-orbit splitting.

An important class of cases for which Eq. (6) can be evaluated is that for which $\tau_e\propto|\epsilon-\epsilon_b|^r$ for any given direction of motion of the carrier within any one of the sub-bands or valleys; r is supposed independent of direction, but the constant of proportionality need not be. For many-sheeted models this can of course only be a good approximation if the spin-orbit splitting is $\gg kT$ or $\ll kT$. Lattice scattering by phonons of long wavelength corresponds to $r=-\frac{1}{2}$, while the extreme case of Conwell-Weisskopf scattering[17] by ionized impurities corresponds to $r=\frac{3}{2}$. For this class of cases Eq. (6) gives

$$|\Delta\epsilon_T/kT|=(5/2)+r. \tag{7}$$

A good estimate of the effective value of r which approximates conditions in any given specimen can be obtained from the ratio of Hall mobility μ_H to drift mobility μ, since for the conventional model of Fig. 1(c)[18]

$$\frac{\mu_H}{\mu}=\frac{\langle v^2\tau_e^2\rangle\langle v^2\rangle}{\langle v^2\tau_e\rangle^2}=\frac{\Gamma(5/2)\Gamma(5/2+2r)}{[\Gamma(5/2+r)]^2}. \tag{8}$$

In addition, the sign of r must be known, since a given μ_H/μ occurs for two values of r, one <0 and one >0. Although the relation (8) requires modification[19] for band structures of the types shown in Figs. 1(a) and 1(b), the modification for Fig. 1(b) can be shown to be slight unless the anisotropy of the ellipsoids is extreme. Although the variation of τ_e with ϵ may sometimes depart considerably from power-law form, a number of calculations which have been made for different τ_e's have given values of $\Delta\epsilon_T/kT$ which, for given μ_H/μ,

[15] T. S. Benedict and W. Shockley, Phys. Rev. **89**, 1152 (1953); T. S. Benedict, Phys. Rev. **91**, 1565 (1953).

[16] See, for example, W. Shockley, *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, Inc., New York, 1950), p. 464.

[17] E. Conwell and V. F. Weisskopf, Phys. Rev. **77**, 388 (1950); H. Brooks, Phys. Rev. **83**, 879 (1951).

[18] Reference 16, p. 280.

[19] C. Herring (to be published).

have usually been within about 0.2 unit of the values obtained by eliminating r between the right-hand expressions (7) and (8). The table recently published by Johnson and Lark-Horovitz,[5] based on the usual combination of lattice with Conwell-Weisskopf scattering, also gives values within this limit.

The actual range of values of $\Delta\epsilon_T/kT$ is a little different from the range 2.0 to 4.0 which one can get by combining ordinary lattice scattering ($r=-\frac{1}{2}$) with various amounts of Conwell-Weisskopf scattering. The ratio μ_H/μ for high-purity p-type germanium is close to 1.8 at 300°K,[20,21] and it is likely[22] that this is due to τ_e decreasing more sharply with $|\epsilon-\epsilon_b|$ than corresponds to $r=-\frac{1}{2}$. A value $r=-0.8$ would fit this μ_H/μ, and for this value, $\Delta\epsilon_T/kT=1.7$. By about 100°K, μ_H/μ has decreased to a value close to the 1.18 which corresponds to $r=-\frac{1}{2}$. At the other extreme, Spitzer and Härm[23] have shown that the limiting value of $\Delta\epsilon_T/kT$ for pure ionized impurity scattering is 3.20, not 4.0, provided the number of ions equals the number of electrons. The difference between this result and that obtained from the Conwell-Weisskopf formula is due to the fact that Spitzer and Härm take electron–electron interactions into account. The value 4.0 is only to be expected if the ions greatly exceed the electrons in number, i.e., for a highly compensated specimen.

At high carrier and impurity concentrations, we may well question the legitimacy of computing the position of the Fermi level as was done for Eq. (4), and of treating the impurities merely as a source of scattering. To get an idea of the size of the error in Eq. (4) due to fuzzing of the band edge, etc., we may try to modify the relation of n to ϵ_F by the Debye-Hückel correction.[24] However, to be consistent we must also correct $\Delta\epsilon_T$ for the potential energy which the carriers transport. For the simplest case $r=0$ the effect on Π_e due to the change in ϵ_F turns out just to compensate that due to the change in $\Delta\epsilon_T$; this suggests that the correction to Eq. (4) is small under all nondegenerate conditions. When impurity band conduction becomes appreciable, however, the theory leading to Eq. (4) is of course very questionable.

The dashed line in Fig. 2 shows values of Q_e computed from Eq. (4) for a typical high-purity specimen of p-type germanium, for reasonable guesses at $m^{(N)}$ and $\Delta\epsilon_T$. For this sample n was obtained from Hall measurements with correction for the roughly known[20,21] ratio of Hall to drift mobility. The observed Q values, shown for comparison in the full curve, show a marked rise at low temperature, which occurs very much sooner than the rise in the theoretical Q_e due to depletion of the concentration of free carriers. It is clear that the observed values could be fitted to Eq. (4) only by assuming fantastic—and highly temperature-dependent—values of $m^{(N)}$ or the exponent r in Eq. (7). Provided the impurity content is not too high, all specimens both of p- and n-type germanium studied in references 1 and 2 show the same marked deviation from Eq. (4) at low temperatures. Inhomogeneity of the samples can hardly be blamed for the high Q values; not only would inhomogeneity be hard to reconcile with the consistency and reproducibility of the data but, as a consideration of simple series and parallel circuits shows, it would be likely to act in the wrong direction. We are therefore forced to consider the only other mechanism of energy transport which can contribute to Π, namely, transport by traveling elastic waves in the lattice.

III. ELEMENTARY CONSIDERATIONS ON ENERGY TRANSPORT BY PHONONS

We shall show first that there is an approximate proportionality between energy flux and crystal momentum in the phonon system. Let $N_{q\alpha}$ be the mean number of quanta of excitation present in the traveling wave normal mode of wave vector $\mathbf{q}$ and polarization α. In thermal equilibrium, with no current, this of course takes the Planck value

$$N^{(0)}(\mathbf{q},\alpha)=[\exp(\hbar\omega(\mathbf{q},\alpha)/kT)-1]^{-1}, \qquad (9)$$

where $\omega(\mathbf{q},\alpha)$ is the frequency of the mode. In the presence of a current, $N(\mathbf{q},\alpha)$ will differ from Eq. (9), but we suppose it to be still a smoothly varying function

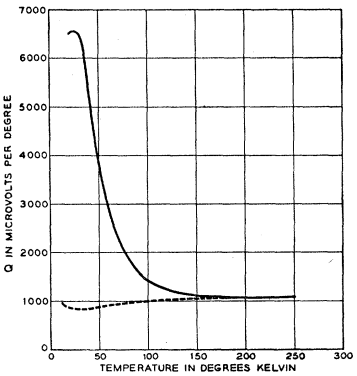

FIG. 2. Comparison of observed thermoelectric power Q with the theoretical electronic contribution. Full curve: values observed by Geballe and Hull, reference 2, for their specimen No. 7 (p-type Ge, 1.5×10^{14} excess acceptors/cc). Dashed curve: electronic contribution calculated from (4), with $m^{(N)}/m=0.6$ and with $\Delta\epsilon_T/kT$ assumed (see text) to increase from 1.7 at 250° to 2.0 an and below 100°.

[20] F. J. Morin, Phys. Rev. **93**, 62 (1954).
[21] M. B. Prince, Phys. Rev. **92**, 681 (1953).
[22] A different interpretation is indicated by recent work of Harman, Willardson, and Beer, Phys. Rev. **94**, 1065 (1954) and a forthcoming publication. According to these authors, the primary cause of the large room-temperature Hall coefficient is the parallel participation of high- and low-mass holes, as required by the two-band model of reference 32. I am indebted to Dr. Beer for supplying me with details of this work in advance of their publication.
[23] L. Spitzer, Jr., and R. Härm, Phys. Rev. **89**, 977 (1953).
[24] See for example R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (Cambridge University Press, London, 1939).

of $\mathbf{q}$. The flux density of energy flow is

$$\mathbf{j}=\sum_{\mathbf{q},\alpha}N(\mathbf{q},\alpha)\hbar\omega(\mathbf{q},\alpha)\mathbf{v}(\mathbf{q},\alpha)$$

$$=(2\pi)^{-3}\sum_{\alpha}\int N(\mathbf{q},\alpha)\hbar\omega(\mathbf{q},\alpha)\mathbf{v}(\mathbf{q},\alpha)d\mathbf{q}, \quad (10)$$

where $\mathbf{v}(\mathbf{q},\alpha)=\nabla_{\mathbf{q}}\omega(\mathbf{q},\alpha)$ is the group velocity, and where the summation is over a density of $\mathbf{q}$ values corresponding to unit volume of material. If we have to deal only with modes of long wavelength, we can set $\omega(\mathbf{q},\alpha)=c(\mathbf{q},\alpha)\mathbf{q}$, where $c(\mathbf{q},\alpha)$, the phase velocity of the mode, depends on the direction of $\mathbf{q}$ but not on its magnitude. This will certainly be the case at low temperatures, and it will in fact turn out to be the case for our problem at all temperatures. If the crystal is taken to be elastically isotropic we can make the further simplification $\mathbf{v}(\mathbf{q},\alpha)=c_\alpha\mathbf{q}/q$, and even for an anisotropic crystal this may be a fair approximation to use in Eq. (10), if suitable average sound velocities are used. To this approximation

$$\mathbf{j}\approx\sum_{\mathbf{q},\alpha}N(\mathbf{q},\alpha)(\hbar\mathbf{q})c_\alpha^2=\mathbf{P}_l c_l^2+\mathbf{P}_t c_t^2=(\mathbf{P}_l+\mathbf{P}_t)\bar{c}^2, \quad (11)$$

where the subscripts l and t refer to longitudinal and transverse waves, respectively, $\bar{c}$ is a suitable average of the two sound velocities, and

$$P_\alpha=\sum_{\mathbf{q}}N(\mathbf{q},\alpha)\hbar\mathbf{q}=(2\pi)^{-3}\int N(\mathbf{q},\alpha)\hbar\mathbf{q}d\mathbf{q} \quad (12)$$

is the total crystal momentum per unit volume of all modes of polarization type α, the summation being, as before, over a density of $\mathbf{q}$ values appropriate to unit volume of material.

The interest of the relation in Eq. (11) is, of course, that crystal momentum is conserved in collisions of phonons with the charge carriers, and also in phonon-phonon collisions, unless these are of the Umklapp type. We may therefore calculate $\mathbf{j}$ by equating the rate at which the phonon system receives crystal momentum from the electronic carriers to the rate at which it loses it via Umklapp collisions, boundary scattering, etc. Since an electric field $\mathbf{E}$ imparts crystal momentum to the carriers at a rate $\mp ne\mathbf{E}$ per unit volume, where n is the density of carriers, the former rate may be written

$$[d(\mathbf{P}_l+\mathbf{P}_t)/dt]_e=\mp fne\mathbf{E}, \quad (13)$$

where the upper sign is for electrons, the lower for holes, and where f is the fraction of the crystal momentum lost by the electrons which is delivered to the lattice vibrations ($f=1$ in the absence of impurity scattering). For the other term, let us define an effective relaxation time $\bar{\tau}$ for the phonon system by setting for the rate of loss of crystal momentum via the various types of collisions which do not conserve this quantity,

$$[d(\mathbf{P}_l+\mathbf{P}_t)/dt]_c=-(\mathbf{P}_l+\mathbf{P}_t)/\bar{\tau}. \quad (14)$$

We may now calculate $\mathbf{j}$ from Eqs. (11), (13), and (14), eliminating $\mathbf{E}$ in favor of the current density $\mathbf{J}=ne\mathbf{E}\mu$, where μ is the mobility. The result is

$$\mathbf{j}\approx\mp(\bar{c}^2 f\bar{\tau}/\mu)\mathbf{J}. \quad (15)$$

The coefficient of $\mathbf{J}$ in this expression is the phonon contribution Π_p to the Peltier heat. It may be placed in a form more suitable for comparison with the electronic contribution in Eq. (2) by writing $\mu=(e/m^*)\bar{\tau}_e$, where m^* and $\bar{\tau}_e$ are suitably defined averages of the effective mass and relaxation time of the carriers, respectively. Thus,

$$e\Pi_p\approx\mp\bar{c}^2 ef\bar{\tau}/\mu=\mp m^*\bar{c}^2(f\bar{\tau}/\bar{\tau}_e). \quad (16)$$

The factor $m^*\bar{c}^2$ is much smaller than the right of Eq. (2) for a nondegenerate semiconductor, being only of the order of 10^{-4} ev for germanium. Thus for the phonon contribution to exceed the electronic, the weighted ratio of phonon to electron or hole relaxation times must be of the order of hundreds of thousands.

A few words are in order at this point regarding the meaning of the $\bar{\tau}$ occurring in Eq. (16). We note, to begin with, that most of the crystal momentum fed in from the charge carriers to the phonon system, or at least a major part of it, is fed in to a very small fraction of the normal modes, namely, those with very small wave vectors $\mathbf{q}$, of the order of the wave vector of a thermal electron. The course of the crystal momentum given to the phonon system by the charge carriers is therefore roughly describable by the hydraulic analogy shown in Fig. 3. Crystal momentum, represented by water, is fed into an upper container, representing the modes of low q. From this it can either escape to the modes of higher q (water passing through the orifice on the right of the upper container) or be lost forever by, for example, scattering of the low-q phonons from the boundaries of the specimen (water passing through the orifice on the left). That part of the crystal momentum which reaches the modes of high q, represented by the lower container, is eventually lost by Umklapp, impurity, or boundary scattering (water passing through the orifice on the extreme right). Now because of their long wavelength the modes of small q are less susceptible to all kinds of scattering than are the rest of the modes, except

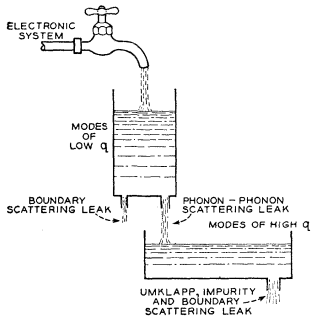

FIG. 3. Hydraulic analogue of the transfer of crystal momentum $\mathbf{P}$ from the electronic system to lattice modes of low and high wave number q, and the ultimate destruction of $\mathbf{P}$.

at very low temperatures where boundary scattering affects all modes alike. This means that both orifices leading from the inner container are narrow. It is therefore quite possible that in a steady state there will be more crystal momentum in the modes of low q than in those of high (more water in the inner container than the outer). In fact, if the scattering of low-q phonons by high-q phonons (middle leak) is sufficiently slow, the rate of annihilation of the crystal momentum of the high-q phonons (leak on extreme right) may be quite inconsequential in determining the total amount of crystal momentum (water) in the system, i.e., $\bar{\tau}$ may be determined almost entirely by the rates of the two leaks out of the low-q system.

To orient ourselves regarding the relative rates of the various leaks in Fig. 3 for the germanium experiments we are trying to intepret, we shall consider two limiting cases. In a thermal conduction experiment, with no charge carriers, the thermal gradient feeds in crystal momentum to each mode $\mathbf{q}$, α at a rate which depends only on the direction of $\mathbf{q}$ and not on its magnitude, as long as $\hbar\omega(\mathbf{q},\alpha) \lesssim kT$.[25] This corresponds to placing the containers of Fig. 3 out in the rain, instead of feeding only the inner one from a faucet. If, as we have been speculating, the leaks from the inner container are slower than that from the outer, a given rate of feeding in water will yield a steady state with less water in the containers for feeding by rain than for feeding by faucet. In other words, the effective phonon mean free time for thermal conduction, τ_c, will be less than the effective $\bar{\tau}$ to be used in Eq. (16). To see whether this is in fact the case, we may try calculating Eq. (16) using for $\bar{\tau}$ the τ_c determined from Debye's formula,[26]

$$\kappa = \tfrac{1}{3} C c^2 \tau_c, \tag{17}$$

for the thermal conductivity κ, in which C is the specific heat per unit volume, and c an average sound velocity. With the values of κ measured by Geballe[27] for fairly pure germanium, values of $Q_p = \Pi_p/T$ were computed from Eqs. (17) and (16), and are shown in the lower curve of Fig. 4, together with values of $Q - Q_e$ from Fig. 2. We have taken $f=1$, its maximum possible value, for this comparison. The suspicions outlined in the preceding paragraph are confirmed: Use of τ_c for $\bar{\tau}$ gives values of Q_p which are vastly too small.

Having thus verified that $\bar{\tau} \gg \tau_c$, we now explore the other extreme. If nothing else puts a lower limit on the loss of crystal momentum by the low-q phonons, boundary scattering will. For the specimens used in most of the work of Geballe and Hull it is reasonable to assume diffuse reflection of phonons at the boundaries of the specimen, and to take a mean free path L of the order of 0.15 cm between such reflections. With $\bar{c} \approx 5 \times 10^5$ cm/sec., we obtain $\tau_b = L/\bar{c} \approx 3 \times 10^{-7}$ sec. Using this for $\bar{\tau}$,

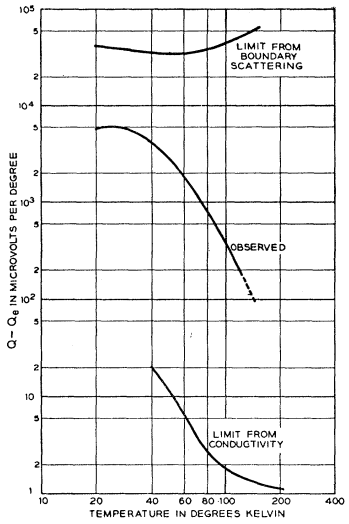

FIG. 4. Middle curve: Empirical phonon contribution Q_p to the thermoelectric power of the Ge specimen of Fig. 2, obtained by subtracting the Q_e curves from the observed Q curves. Upper curve: Approximate upper limit to Q_p, obtained from (16) with $f=1$ and a phonon mean free path of 0.15 cm, the limit set by boundary scattering. Lower curve: Lower limit to Q_p, obtained by inserting the mean free path for thermal conduction into (16). The curvature of the upper plot, which differs from the behavior of the more refined expression (42) derived below, is due to the use of the assumption $f=1$, which allows the irregular variation of the mobility to be felt.

and again taking $f=1$ for simplicity, we get the upper curve of Fig. 4. We see that at the lowest temperatures $\bar{\tau}$ does indeed seem to be starting to approach the boundary scattering value, but that over the range from say 50° to 200°K, $\bar{\tau} \ll \tau_b$.

We have thus arrived at the following picture, which we must develop quantitatively in the next two sections: In an isothermal specimen carrying a current, most of the crystal momentum of the phonon system—and therefore most of its heat current—resides in the normal modes of long wavelength. Except at very low temperatures, crystal momentum is lost from these modes principally by interactions with modes of shorter wavelength, and the relaxation time for these interactions may be used as the $\bar{\tau}$ in Eq. (16), even though these interactions do not themselves destroy any crystal momentum. At the lowest temperatures, $\bar{\tau}$ is compounded out of the relaxation time for these interactions and that for boundary scattering. At no temperature is $\bar{\tau}$ likely to be appreciably affected by scattering processes which involve only phonons of shorter wavelength than those with which the charge carriers interact; in other words, the rate of the right-hand leak in Fig. 3 is never likely to be important.

IV. RATE OF TRANSFER OF CRYSTAL MOMENTUM TO PHONONS

The picture developed in the preceding section, though valid as an over-all description, needs con-

[25] P. G. Klemens, Proc. Roy. Soc. (London) **A208**, 108 (1951).
[26] P. Debye, *Vorträge über die kinetische Theorie der Materie und der Elektrizität* (Teubner, Berlin, 1914), pp. 19–60.
[27] T. Geballe (private comminication).

siderable refinement if it is to yield specific quantitative results. We have just seen that the $\bar{\tau}$ to be used in (15) and (16) is much longer than the effective relaxation time τ_e for thermal conduction, because of the great variation of the mean free path of a phonon with its wavelength. This suggests that even among the modes of low q, with which the charge carriers interact, we should use a different relaxation time $\tau(\mathbf{q},\alpha)$ for each mode. Thus if we define $\mathbf{R}(\mathbf{q},\alpha)$ to be the crystal momentum fed into the mode $\mathbf{q}\alpha$ per unit time, and define $\mathbf{p}(\mathbf{q},\alpha)$ to be the crystal momentum of this mode, we write

$$\mathbf{p}(\mathbf{q},\alpha)-\mathbf{p}^{(0)}(\mathbf{q},\alpha)=\tau(\mathbf{q},\alpha)\mathbf{R}(\mathbf{q},\alpha), \tag{18}$$

where $\mathbf{p}^{(0)}(\mathbf{q},\alpha)=N^{(0)}(\mathbf{q},\alpha)\hbar\mathbf{q}$ is the value in equilibrium. The total crystal momentum of the phonons of branch α in unit volume is

$$\mathbf{P}_\alpha=(2\pi)^{-3}\int \mathbf{p}(\mathbf{q},\alpha)d\mathbf{q}=(2\pi)^{-3}\int \tau(\mathbf{q},\alpha)\mathbf{R}(\mathbf{q},\alpha)d\mathbf{q}. \tag{19}$$

In this section we shall calculate $\mathbf{R}(\mathbf{q},\alpha)$, and in the next one we shall combine it with $\tau(\mathbf{q},\alpha)$ and evaluate the integral in Eq. (19), which we can then substitute into Eq. (11) to get the phonon contribution to the Peltier heat flow. The next section will also include a discussion of the legitimacy of using the mean free time concept in Eq. (18).

To avoid excessive mathematical complication in the initial development of the theory, we shall make several simplifications in this section. Specifically, we shall assume Maxwellian statistics and the conventional model of Fig. 1(c). For this model only the longitudinal modes are important. A more general treatment is given in Appendix A, which assumes the many-valley model of Fig. 1(b), a model for which transverse as well as longitudinal phonons can scatter the charge carriers. This treatment leads to expressions for $\mathbf{R}(\mathbf{q},\alpha)$ and Q_p which are similar to those derived for the simpler model in this section and the next, provided the effective mass in the simple model is taken to be a suitable average of the effective masses in the three principal directions of a valley, an average which may be different for longitudinal and transverse modes. The many-sheeted model of Fig. 1(a), unfortunately, seems too complicated to treat quantitatively; however, it, too, should have properties similar to those of the simple model. In this section where we use the simple model, we shall drop the index α, since only the longitudinal branch is important. In the following section, however, where we shall derive formulas for practical application, we shall resume the distinction between longitudinal and transverse branches, in order to have equations which may be placed in correspondence with those of Appendix A.

The quantity $\mathbf{R}(\mathbf{q})$ is compounded out of the distribution function of the carriers and the scattering probability for emission or absorption of a phonon. The former is easily obtained if we assume that the combined effect of all the scattering processes acting on the charge carriers can be described by a relaxation time $\tau_e(K)$, dependent on the magnitude but not on the direction of the wave vector $\mathbf{K}$. Then an electric field in the x direction produces a first-order perturbation $f^{(1)}$ of the distribution function of the charge carriers which, by (5), is

$$f^{(1)}(\mathbf{K})\propto\tau_e(K)K_x\exp(-\hbar^2K^2/2m^*kT). \tag{20}$$

In this and the following equations all factors independent of $\mathbf{K}$ and $\mathbf{q}$ are absorbed into the proportionality sign.

For a crystal free from flaws and impurities the transition probability from state $\mathbf{K}$ to state $\mathbf{K}+\mathbf{q}$ with absorption of a longitudinal phonon $\mathbf{q}$, or to $\mathbf{K}-\mathbf{q}$ with emission, is

$$\begin{aligned}W_{a,e}(\mathbf{K}\rightarrow\mathbf{K}\pm\mathbf{q})&\propto|M|^2\delta[\epsilon(\mathbf{K})-\epsilon(\mathbf{K}\pm\mathbf{q})\pm\hbar\omega(\mathbf{q})]\\&\propto\begin{pmatrix}N(\mathbf{q})\\N(\mathbf{q})+1\end{pmatrix}q\delta[\epsilon(\mathbf{K})-\epsilon(\mathbf{K}\pm q)\pm\hbar\omega(\mathbf{q})],\end{aligned} \tag{21}$$

since the squared matrix element for the transition, $|M|^2$ is independent of $\mathbf{K}$ and depends on $\mathbf{q}$ only through a factor very nearly proportional to q and to the occupation numbers $N(\mathbf{q})$ (absorption) and $N(\mathbf{q})+1$ (emission).[28] For the purpose of this section and the next two, the occupation numbers $N(\mathbf{q})$ can be set equal to their thermal equilibrium values in Eq. (9); this means that for the present we are considering only semiconductors with such low concentrations of carriers that the unbalance of the phonon system in the presence of a current is much less than the unbalance in the distribution of the carriers. In Sec. VII we shall consider the case of such high carrier concentrations that the unbalance of the phonons is of the same order as that of the carriers.

It is generally assumed that (21) continues to hold even when there are enough impurities present to cause appreciable impurity scattering. In other words, it is generally assumed that when two or more scattering mechanisms operate, the probability for a given electron to suffer any one type of scattering is the same as it would be if all the other types were absent, so that the reciprocal relaxation times for the different mechanisms add. We shall use this assumption here, although its validity at moderately high-impurity concentrations may be questioned.

Returning to (20) and (21), we now combine these equations to get the rate $\mathbf{R}(\mathbf{q})$ of crystal momentum transfer to the mode $\mathbf{q}$. Remembering that the equilibrium carrier distribution $f^{(0)}$ contributes nothing, that $f^{(1)}$ is an odd function of $\mathbf{K}$, and that the W_a, W_e of (21) are unchanged if we reverse the directions

[28] See for example reference 16, pp. 520–531. This proportionality involves the assumption of elastic isotropy of the crystal, but deviations from isotropy are not likely to modify it greatly.

of both $\mathbf{K}$ and $\mathbf{q}$, we can write

$$\mathbf{R}(\mathbf{q})=\hbar\mathbf{q}[dN(\mathbf{q})/dt]_e\propto\mathbf{q}\int f^{(1)}\times[W_e(\mathbf{K}\to\mathbf{K}-\mathbf{q})-W_a(\mathbf{K}\to\mathbf{K}+\mathbf{q})]d\mathbf{K}$$

$$=\mathbf{q}\int f^{(1)}[W_e(\mathbf{K}\to\mathbf{K}-\mathbf{q})+W_a(\mathbf{K}\to\mathbf{K}-\mathbf{q})]d\mathbf{K}$$

$$\propto\mathbf{q}\int\tau_e(\mathbf{K})K_x\exp(-\hbar^2K^2/2m^*kT)\times\{q[N(\mathbf{q})+1]\delta(2\mathbf{K}\cdot\mathbf{q}-q^2-2K_cq)+qN(\mathbf{q})\delta(2\mathbf{K}\cdot\mathbf{q}-q^2+2K_cq)\}d\mathbf{K}, \quad (22)$$

where $K_c=m^*c/\hbar$. An approximation commonly made in the theory of semiconductors[29] is to neglect the energy change suffered by an electron when it is scattered by a phonon. In this section we shall adopt this approximation, which is equivalent to neglecting the terms $\pm2K_cq$ in the arguments of the δ functions in (22). Although the legitimacy of this must certainly fail at sufficiently low temperatures, it can be shown from a treatment using the full (22) that the results obtained with the present approximation are adequate throughout the liquid hydrogen range.[30] When K_c is neglected the two δ functions in (22) become the same. Choosing a polar coordinate system for $\mathbf{K}$ with axis in the direction of $\mathbf{q}$, we have

$$d\mathbf{K}=K^2d\varphi d(\cos\theta)dK=Kd\varphi d(\mathbf{K}\cdot\mathbf{q}/q)dK, \quad (23)$$

$$\text{Average of } K_x \text{ on } \varphi=\mathbf{K}\cdot\mathbf{q}q_x/q^2. \quad (24)$$

Thus with the equipartition value of $N(q)(\gg1)$, (22) reduces to

$$\mathbf{R}(\mathbf{q})=BJ(\mathbf{q}q_x/q)\int_{q/2}^{\infty}\tau_e(K)\times\exp(-\hbar^2K^2/2m^*kT)KdK, \quad (25)$$

where B is a constant and J is the current density; the lower limit $q/2$ represents the minimum value of K for which a θ value can be found to make the argument of the δ function vanish. The dependence of $R_x(\mathbf{q})$ on q is shown in Fig. 5 for different assumptions regarding τ_e, and also for a typical case of degenerate statistics, worked out in Appendix A.

The constant of proportionality B in Eq. (25) is independent of $\mathbf{q}$ and for given electric field $\mathbf{E}$ it is independent of the amount of impurity scattering, etc., contained in $\tau_e(K)$. We can evaluate it by equating the sum of $\mathbf{R}(\mathbf{q})$, over all $\mathbf{q}$ in the longitudinal branch in unit volume, to its known value Jf/μ, where μ is the mobility and f is the fraction of the crystal momentum lost by the charge carriers which is delivered to low-energy acoustic modes of the lattice. The sum can be simplified by interchanging the order of integration on $\mathbf{K}$ and $\mathbf{q}$; it is

$$(2\pi)^{-3}\int R_x(\mathbf{q})d\mathbf{q}=(6\pi^2)^{-1}BJ\int_0^{\infty}q^3\int_{q/2}^{\infty}\tau_e(K)\times\exp(-\hbar^2K^2/2m^*kT)KdKdq$$

$$=(2/3\pi^2)BJ\int_0^{\infty}K^5\tau_e(K)\times\exp(-\hbar^2K^2/2m^*kT)dK. \quad (26)$$

Let $\tau_{eL}(K)$ be the relaxation time for scattering by low-energy acoustic modes only, and let μ_L be the value which the mobility would have if this were the only scattering mechanism operating. Let averages over the Maxwellian distribution be denoted by angular brackets Then[31]

$$\mu=(e/m^*)\langle K^2\tau_e\rangle/\langle K^2\rangle, \quad (27)$$

$$\mu_L=(e/m^*)\langle K^2\tau_{eL}\rangle/\langle K^2\rangle, \quad (28)$$

$$f=\langle K^2\tau_e/\tau_{eL}\rangle/\langle K^2\rangle. \quad (29)$$

Since $\tau_{eL}\propto K^{-1}$ we have

$$f\mu_L=(e/m^*)\langle K\rangle\langle K^3\tau_e\rangle/\langle K^2\rangle^2, \quad (30)$$

$$f/\mu=\langle K\rangle\langle K^3\tau_e\rangle/\mu_L\langle K^2\rangle\langle K^2\tau_e\rangle. \quad (31)$$

Note that $\langle K^3\tau_e\rangle/\langle K^2\tau_e\rangle$ equals the integral in Eq. (26) divided by a similar integral with K^5 replaced by K^4. Equating Eq. (26) to J times Eq. (31) we have finally

$$B=\frac{3\pi^2}{2}\frac{\langle K\rangle}{\mu_L\langle K^2\rangle\int_0^{\infty}K^4\tau_e(K)\exp(-\hbar^2K^2/2m^*kT)dK}. \quad (32)$$

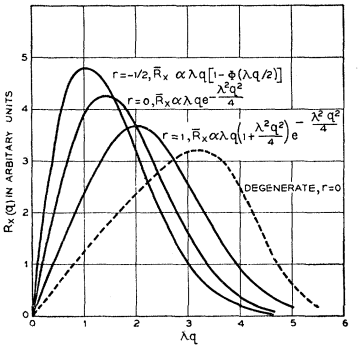

FIG. 5. Dependence of $\bar{R}_x(q)$, the spherical average of (25) or (A4), on λq, where λ is the thermal electron wavelength defined by (40). Full curves: Classical statistics, Eq. (25), for various values of the exponent r in $\tau_e\propto|\epsilon-\epsilon_b|^r$. Dashed curve: Typical case of Fermi degeneracy, Eq. (A4), with $|\epsilon_F-\epsilon_b|=4kT$, $r=0$, isotropic effective masses assumed. Areas under different portions of any curve represent relative contributions to the integral (19), hence to Q_p, if $s=0$ is assumed in (36). All curves have been arbitrarily normalized to the same area.

[29] See for example reference 16, p. 252.
[30] C. Herring (to be published).
[31] See, for example, reference 16, p. 276.

Regarding the applicability of these results to the many-valley model, the conclusions reached in Appendix A are the following: When a relaxation time dependent only on energy exists, an integral expression for the $R(\mathbf{q},\alpha)$ of each branch can be written down which is a superposition of expressions of the form in Eq. (25) with different m^*'s ranging from the highest to the lowest of the principal m^* of a single valley. Here K is to be defined by setting the $\hbar^2K^2/2m^*$ equal to the energy. We may note that if the energy surfaces in a single valley happen to be nearly spherical, Eq. (25) will be quite good as it stands, for either branch. The constant B_α going with branch α is to be evaluated by setting the right of Eq. (26) equal to Jf_α/μ, where f_α is the fraction of the crystal momentum of the carriers which is lost to branch α. The variation of f_α/μ with temperature, amount of impurity scattering, etc., is similar to that given by Eq. (31) if μ_L is replaced by the corresponding partial mobility μ_α and a constant factor of the order of unity is inserted.

In Sec. VI we shall consider temperatures so low that only boundary scattering is important for the phonons. For this case the $\tau(\mathbf{q},\alpha)$ in Eq. (19) becomes independent of the magnitude of q, and it is unnecessary to evaluate the $R(\mathbf{q},\alpha)$'s; we shall merely use Eqs. (16) and (31) with a $\bar{\tau}$ determined by the dimensions of the specimen. We shall then need to know the ratio of the quantities in angular brackets in Eq. (31); values of this ratio for several limiting cases are given in Table I.

V. PHONON CONTRIBUTION TO Q FOR SMALL CARRIER CONCENTRATIONS, WITH NEGLECT OF BOUNDARY SCATTERING

We have now to evaluate the integral in Eq. (19) for the total crystal momentum of the phonons in a current-carrying specimen. In accordance with the remarks made near the end of the preceding section, we shall use an expression of the form in Eq. (25) for the $R(\mathbf{q},\alpha)$ of each branch α of the acoustic spectrum. In the temperature range where boundary and mosaic scattering of the phonons is unimportant, we should use the law of variation of $\tau(\mathbf{q},\alpha)$ with T and q which is predicted by the theory of phonon-phonon collisions.[32] According to this theory most of the collisions suffered by phonons of very small wave number q at moderate and low temperatures

TABLE I. Values of $\mu_L f/\mu$ from (31), and of the coefficient of λ^{2+} in (39), for $\tau_e \propto K^{2r}$.

r	$f'=\mu_L f/\mu$	Coeff. λ^2/μ_α, $s=0$	Coeff. λ^3/μ_α $s=1$
−1.00	0.85	0.85	1.50
−0.75	0.93	0.74	1.00
−0.50	1.00	0.67	0.75
0	1.13	0.57	0.50
1.00	1.36	0.45	0.30
1.50	1.46	0.42	0.25

[32] C. Herring, Phys. Rev. **95**, 954 (1954).

are with phonons of energy $\sim kT$, and very few are with each other. It is therefore legitimate for our purposes to regard each such mode $\mathbf{q},\alpha$ as having a relaxation time $\tau(\mathbf{q},\alpha)$ equal to the time required for this mode to share its lattice momentum with other modes—i.e., the time constant of the middle leak of Fig. 3.

At temperatures well below the Debye temperature the theory for an elastically anisotropic cubic crystal such as germanium predicts that as $q\to 0$,

$$1/\tau_l(\mathbf{q})\sim A_l T^3 q^2, \tag{33}$$

where $\tau_l(\mathbf{q})$ is the value of $\tau(\mathbf{q},\alpha)$ when α is the highest-frequency (longitudinal) acoustic branch, and

$$1/\tau_t(\mathbf{q})\sim A_t T^4 q \tag{34}$$

for the other two (transverse) branches. The coefficients A_l, A_t may depend slightly on the direction of $\mathbf{q}$. For an elastically isotropic crystal (34) is still valid, but (33) is replaced by

$$[1/\tau_l(\mathbf{q})]_i\sim A_{li} T q^4. \tag{35}$$

A rough estimate made for germanium[31] suggests that (33) should be applicable when $\hbar\omega_q \lesssim 0.15kT$; (34) should apply over a wider range, while for a typical elastically isotropic crystal the validity of (35) would be restricted to a rather narrower range. Whether or not q is small enough for any of these asymptotic expressions to apply, it is always true that $1/\tau(\mathbf{q},\alpha)\propto T^5$ if $\mathbf{q}$ and T are scaled proportionately, provided T and $\hbar\omega(\mathbf{q})/k$ are $\ll$ the Debye temperature. A crystal with a very small elastic anisotropy should have a longitudinal relaxation time following (35) down to quite small q, and then changing over to (33).

Although we are only concerned with temperatures rather smaller than the Debye temperature, we may get some insight into the probable direction of departures from (33) and (34) in the upper part of this range by considering the extreme of very high temperatures. For this case, $1/\tau(\mathbf{q})$ for any given mode is proportional to T, the q dependence being similar to that at lower temperatures.[31]

Since we expect the longitudinal and transverse relaxation times to be of the same order when $\hbar\omega(\mathbf{q})\sim kT$, (33), (34), and (35) suggest that $\tau_l(\mathbf{q})\gg\tau_t(\mathbf{q})$ if $\hbar\omega(\mathbf{q})\ll kT$, as is the case for the phonons active in scattering the charge carriers. Thus we expect the longitudinal modes to contribute much more than the transverse ones to the thermoelectric power, over most of the temperature range. At low temperatures, however, boundary scattering reduces the longitudinal contribution much more than the transverse, so the two contributions may become comparable. We shall therefore carry through the analysis of this section in a form applicable to either branch. We shall use the adjustable assumption,

$$1/\tau(\mathbf{q},\alpha)=A_\alpha T^{3-s-\gamma} q^{2+s}. \tag{36}$$

For simplicity A_α is assumed independent of the direction of $\mathbf{q}$. To a first approximation we may set $s=\gamma=0$, as this corresponds to the ideal case of longitudinal modes, small q, and low T. In a more refined calculation we may add to this contribution Q_{pl} of the longitudinal modes a smaller contribution Q_{pt}, obtained from the formulas of this section with $s=-1$. Departure of s from the ideal value 0 for longitudinal modes may occur if the important range of q's is not sufficiently small; we expect $|s|$ to be rather smaller than unity. For a temperature range covering sizable fractions of the Debye temperature a positive γ may need to be used. Mosaic scattering may be taken into account crudely by adjusting γ and s, or in a more refined way by a theory like that of the next section.

With $\tau(\mathbf{q})$ from Eq. (36) and $\mathbf{R}(\mathbf{q})$ from Eq. (25), the integral in Eq. (19) for P_α becomes a double integral over $\mathbf{K}$ and $\mathbf{q}$. As in Eq. (26), we can interchange the order of integration and carry out the integration on $\mathbf{q}$. The result is

$$P_\alpha=\frac{2^{2-s}}{6\pi^2(2-s)}\frac{B_\alpha J}{A_\alpha T^{3-s-\gamma}}\int_0^\infty K^{3-s}\tau_e(K)\times\exp(-\hbar^2K^2/2m^*kT)dK. \quad (37)$$

We may eliminate B_α by using Eq. (32), if the simple model applies, or more generally, by using Eq. (26) equated to Jf_α/μ, where as before f_α is the fraction of the crystal momentum of the charge carriers which is delivered to phonons of branch α. The phonon contribution to the thermoelectric power is related to P_l and P_t by Eq. (11), and is of the form $Q_{pl}+Q_{pt}$; we find

$$Q_{p\alpha}=\Pi_{p\alpha}/T=\frac{\bar{c}_\alpha^2 f_\alpha}{2^s(2-s)A_\alpha T^{4-s-\gamma}\mu}\frac{\langle K^{1-s}\tau_e\rangle}{\langle K^3\tau_e\rangle}, \quad (38)$$

where as before the angular brackets denote Maxwellian averages. This applies, of course, only at temperatures high enough for boundary scattering to be unimportant. As is shown in Appendix A, Eq. (38) applies to the many-valley model if a suitable average of the effective masses in the different directions is used in relating K to the energy, i.e., in Eqs. (39) and (40) below.

If for the moment τ_e is assumed to follow the same law of variation with K at all temperatures, the ratio of angular brackets in Eq. (38) is proportional to $T^{-1-\frac{1}{2}s}$ and f_α/μ goes as $T^{\frac{3}{2}}$, so $Q_{p\alpha}$ varies as $T^{-\eta}$ with $\eta=(7/2)-\frac{1}{2}s-\gamma$. Though our derivation has assumed the simple model of Fig. 1(c), or, in Appendix A, the many-valley model of Fig. 1(b), it is not hard to see that this law of temperature variation has a general validity. In particular, it applies to the many-sheeted model of Fig. 1(a), provided the splitting introduced into this model by spin-orbit coupling is either $\gg kT$ or $\ll kT$. For if the scattering probability for transition of a carrier from state $\mathbf{K}$ to state $\mathbf{K}'$ depends only on $\mathbf{K}/T^{\frac{1}{2}}$, $\mathbf{K}'/T^{\frac{1}{2}}$, the function $\mathbf{R}(\mathbf{q},\alpha)$ for one temperature will be derivable from $\mathbf{R}(\mathbf{q},\alpha)$ for any other temperature by a simple scale change, the q scale going as $T^{\frac{1}{2}}$. Thus the $\bar{\tau}$ of Eq. (16), which is an average of the $\tau(\mathbf{q},\alpha)$ of Eq. (36) with weights proportional to $\mathbf{R}(\mathbf{q},\alpha)$ will be given by Eq. (36) with an effective q proportional to $T^{\frac{1}{2}}$. This leads to the same exponent η as Eq. (38).

Since ideally $s=\gamma=0$, the ideal value of η is 7/2. This accords fairly well with the observation of Geballe and Hull[2] that for p germanium Q_p goes nearly as $T^{-3.2}$. However, there are a number of refinements which need to be made in an accurate comparison of theory and experiment, of which we shall consider one here and others in the following sections, and it is therefore best to defer a detailed comparison until Sec. IX. The refinement we shall consider here arises from variations in the way τ_e depends on K. These variations, which affect both f_α/μ and the last fraction in Eq. (38), may arise from impurity scattering, or, for the many-valley model of Fig. 5(b), because the importance of intervalley scattering decreases with decreasing T. If we approximate the behavior of $\tau_e(K)$ at any temperature by K^{2r}, both these effects cause the effective exponent $2r$ to increase with decreasing T. For such a law we have, when Eq. (31) is valid,

$$\frac{f_\alpha}{\mu}\frac{\langle K^{1-s}\tau_e\rangle}{\langle K^3\tau_e\rangle}=\frac{\Gamma(2+r-\frac{1}{2}s)}{\Gamma(5/2)\Gamma(5/2+r)}\cdot\frac{\lambda^{2+s}}{\mu_\alpha}, \quad (39)$$

where

$$\lambda=(\hbar^2/2m^*kT)^{\frac{1}{2}} \quad (40)$$

is a thermal electron wavelength, and $\mu_\alpha\propto T^{-\frac{3}{2}}$ is the mobility which would occur if the only scattering of the carriers were that by phonons of branch α. The principal correction required for the general many-valley model is, according to Appendix A, the insertion of a constant factor independent of T and of the behavior of $\tau_e(K)$; this might be absorbed into the definition of μ_α, and does not affect the temperature dependence of Eq. (39).

If $s>-1$, the value of Eq. (39) decreases as r increases, as is shown in Table I. Thus the effect of intervalley scattering, or of impurity scattering if the assumption of Eq. (21) is valid, is to make the exponent η in $Q_pT^{-\eta}$ smaller than $(7/2)-\frac{1}{2}s-\gamma$. This is what one expects physically. The greater the amount of impurity scattering, or the less the amount of intervalley scattering, the larger will be the effective average K of the carriers which contribute most to the current, hence the larger will be the effective average q of the modes to which most of the crystal momentum is delivered. Although for given current this means that more crystal momentum will be delivered to the low-energy longitudinal phonons, as shown by the $\mu_l f/\mu$ values of Table I, it also means that the average $\tau(\mathbf{q})$ of the modes to which it is delivered will be less, and the latter effect outweighs the former because of the rapidity of the variation of $\tau(\mathbf{q})$ with q. Use of a more realistic law than K^{2r} for the dependence of τ_e on K turns out not to change this conclusion qualitatively. However, it is

conceivable that when impurity scattering is important the increase in the delivery of crystal momentum to the phonons may be greater than one would compute by the present procedure of merely using an altered distribution of electron velocities with the same expression in (21) for the lattice scattering probability. If this occurs the increase in delivery may compensate or outweigh the decrease in the average $\tau(\mathbf{q})$.

We shall try to make some rough quantitative estimates of the effect of temperature variation of the carrier scattering law on the temperature variation of Q_p. Consider first the effect, in pure p-type germanium, of intervalley scattering, if the model of Fig. 1(b) applies, or of optical mode scattering, if Fig. 1(a) is applicable[33] with spin-orbit splitting $\gg kT$. We have seen in Sec. II that this can be described roughly by letting r vary from -0.8 at 300°K to -0.5 at 100°K. For such a variation we find from (39) and Table I that the effective temperature exponent η for Q_p over this range should fall below 7/2 by 0.1 or 0.15 if $s=0$, or that it should fall below 3 by 0.3 or more if $s=1$. Similar crude estimates of r and the lowering of η can be made from mobility data on samples with impurity scattering.[21,34] According to the existing theory of impurity scattering,[34,35] an amount of impurity scattering sufficient to reduce the mobility to half that in pure material should reduce the ratio of Hall to drift mobility to a minimum value only slightly above unity. If we therefore assume that this amount of impurity scattering raises the effective r from $-\frac{1}{2}$ to 0, we can deduce from the observed temperature change required to increase the role of impurity scattering by this amount that for $s=0$ the temperature exponent η should fall below 7/2 by an amount of the order of 0.1 or 0.2 when the mobility has been reduced to half or so of the lattice mobility. At this stage, as Table I shows, Q_p should be about 15 percent below the value in pure material. If $s=1$ the reduction of η below 3 is 0.3 or so, and Q_p is about $\frac{2}{3}$ the value for pure material when the mobility has been reduced to half.

VI. EFFECT OF FINITE SIZE OF THE SPECIMEN

The analysis of the preceding section, giving a variation of Q_p as an inverse power of T, was based on the assumption in Eq. (36) for the variation of phonon relaxation time with temperature and wave number. This assumed law will break down at very low temperatures because of scattering of the phonons by the boundaries of the specimen. Such boundary scattering is the obvious explanation for the fact[2] that at the lowest temperatures Q_p falls below the power law extrapolation (see for example Fig. 4). It was therefore predicted while the experiments were in progress that in this temperature range the effective thermoelectric power could be appreciably changed by changing the diameter of the specimen. This prediction was confirmed.[2] The interest of this boundary effect lies in the opportunity which it provides not only to check the general theoretical picture we have presented, but also to extract from the experimental data some new information on the properties of the semiconductor. We shall therefore devote this section to analyzing the effect of boundary scattering in some detail. We shall consider here only the case of low carrier concentrations, so that Fermi degeneracy and the saturation effect discussed in the next section do not occur.

We can get a little orientation on the problem of boundary scattering by recalling, and slightly amplifying, the simple calculation which was made in Sec. II to establish an upper limit to Q_p. This calculation, after suitable refinement, gives the asymptote which Q_p must approach at very low temperatures. The necessary refinement is easy to make. We replace the expression in Eq. (16) by a sum of two terms, one for longitudinal and one for transverse phonons. We neglect anisotropy of sound velocity within each branch. The values $\bar{\tau}_{bl}$, $\bar{\tau}_{bt}$ to be used for the mean phonon relaxation time $\bar{\tau}$ are of course averages over the boundary scattering times of phonons with different directions of motion. Fortunately, however, these averages are to be made with a weighting factor q_x^2/q^2, as Eqs. (19) and (25) show; since this is the same weighting factor as occurs in the theory of thermal conduction,[25] we can use the relation between $\bar{\tau}_b$ and the dimensions of the specimen which has been derived by Casimir[36] for the latter case. This type of analysis gives

$$\bar{c}_l\bar{\tau}_{bl}=\bar{c}_t\bar{\tau}_{bt}=L, \tag{41}$$

where $\bar{c}_l$ and $\bar{c}_t$ are mean longitudinal and transverse sound velocities respectively, and where for a specimen of square cross section with completely diffuse scattering at the boundaries L is 1.1 times the square side. Noting that both for the simple model and for the many-valley model of Appendix A the variation of f_l/μ and f_t/μ with temperature and composition is proportional to that of Eq. (31), we have from Eqs. (16) and (1):

$$\text{asymptotic } Q_p=\mp(L/\mu T)(f_l\bar{c}_l+f_t\bar{c}_t)\propto f'T^{\frac{1}{2}}, \tag{42}$$

where as before f_l, f_t are the fractions of the crystal momentum of the carriers lost to longitudinal and transverse acoustic modes, respectively, and where f' is the ratio of angular brackets in Eq. (31). This quantity f', tabulated in Table I, depends on the amount and kind of impurity scattering. The absolute value of Q_p for a given lattice mobility μ_L depends slightly on the ratio of f_l to f_t; it would be independent of this ratio if $\bar{c}_l$ were the same as $\bar{c}_t$.

Since, as Fig. 4 shows, Q_p is still far below the asymptote in Eq. (42) at the lowest temperatures used by

[33] Suggested by C. Kittel, Phys. Rev. **94**, 768 (1954).
[34] P. P. Debye and E. M. Conwell, Phys. Rev. **93**, 693 (1954).
[35] H. Jones, Phys. Rev. **81**, 149 (1951); V. A. Johnson and K. Lark-Horovitz, Phys. Rev. **82**, 977 (1951); E. Conwell, Proc. Inst. Radio Engrs. **40**, 1331 (1952).
[36] H. B. G. Casimir, Physica **5**, 495 (1938).

Geballe and Hull, it is clear that to get a theoretical Q_p for these temperatures we shall have to consider the case where boundary scattering and phonon-phonon scattering are of comparable importance. However, there are two points worth noting in the simpler theory. One is that at the lowest temperatures the contribution of transverse phonons to Q_p may well be comparable with that of longitudinal ones, and so it may be that even at temperatures where Q_p is well below its asymptote the transverse contribution is perceptible. The other is that, as Table I shows, a semiconductor with ionized impurity scattering predominant should have a higher asymptotic Q_p than one with lattice scattering only, while one with neutral impurity scattering predominant should be intermediate. This effect of impurity scattering is opposite to that obtained from the theory of Sec. IV for temperatures above the boundary scattering range. This is to be expected: for the higher range of temperatures the decrease of the mean $\tau(\mathbf{q})$ with increasing impurity scattering outweighs the increase in the total crystal momentum delivered to the phonons by a given current, at least if (21) is valid; when boundary scattering is dominant, however, $\tau(\mathbf{q})$ is constant and the latter effect predominates. We shall see that this reversal in the effect of impurity scattering occurs at temperatures where Q_p is well below its asymptote.

Let us now try to calculate Q_p when boundary scattering and phonon-phonon scattering are of comparable importance. The usual procedure in cases of this sort is to assume that the reciprocal relaxation time for each mode is the sum of the contributions which the two types of scattering would yield separately. However, as we shall show, this assumption is not accurate enough for our present purpose. A better procedure is to consider how the crystal momentum in a volume element $\Delta_\mathbf{q}$ of q space is distributed spatially over the cross section of the specimen. In the presence of a current $\mathbf{J}$, crystal momentum is fed into this range of modes at a rate which is independent of position and proportional to the quantity $\mathbf{R}(\mathbf{q})$ defined in Section IV. Phonon-phonon scattering removes crystal momentum from these modes in every element of the volume, at a rate proportional to the local density $\Delta\mathbf{P}(\mathbf{q};\mathbf{r})$ of crystal momentum. Boundary scattering also removes crystal momentum but only from regions adjoining the boundaries; the rate is proportional to the value of $\Delta\mathbf{P}$ at the boundary and to the component of phonon velocity toward the boundary, if the latter is positive, while the rate is zero if the normal velocity is negative. For the simplest geometry, that of a plane-parallel slab with faces normal to the y axis, current in the x direction, we have in a steady state:

$$(2\pi)^{-3}\mathbf{R}(\mathbf{q})\,\Delta_\mathbf{q}-c(q_y/q)\partial\Delta\mathbf{P}/\partial y-\Delta\mathbf{P}/\tau(\mathbf{q})=0, \quad (43)$$

where $\tau(\mathbf{q})$ is the phonon-phonon relaxation time and where we assume for simplicity that the group velocity

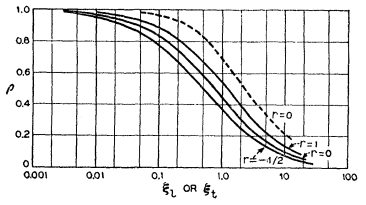

FIG. 6. Full curves: ratio ρ_l of the longitudinal phonon contribution to Q_p for a specimen of finite diameter to the value for an infinite specimen, as a function of the quantity ξ_l defined by (46), for different values of the exponent r in $\tau_e \propto |\epsilon-\epsilon_b|^r$. Dashed curve: corresponding ratio ρ_t for the transverse phonon contribution, abscissa ξ_t given by Eq. (47).

is $c\mathbf{q}/q$. If $q_y>0$ the boundary condition is that $\Delta\mathbf{P}$ must vanish on the back surface, which we take for simplicity to be the plane $y=0$. The solution of Eq. (43) is then

$$\Delta\mathbf{P}=\Delta\mathbf{P}_0[1-\exp(-yq/q_yc\tau)], \quad (44)$$

where

$$\Delta\mathbf{P}_0=(2\pi)^{-3}\mathbf{R}\tau\Delta_\mathbf{q}. \quad (45)$$

Thus for large $y/c\tau$, $\Delta\mathbf{P}$ is distributed uniformly over the specimen except for a small region near the back face (i.e., the face $\mathbf{q}$ points away from), while when $y/c\tau$ is very small $\Delta\mathbf{P}$ increases nearly linearly from back to front. Consequently, for a given total crystal momentum in a specimen of given size, the rate of removal by boundary scattering is twice as great for large τ as for small.

The ratio ρ of the Q_p of a finite specimen to that of an infinite specimen is simply the ratio of the integral of $\Delta\mathbf{P}$ over the cross section and over $\mathbf{q}$ to the integral of $\Delta\mathbf{P}_0$. An approximate evaluation of this ratio is given in Appendix B, the evaluation being exact as $T\to 0$ and in error as $T\to\infty$ by only 10 percent or 20 percent of the deviation of the ratio from unity. For longitudinal modes—$s=0$ in Eq. (36)—ρ_l is a function of the dimensionless variable,

$$\xi_l=3\pi\bar{c}_l\lambda^2/32A_lT^3L\propto 1/T^4L, \quad (46)$$

where λ is the thermal electron wavelength in Eq. (40). Physically ξ_l is the ratio of the phonon-phonon scattering time to $L/\bar{c}_l$ for phonons of wave number $(32/3\pi)^{\frac{1}{3}}/\lambda$. Graphs of ρ_l against ξ_l are given in Fig. 6 for three values of the exponent r in $\tau_e\propto|\epsilon-\epsilon_b|^r$, and some of the values are given in Table II. The calculation for trans-

TABLE II. Values of ρ_l plotted in Fig. 6.

	ρ_l	
ξ_l	$r=-\frac{1}{2}$	$r=0$
0.003	0.981	0.989
0.01	0.953	0.969
0.03	0.898	0.928
0.1	0.778	0.835
0.3	0.604	0.676
1.	0.371	0.440
3.	0.195	0.238
10.	0.085	0.10[illegible]

verse modes ($s=-1$) is similar: ρ_t depends on

$$\xi_t=3\pi\bar{c}_t\lambda/8A_tT^4L\propto 1/T^{9/2}L. \quad (47)$$

Figure 6 shows a plot of ρ_t for the case $r=0$; this is the most important case since boundary scattering becomes important for transverse modes only at very low temperatures, where neutral impurity scattering often dominates τ_e.

In future comparisons of theory and experiment along the lines of that given in Sec. IX below, it may turn out that important inferences can be drawn from the abruptness or gradualness of the onset of boundary scattering. It is only in regard to this feature that the various assumptions used in the theory can be tested, since any theoretical expression for $Q_p(L,T)$ which is asymptotic to $Q_P(\infty,T)$ at high T and to Eq. (42) at low T will give roughly the right sort of bending over of the plot of Q_p against T. The variation of abruptness of onset with different choices of L, A_l, A_t is of course covered by Eqs. (46) and (47) and Fig. 6. We may also ask, however, how the abruptness varies with the value of the exponent s in Eq. (36), or with variation of the nature of the boundary scattering, e.g., if the boundary scattering takes place at grain or mosaic boundaries, or at finely distributed foreign inclusions. The answer to the latter question is supplied, at least in some cases, by the remarks made after Eq. (45). For under conditions where one should assume additivity of phonon-phonon and boundary scattering probabilities, boundary scattering is twice as effective at high T, for a given effectiveness at low T, as when Eq. (44) applies. Thus making the boundary scattering take place uniformly throughout the volume makes the onset as T is lowered more gradual. The effect of varying the exponents can be surmised from a comparison of Eq. (46) with Eq. (47) and of the full with the dashed curve in Fig. 6. The onset of boundary scattering is more abrupt for $s=-1$ (transverse) than for $s=0$ (longitudinal), both because Eq. (47) has a higher exponent of T than Eq. (46) and because the initial downard bend of the dashed curve is more abrupt than that of the full curve. One can easily see, after a little reflection on the form of the curve of $\tau(\mathbf{q})$ against q and its shift with temperature, that this trend to a more gradual onset with increasing s must hold over the entire range of possible values of s. Explicit quantitative calculations, not reproduced here, have given additional confirmation of this fact.

The theory of this section contains only one constant not previously introduced, namely, the boundary scattering length L. This constant is determined by the size and shape of the specimen if it can be assumed that the scattering of elastic waves impinging on the surface of the specimen is completely diffuse; in any case, a calculation on this assumption gives a lower limit to L. Three facts make the assumption of diffuse scattering plausible for specimens like those of reference 2. The first is the approximate agreement of the theory of Casimir,[36] based on this assumption, with thermal conduction experiments[37] on rods of quartz, potassium chloride, sapphire, and diamond, at temperatures low enough to be in the boundary scattering range. The second is that the surfaces of the specimens of reference 2 were sandblasted. The third is that when a longitudinal wave impinges on the surface, a good part of the reflected energy will reside in transverse nodes, and if the temperature is not too low this part of the energy flux will quickly be randomized by phonon-phonon scattering.

VII. THE SATURATION EFFECT IN THE NONDEGENERATE RANGE

In the caluclations of the two preceding sections it has been assumed that the probability of emission or absorption of a phonon by a charge carrier is the same as it would be if the phonon system were in thermal equilibrium. In Sec. IV, where this assumption was introduced, it was mentioned that it is asymptotically valid for small densities of charge carriers, but must fail for sufficiently high densities. It is easy to see that when this assumption breaks down, Q_p will be less than the value given by the theory of the preceding sections. Consider first the II picture. Suppose the charge carriers are drifting in the positive x direction under the action of an applied field. If there are more phonons moving in the positive than in the negative x direction, then scattering processes in which the x component of velocity of a carrier is increased will be more frequent than if the phonons were in thermal equilibrium, while those in which the x component of velocity is decreased will be less frequent. For a given velocity distribution of the carriers, therefore, the resistance offered by the lattice vibrations to the motion of the carriers will be less than if the phonons were in equilibrium, and concomitantly the carriers will deliver less crystal momentum to the phonons. Thus the phonon contribution Π_p to the Peltier flux will be less than one would calculate on the assumption of phonon equilibrium. One can reach the same conclusion using the Q picture: if the density of charge carriers is sufficiently high the thermal conduction current carried by the low-energy phonons in a temperature gradient will be decreased by the interaction of these phonons with the charge carriers, and the phonons will therefore exert less drag on the carriers than they would if the perturbation of the phonon distribution by the carriers were neglected. In this section we shall try to discuss this effect quantitatively.[38]

As was mentioned in Sec. III, the effect we wish to discuss occurs when, for an isothermal current-carrying

[37] W. J. de Haas and Th. Biermasz, Physica **5**, 47, 320, 619 (1938); R. Berman, Proc. Roy. Soc. (London) **A208**, 90 (1951); Advances in Physics **2**, 103 (1953); Berman, Simon, and Ziman, Proc. Roy. Soc. (London) **A220**, 171 (1953). The latter authors have estimated the reflection for diamond to be roughly midway between diffuse and specular.

[38] A study of the influence of this effect on the electrical conductivity and its frequency dependence has been carried out by B. Goodman (unpublished). I am indebted to Dr. Goodman for information about this work.

specimen, "the unbalance of the phonons is of the same order as that of the carriers." This suggests that we begin by formulating a quantitative definition of "unbalance." This is easily done when, as in the present problem, the most important interactions between the various parts of a statistical assembly conserve a vector quantity such as crystal momentum. In the usual formulation of statistical mechanics for a complicated assembly, it is shown that if energy is the only extensive quantity which is conserved in the motion of the assembly and in its interaction with external systems, such as thermal reservoirs, there is a one-parameter family of equilibrium distributions, the single parameter being temperature, an indirect measure of energy content. But if several extensive quantities are conserved, the possible equilibrium distributions contain a corresponding number of parameters.[39] Thus if all the interactions of electrons and phonons with each other conserved crystal momentum as well as energy, the possible equilibrium distributions would be described by parameters T, $\boldsymbol{\beta}$, such that in a given distribution the probability of any quantum state of total energy E and total crystal momentum $\mathbf{p}$ would be proportional to $\exp[(-E+\boldsymbol{\beta}\cdot\mathbf{p}/kT]$. Note that $\boldsymbol{\beta}$ has the dimensions of velocity. In particular, a charge carrier distribution of this sort, with a given value of $\boldsymbol{\beta}$, could be in equilibrium with a phonon distribution having the same $\boldsymbol{\beta}$. A repetition of the usual Fermi-Dirac and Einstein-Bose derivations gives for the single-particle distribution functions of electrons and phonons,[24] respectively,

$$f(\mathbf{K};\boldsymbol{\beta},T)=\left[\exp\left(\frac{\epsilon(\mathbf{K})-\epsilon_F-\boldsymbol{\beta}\cdot\hbar\mathbf{K}}{kT}\right)+1\right]^{-1}, \quad (48)$$

$$N(\mathbf{q};\boldsymbol{\beta},T)=\left[\exp\left(\frac{\hbar\omega(\mathbf{q})-\boldsymbol{\beta}\cdot\hbar\mathbf{q}}{kT}\right)-1\right]^{-1}, \quad (49)$$

where as previously ϵ_F is the Fermi level, $\epsilon(\mathbf{K})$ is the energy of an electronic state of wave number $\mathbf{K}$, etc. Although these equations represent an equilibrium distribution only when all interactions conserve crystal momentum, they form a useful reference point even in the presence of other types of motivations and interactions (impurity scattering, electric field, Umklapp collisions, etc.), because they show the relation between the electron and phonon distribution functions which the electron-phonon intereaction is attempting to establish. Thus if we take all $\boldsymbol{\beta}$'s to be in the field direction x, we may define an unbalance velocity β_{ex} for each electronic state $\mathbf{K}$ by Eq. (48), and similarly a β_{px} for each phonon mode $\mathbf{q}$ by Eq. (49), and may describe the electron-phonon interaction as one which attempts to equalize the β_{ex} and β_{px} of each carrier-mode pair which interact.

[39] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, in *Collected Works* (Longmans, Green, and Company, New York, 1928), Vol. 2, pp. 37–40.

When an infinitesimal current flows in an isothermal conductor the $\boldsymbol{\beta}_e$ and $\boldsymbol{\beta}_p$ just defined are infinitesimal. On equating Eq. (48) to the usual expression $f^{(0)}+f^{(1)}$ for the electronic distribution function in the presence of a field $\mathbf{E}$ one finds easily

$$\beta_{ex}(\mathbf{K})=\pm(eE/m^*)\tau_e(\mathbf{K})=v_{Dx}(\mathbf{K}), \quad (50)$$

where the upper sign is for electrons, the lower for holes, and where $\mathbf{v}_D$ is an effective drift velocity of the carriers in states $\mathbf{K}$ and $-\mathbf{K}$. Here we have for simplicity assumed the conventional semiconductor model of Fig. 5(c), with effective mass m^*, and have assumed a relaxation time $\tau_e(\mathbf{K})$ to exist. Similarly we find from Eq. (49)

$$\beta_{px}(\mathbf{q})=\frac{[\omega(\mathbf{q})]^2\Delta p_x(\mathbf{q})}{kTq_x^2}\approx\left[\frac{q}{q_x}\right]^2\left[\frac{c\tau(\mathbf{q})R_x(\mathbf{q})}{kT}\right]c, \quad (51)$$

where $\Delta\mathbf{p}(\mathbf{q})$ is the excess crystal momentum in the $\mathbf{q}$th mode, given by Eq. (18), $\tau(\mathbf{q})$ is its relaxation time, c is a mean sound velocity, and $\mathbf{R}(\mathbf{q})$ is as before the rate at which the electronic system feeds lattice momentum into the $\mathbf{q}$th mode. Note that in our problem $\Delta p_x \propto q_x^2/q^2$, so β_{px} is roughly independent of the direction of $\mathbf{q}$.

In the approximation we have been using in the preceding sections, $\mathbf{R}(\mathbf{q})$ is given by Eq. (25) or (A4) and $\beta_{px}(\mathbf{q})$ is thus an integral over the various $\beta_{ex}(\mathbf{K})$. When the various β_{px} become comparable with the β_{ex}, however, the $\mathbf{R}(\mathbf{q})$ of Eq. (25) or (A4) must be replaced by an expression which vanishes when $\boldsymbol{\beta}_p=\boldsymbol{\beta}_e$. Since the $\boldsymbol{\beta}$'s are infinitesimal, the correct expression must have the form,

$$\beta_{px}(\mathbf{q})=\int\Phi(\mathbf{q},\mathbf{K})[\beta_{ex}(\mathbf{K})-\beta_{px}(\mathbf{q})]d\mathbf{K}, \quad (52)$$

where Φ can be taken from Eq. (25) or (A4). Since the approximation of the preceding sections consisted in neglecting the β_{px} in the integrand of Eq. (52), we have for the ratio of the Q_p derived from Eq. (52) to the $Q_p^{(0)}$ derived previously

$$\frac{Q_p}{Q_p^{(0)}}=\int\left(\frac{\int\Phi\beta_{ex}d\mathbf{K}}{1+\int\Phi d\mathbf{K}}\right)d\mathbf{q}\Big/\int\int\Phi\beta_{ex}^{(0)}d\mathbf{K}d\mathbf{q}. \quad (53)$$

Here we distinguish $\beta_{ex}^{(0)}$, the value calculated neglecting saturation, from β_{ex}, the true value. Although we shall not attempt such a refinement, the latter should strictly be calculated from an integral expression analogous to Eq. (52) but with the roles of $\mathbf{K}$ and $\mathbf{q}$ interchanged.

We shall not try to evaluate Eq. (53) to as high a degree of refinement or for such a variety of conditions as we have employed in the preceding sections; this is partly because of the greater computational difficulty of Eq. (53) and partly because for conditions where Eq. (53) is significantly below unity some of our basic assumptions are probably not very accurate, e.g., the

neglect of the influence of impurities on the band structure. We shall content ourselves with assuming $\beta_{ex}=\beta_{ex}^{(0)}$ and treating only longitudinal modes for the case $\tau_e=$ constant, a case which is probably not too far from the truth throughout a good part of the concentration range of interest. For this case, by Eq. (50), β_{ex} is independent of $\mathbf{K}$, and it can be factored out of Eq. (53). On comparing Eqs. (25) and (32) with (51) and (52), with the β_{px} in the integrand of the latter omitted, we find, for longitudinal modes $\mathbf{q}$,

$$\int \Phi d\mathbf{K}=\frac{8\pi}{3}\left(\frac{\bar{c}_l^2 ne\lambda^4 f_l'}{kT\mu_L}\right)q\tau(\mathbf{q})\exp(-\tfrac{1}{4}\lambda^2q^2), \quad (54)$$

where as before n is the density of carriers, λ is the thermal electron wavelength in Eq. (40), μ_L is the value the mobility would have if only low-energy phonons scattered, and f_l' is the fraction of the crystal momentum which would be delivered to longitudinal modes under these conditions. Since the most interesting saturation effects occur at low temperatures where boundary scattering may be appreciable, it is desirable to use an expression for $\tau(\mathbf{q})$ which at least approximates those of the preceding section. We shall therefore assume

$$1/\tau(\mathbf{q})=A_lT^3q^2+(\bar{c}_l/L). \quad (55)$$

This is a crude version of the additive probability assumption which in the preceding section we rejected as inadequate. However, it is good enough for the present purpose, and it is easy to see that an approach paralleling that of Eq. (43) would lead to great mathematical complications. Insertion of Eqs. (55) and (54) into Eq. (53) and changing of the variation of integration to $\exp(-\frac{1}{4}\lambda^2q^2)$ give, for the case $\tau_e=$ constant,

$$\frac{Q_{pl}}{Q_{pl}^{(0)}}=\frac{\int_0^1[1+(\lambda^2\bar{c}_l/4LA_lT^3)(-\ln x)^{-1}+(4\pi/3)(\bar{c}_l^2ne\lambda^5/A_lkT^4\mu_l)x(-\ln x)^{-\frac{1}{2}}]^{-1}dx}{\int_0^1[1+(\lambda^2\bar{c}_l/4LA_lT^3)(-\ln x)^{-1}]^{-1}dx}. \quad (56)$$

The denominator of Eq. (56) represents the value of ρ_l, the ratio of Q_{pl} (without saturation) for a finite specimen to that for an infinite specimen, as given by the present approximation in Eq. (55). The ratio in Eq. (56) can therefore be described as a function of the two parameters ρ_l and n/n_1, where

$$n_1=(3/4\pi)(kT^4\mu_LA_l/f'\bar{c}_l^2e\lambda^5)=\frac{1}{4\pi}\left(\frac{T}{T_0}\right)^5\frac{(k/e)}{Q_{pl}^{(0)}(T_0)}\cdot\frac{1}{[\lambda(T_0)]^3} \quad (57)$$

by Eq. (38), where T_0 is any convenient reference temperature at which it is legitimate to take $\mu=\mu_L$ (so $r=-\frac{1}{2}$), and where $Q_{pl}^{(0)}(T_0)$ is to be evaluated for $L=\infty$.

Figure 7 shows some plots of $Q_{pl}/Q_{pl}^{(0)}$ against n/n_1 for various values of ρ_l, as obtained by numerical evaluation of Eq. (56). The behavior of these is qualitatively understandable. Increasing the amount of boundary scattering decreases each $\tau(\mathbf{q})$, hence decreases Eq. (54). Since the decrease in this quantity is always by a greater ratio than the decrease in this quantity divided by one plus itself, the denominator of Eq. (53) decreases more than the numerator. Thus for a given A_l, the greater the boundary scattering, the larger Eq. (53).

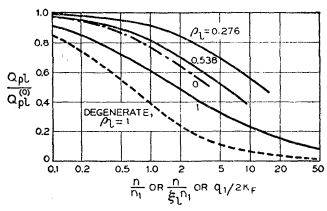

FIG. 7. Full curves: saturation ratio $Q_{pl}(T,L,n)/Q_{pl}^{(0)}$, for nondegenerate statistics, as a function of carrier density n, measured in units of the n_1 of Eq. (57). Here $Q_{pl}^{(0)}$ is $Q_{pl}(T,L,0)$, and the three curves are for different values fo $\rho_l\equiv Q_{pl}(T,L,0)/Q_{pl}(T,\infty,0)$. Dot-dash curve: same for case of pure boundary scattering ($\rho_l\to 0$), abscissa $n/\xi_l n_1$, where ξ_l is given by Eq. (46). Dashed curve: saturation ratio for degenerate statistics, $L=\infty$, as a function of $q_1/2K_F$, q_1 given by Eq. (61).

The same sort of reasoning gives us some insight into the extent of the errors introduced by the assumption $\beta_{ex}=$ constant. The quantity $\int\Phi d\mathbf{K}$ is given by Eq. (54) independently of the law of variation of β_{ex}, which enters only as shown in Eq. (53). Since $\int\Phi d\mathbf{K}$ decreases with increasing q except when $\tau(\mathbf{q})$ is close to the boundary scattering value, we may expect that at not too low temperatures any increase in the $|\beta_{ex}|$ of states of large K relative to those of small K will increase $Q_p/Q_p^{(0)}$. Now there are two things which may cause β_{ex} to depart from constancy. One is that the combination of lattice and impurity scattering may cause the effective exponent r in $\tau_e(K)\propto K^{2r}$ to differ from the value zero which we have assumed in this section. From what has just been said we may describe this effect by saying that the more impurity scattering, the less saturation. The other effect is the difference between β_{ex} and $\beta_{ex}^{(0)}$ in Eq. (53). It is clear that β_{ex} is larger than the $\beta_{ex}^{(0)}$ given by Eq. (50) with the unsaturated τ_e, since the phonons do not slow down the charge carriers as effectively as previously assumed. Therefore the ratio $Q_p/Q_p^{(0)}$ will be greater than it would be for $\beta_{ex}=\beta_{ex}^{(0)}$, i.e., the calculation leading to Eq. (56) slightly overestimates the amount of saturation.

It can be verified that if, for given $Q_p^{(0)}$, we make the

decrease of $\tau(\mathbf{q})$ with q more gradual by using an $s<0$ in Eq. (36), the value of Eq. (53) will be increased.

Finally, we note that if transverse modes contribute appreciably to Q_p, $Q_p/Q_p^{(0)}$ will be less than the $Q_{pl}/Q_{pl}^{(0)}$ computed here, because the transverse modes saturate less readily than the longitudinal.

VIII. EFFECTS OF FERMI DEGENERACY

All the calculations of Sec. II and Secs. IV through VII have assumed the charge carriers to obey Maxwell-Boltzmann statistics. We wish now to examine briefly the way in which the electronic and phonon contributions to the thermoelectric power are modified when the carrier density is so high that degenerate Fermi statistics must be used. Our treatment of this case will be very rough, for two reasons: in the first place, as was pointed out in the introduction, we do not yet know how to treat degenerate semiconductors accurately; secondly, the phonon contribution, which is the principal concern of this paper, is much less important at high carrier densities than at low, partly because of the saturation effect discussed in the preceding section.

Wright[7] has given an expression for the thermoelectric power of the conventional semiconductor model [Fig. 1(c)], for arbitrary degree of degeneracy. Just as in Sec. II, if the charge carriers have a relaxation time τ_e dependent only on energy the conventional model may be expected to give the same Q_e as the models of Figs. 1(a) and 1(b), provided that any spin-orbit splitting of the former model is either very large or very small. If $\tau_e \propto |\epsilon-\epsilon_b|^r$, Wright's expression for eTQ_e is the same as Eq. (2) with

$$\Delta\epsilon_T = \pm\left(\frac{5+2r}{3+2r}\right)\frac{F_{r+\frac{3}{2}}}{F_{r+\frac{1}{2}}}, \tag{58}$$

where

$$F_\nu(\zeta^*) = \int_0^\infty \frac{\zeta^\nu d\zeta}{1+e^{\zeta-\zeta^*}}, \quad \zeta^* = \pm\frac{(\epsilon_F-\epsilon_b)}{kT},$$

and where the upper sign is for n type, the lower for p. As in Sec. II, ϵ_F is the Fermi level, ϵ_b the band edge energy. The F_ν for half-integral ν values can be obtained from the tables of McDougall and Stoner.[40] Figure 8 shows the Q_e values calculated in this way from Eqs. (2) and (58), for $r=0$ and 1, as functions of ζ^*. Curves of Q_e *vs* T can be calculated from these, for any desired case, by using the relation of ζ^* to T given by the dashed curve. The mobilities of typical specimens just starting to become degenerate, analyzed in terms of the Conwell-Weisskopf-Brooks formula,[17] suggest values of r in the range 0.8 to 1.2. As extreme degeneracy is approached, however, it is not likely that r gets any closer to the Conwell-Weisskopf limit 1.5, because the scattering by the impurity ions becomes more nearly isotropic and less dependent on energy as their screening radius becomes shorter. Moreover, as we have already noted for the non-degenerate case, electron-electron scattering makes r closer to zero than in the one-electron theory.

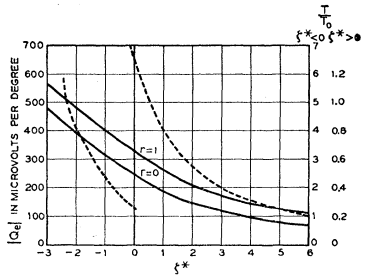

FIG. 8. Full curves: Q_e for the simple model with degenerate statistics, as a function of $\zeta^*=\pm(\epsilon_F-\epsilon_b)/kT$, for two values of the exponent r in $\tau_e \propto |\epsilon-\epsilon_b|^r$. Dashed curve (two scales on right): relation of ζ^* to T/T_0, where $T_0=(n/5.30\times10^{14})^{\frac{2}{3}}m/m^{(N)}$.

We turn now to the phonon contribution Q_p under conditions of Fermi degeneracy. If the saturation effect of the preceding section were absent, the principal effect of degeneracy ought to be to increase the effective average q of the modes to which the carriers deliver their crystal momentum, hence to decrease the effective $\tau(\mathbf{q})$ and so decrease Q_p below the value given by Eq. (38). If we make the assumption, now definitely risky, that lattice and impurity scattering act independently, so that (21) holds, the effect can be evaluated quantitatively. Although saturation is usually quite advanced before strong degeneracy sets in, a calculation neglecting saturation is still of some value, as it sets an upper limit to Q_p. The calculation is easy if $|\epsilon_F-\epsilon_b| \gg kT \gg \hbar\omega(\mathbf{q})$, where the $\omega(\mathbf{q})$ are the frequencies of the phonons which do most of the scattering. The calculation for this case is given in Appendix A, and leads to Eq. (A10). The relation of this equation to Eq. (38) or (A8) is simple: the ratio of averages in the last factor of Eq. (38) has been replaced by the ratio of the same quantities evaluated at the Fermi surface; this takes account of the shorter phonon lifetime due to the larger q's with which the degenerate electrons interact. This excitation of larger q values is shown in more detail by the sample plot ln Fig. 5 of $\bar{R}_x(q)$ for a degenerate specimen, a plot derived from Eq. (A4). However, it must be remembered that Eq. (A8), and hence Eq. (A10), cease to be valid when the degeneracy is so advanced that the phonon energies are not $\ll kT$. For such high degeneracy the average q of the phonons effective in lattice scattering will be smaller than we have assumed, and the last factor in Eq. (A10) should be replaced by something larger, probably of the order of $(m^*c^2/kT)^{1+\frac{1}{2}s}$. However, Q_p continues to decrease, since under these conditions τ_{el}^{-1} decreases with increasing degeneracy, at least when (21) is assumed to hold.

In actual cases the phonon part of the thermoelectric power will of course be reduced by the saturation effect. The expression in Eq. (53) of the preceding section gives the factor by which the calculation of the preceding

[40] J. McDougall and E. C. Stoner, Trans. Roy. Soc. (London) **A237**, 67 (1938).

paragraph must be reduced. For the present case we have, analogously to Eq. (54),

$$\int \Phi d\mathbf{K} = (\bar{c}_l^2 ne/kTJ)[q^2 R_x^{(0)}(\mathbf{q})/q_x^2]\tau(\mathbf{q}), \qquad (59)$$

where $R_x^{(0)}$ is the value computed with neglect of saturation. The latter quantity is given by Eq. (A4) of Appendix A. We shall consider here only the simplest case, that of complete degeneracy, longitudinal phonons ($s=0$), and single-valley spherical energy surfaces. For this case $q^2R_x^{(0)}/q_x^2$ is proportional to q up to a value of q equal to twice the radius K_F of the Fermi surface, and is zero for $q>2K_F$. In other words, the dashed curve of Fig. 5 would be simply a triangle if degeneracy were complete. The factor of proportionality in $R_x^{(0)}$ can be evaluated by the condition,

$$(2\pi)^{-3}\int R_x(\mathbf{q})d\mathbf{q} = (\tau_e/\tau_{eL})_{\epsilon_F} enE,$$

where as in Sec. IV τ_{eL} is the relaxation time for scattering by the low-energy longitudinal modes. With $s=0$ in Eq. (36) for $\tau(\mathbf{q})$ the expression in Eq. (59) reduces to

$$\int \Phi d\mathbf{K} = q_1/q, \quad (q<2K_F)$$
$$= 0, \quad (q>2K_F), \qquad (60)$$

where

$$q_1 = \tfrac{3}{2}\pi^2(\tau_e/\tau_{eL})_{\epsilon_F}\lambda^4 k\bar{c}_l^2 ne/\mu\epsilon_F^2 A_l T^2. \qquad (61)$$

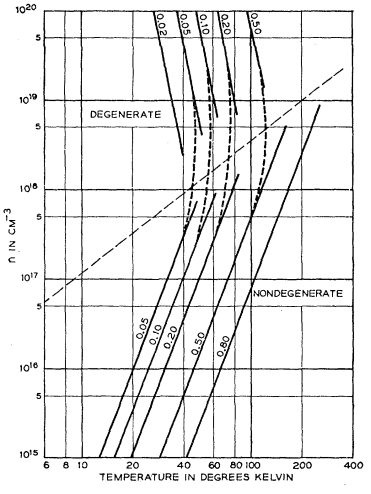

FIG. 9. Contours of constant saturation ratio $Q_{pl}/Q_{pl}^{(0)}$ in the $n-T$ plane, for the case $m^{(N)}=m$, $m^*=0.3m$, $Q_{pl}^{(0)}(80°)=1000$ μv/deg. The dashed curve is the boundary between degenerate and nondegenerate conditions, arbitrarily defined as the locus of values for which the Fermi level coincides with the band edge. The full lines are taken from Eq. (57) and Fig. 7 (nondegenerate) and from Eqs. (61) and (62) (degenerate). The dotted lines indicate qualitatively the connection of the regions of validity of the degenerate and nondegenerate formulas. Boundary scattering has been neglected.

Inserting Eq. (60) into (Eq. (53) we find, after an elementary integration

$$Q_{pl}/Q_{pl}^{(0)} = 1-(q_1/K_F) + \tfrac{1}{2}(q_1/K_F)^2 \ln[1+(2K_F/q_1)]. \qquad (62)$$

As n increases, q_1/K_F decreases, since it goes as n/ϵ_F^2 if τ_{eL} is proportional to K_F^{-1}. Thus Eq. (62) predicts less saturation the larger n. This is clearly because increasing K_F increases the mean q of the participating phonons, and so, by decreasing the mean $\tau(\mathbf{q})$, makes the phonons harder to saturate. The effect would be even more pronounced if we assumed an exponent $s>0$ in Eq. (36). However, as is shown by the sample contours of constant saturation ratio in Fig. 9, saturation is always marked at degenerate densities.

It must be borne in mind that the calculations of this section, including in particular Fig. 9, have many limitations. We have assumed the conventional model of a semiconductor, and so have negelcted such things as perturbation of the band structure by the impurity ions, electron interactions, and nonadditivity of scattering probabilities. Moreover, our calculations of $Q_{pl}^{(0)}$ and $Q_{pl}/Q_{pl}^{(0)}$ are crude in that they assume $|\epsilon_F-\epsilon_b| \gg kT$ but $\hbar cK_F \ll kT$. Finally, both Eq. (56) and Eq. (62) must become untrustworthy when they predict extremely small saturation ratios, since they are based on the assumption that the crystal momentum of a low-energy phonon disappears when it suffers a collision; actually, it is merely delivered to higher-frequency modes with rather short relaxation times. Very roughly, we may say that the saturation effect should not decrease Q_{pl} below the value of Fig. 4 corresponding to the use in Eq. (16) of the conductivity relaxation time τ_c given by Eq. (17). However, this limit is not likely to be reached when the contribution of transverse modes to Q_p is appreciable, since these saturate much more slowly than the longitudinal ones we have treated.

IX. COMPARISON WITH EXPERIMENT

The theoretical notions which have been presented in the preceding sections account qualitatively for most of the features of the thermoelectric power data of Geballe and Hull[2]. Specifically, they account for the rise of the thermoelectric power Q as T is lowered (Sec. IV) the bending over at the lowest temperatures (Sec. V), the decrease of the low temperature Q when the diameter of the specimen is reduced (Sec. V), the approximate lack dependence of Q on carrier concentration n at low n (Sec. IV), and the decrease of Q at high n (Secs. VI and VII). Table III summarizes the way in which the principal adjustable constants of the theory enter into these various phenomena. Only those constants are listed which are not usually well known from other types of measurement and which have a major, rather than just a minor, effect on the phenomenon under consideration. In each entry an effort has been made to arrange the constants in order of decreasing importance. In the

middle two rows the value of Q_p at moderate temperature and low n is assumed known. In the last two rows (degenerate) the effective masses have been primed to emphasize the fact that they need not be the same as for nondegenerate material.

Because of the probable availability in the near future of more accurate and detailed experimental data, it seems best to give merely an illustrative discussion of the data of reference 2, without attempting to extract accurate values of the constants or to discuss the uniqueness of the fit critically. We shall try to show only (i) that approximate quantitative agreement of theory and experiment is possible with entirely reasonable values of the constants; (ii) that even the gross features of the data lead to some significant new conclusions about germanium; and (iii), that the study of thermoelectric power at low temperatures is a promising tool for getting a great deal of useful information on semiconductors.

Most of the comparisons to be made below will be for p-type rather than n-type germanium. This is because it is more likely for p- than for n-germanium that the band structure is one to which most of the analysis of this paper is applicable in the medium- and high-temperature range. Specifically, the band structure of p germanium is very probably either of the many-valley type of Fig. 1(b), or of the single-valley degenerate type with a fairly large spin-orbit splitting and energy surfaces which near the band edge are nearly concentric spheres.[33] In the former case all the analysis of this paper applies at any temperature, and in the latter case it should apply when kT is small compared to the spin-orbit splitting, a condition which should be fulfilled below 100°K, though perhaps not at room temperature. For n-germanium, on the other hand, the disparity between the effective masses measured in the cyclotron resonance experiment[41] and the inertial average measured by Benedict and Shockley[15] suggests that the model for this case may be of the degenerate many-valley type, with a spin-orbit splitting small enough so that departures from constant curvature in the $\epsilon-K$ relation may be appreciable at 100°K. If this is the case, the mathematics of this paper is inapplicable except at very low temperatures. Even if the structure is of the simple many-valley type, Fig. 1(b), the extremely high anisotropy[41,42] of the effective mass will make the value of the m^* of Appendix A very uncertain.

1. *P*-Type Germanium, Low *n*

In the range around room temperature the phonon contribution Q_p is negligible compared with the electronic term Q_e. As Geballe and Hull[2] have shown, the data in this region can be fitted by using Eqs. (4) and (7) for the Q_e of the holes, and combining this, when intrinsic conduction is appreciable, with a similar formula for the Q_e of the electrons, according to the rule for emf's in parallel. They found $m^{(N)}/m=0.75\pm0.15$, assuming the effective exponent r in $\tau_e\propto|\epsilon-\epsilon_b|^r$ to vary from -0.5 to -0.8 in the way proposed in Sec. II above. The result is not especially sensitive to this assumption. This is roughly consistent with an inertial mass $m^{(I)}=0.3m$ and a six-valley model. If on the other hand one assumes a similar inertial mass but a band edge at $K=0$ with spin-orbit splitting, as cyclotron resonance experiments[41,43] suggest,[33] the only possibility for fitting the thermoelectric data is to have the spin-orbit splitting small enough to cause appreciable departures from Eqs. (3) and (6).

TABLE III. Constants determining various features of the thermoelectric power $Q(n,T,L)$.

Feature or region	Principal constants	Eqs.
High T, low n, Q_e	$m^{(N)}$, $(\Delta\epsilon_T/kT)$ or r	(4)
Medium T, low n, Q_p	$A_l m^*/f_l$	(38)
Low T, low n, effect of size on Q_p	Lf_l, and perhaps $A_l m^{*\frac{1}{2}}L$, $(1-f_l)L$	(46) (47), (38)
Saturation of Q_{pl}, nondegenerate	m^*	(57)
High T, degenerate, Q_e	$m^{(N)\prime}$, r	(58)
Medium T, degenerate, Q_p	$A_l m^{*\prime}/m^{(N)\prime\frac{1}{2}}f_l$, $m^{(N)\prime}/m^{*\prime}$ $A_l m^{*\prime\frac{1}{2}}/m^{(N)\frac{1}{2}}f_l$, etc.	(A10), (61), (62),

In the range around liquid air temperature Q_p can be determined fairly accurately as $Q-Q_e$, with Q_e from Eq. (8) with the previously determined $m^{(N)}$ and r. Geballe and Hull[2] have plotted, in their Fig. 6, values obtained in this way, and we have already remarked, in Sec. V, the approximate agreement of the predicted and observed values of the exponent η in $Q_p\propto T^{-\eta}$. Although we shall see below that boundary scattering and possibly even the contribution of the transverse phonons can affect η appreciably, the absolute value of Q_p at say 80°K is only slightly different from the contribution of Eq. (38) of the longitudinal phonons in an infinite medium. By Eqs. (38), (39), and (40) this value therefore determines the quantity $A_l(m^*)^{1+\frac{1}{2}s}/f_l$, if r and s are known or assumed. For the purest specimens we may take $r=-\frac{1}{2}$ at 80°. For the ideal case $s=0$, the last factor of Eq. (38) reduces to $\frac{2}{3}\lambda^2$, and with $\gamma=0$ we have

$$A_l m^*/f_l m=\hbar^2\bar{c}_l^2/6kT^5\mu m Q_{pl}. \tag{63}$$

Numerical values are

$$\bar{c}_l=5.33\times10^5 \text{ cm/sec},$$
$$\mu=3.8\times10^4 \text{ cm}^2/\text{volt sec},$$
$$Q_p=9.4\times10^{-4} \text{ volt/deg},$$

for specimen 8 of the Geballe-Hull paper at 80°. These give the value 3.6×10^{-12} cm^2/deg^3 sec for the right of Eq. (63). This value needs to be corrected for boundary scattering and possibly for the contribution of transverse phonons; as we shall see, the corrections are

[41] Lax, Zieger, Dexter, and Rosenblum, Phys. Rev. **93**, 1418 (1954).
[42] S. Meiboom and B. Abeles, Phys. Rev. **93**, 1121 (1954).
[43] Dresselhaus, Kip, and Kittel, Phys. Rev. **92**, 827 (1953).

probably such as to increase the value given by something like 20 percent.

Comparisons of theory and experiment in the boundary scattering range have been made for a number of different choices of the amount of participation of transverse modes and various values of the ratio $f_l L/L_0$, i.e., the product of the fraction of the scattering of the carriers due to longitudinal modes by the ratio of the true boundary scattering length to that for diffusely scattering boundaries. Figure 10 shows an example for which almost no participation of transverse phonons is required, i.e., for which Q_{pt} is negligible over the termperature range covered, even though $f_t=1-f_l$ is sizable. The theoretical curve for $r=-\frac{1}{2}$ was constructed by choosing arbitrarily a value for the ξ_l of Eq. (46) at 80°, hence of ρ_l (80°), determining the Q_{pl} of an infinite specimen at 80° so that $\rho_l Q_{pl}$ (80°,∞) equalled the observed Q_p, determining $\rho_l(T)$ from Eq. (46) and Fig. 6, and finally multiplying ρ_l by $(80/T)^{3.5}Q_{pl}(80°,\infty)$. The value shown for $f_l L/L_0$ is, of course, merely the value required to give the chosen $\xi_l(80°)$ and $Q_{pl}(80°,\infty)$, when Eqs. (46) and (38) are combined. The curve for $r=0$, constructed to correspond to the same atomic constants, is shown to illustrate the order of magnitude of the effect which might be produced by the increasing importance of impurity scattering at the lower temperatures.

Several conclusions can be drawn from the comparisons of theory and experiment which have been made to date, and they can all be illustrated by reference to Fig. 10. The most noteworthy and striking of these is that

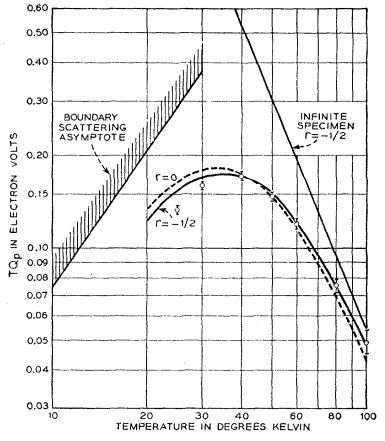

FIG. 10. Sample comparision of theory and experiment on Q_p for p type germanium with $n=1.9\times10^{14}$ cm^{-3}, $L_0=0.153$ cm. Points: observed Q (Geballe and Hull, specimen 8) minus Q_e obtained from (8) with $n^{(N)}=0.75m$, $\Delta\epsilon_T=2kT$. The vertical lines give the order of magnitude of the uncertainty due to scatter in the observed Q's; no estimate of systematic errors has been included. Full curve: Q_{pl} for $r=-\frac{1}{2}$, computed as described in the text from the theory of Sec. VI, with fit to the observed value at 80°, and with the parameter choice $f_l(80°)$ $L/L_0=0.46$. Dashed curve: Q_{pl} for $r=0$. The line at the right represents the values Q_{pl} would attain if there were no boundary scattering; that at the left gives the lower limit to $Q_{pl}+Q_{pt}$ when phonon-phonon scattering becomes negligible compared with boundary scattering, as calculated from Eq. (42) with $L=L_0$, $(f_l+f_t)/\mu=(Q/80)^{\frac{3}{2}}/\mu(80°)$.

an f_l value at 80° of the order $\frac{1}{2}$ or less is required to fit the data, i.e., that about half or more of the lattice scattering at 80° must be due to shear modes. This is qualitatively obvious from the fact that the observed points bend away from the "infinite specimen" line so far below where this line intersects the lower limit of the "boundary scattering asymptote" derived from Eq. (42). If this early bendover is not due entirely to the saturation effect, a possibility which seems ruled out by Fig. 4 of reference 2, it must mean that the contribution Q_{pl} of the longitudinal modes is dominant and is approaching an asymptote considerably lower than that of the total $Q_{pl}+Q_{pt}$. Further calculations which have been made show that no fit is possible with $f_l>0.6$. and if the boundary scattering is not completely diffuse the limit will be even lower. However, values of $f_l L/L_0$ smaller than the 0.46 used for Fig. 10 are possible if the contribution Q_{pt} from transverse modes is assumed to come in at a high enough temperature to partially compensate the early bendover of Q_{pl}.

The high- and low-temperature extremes of Fig. 10 are worth noting. Even at 100° the theoretical curve of Q_{pl} is appreciably below the "infinite specimen" line, the difference corresponding to $\rho_l=0.89$. This illustrates the fact, implicit in (B5) of Appendix B, that at high T, where only a negligible proportion of the modes are dominated by boundary scattering, the effect of boundary scattering on the remaining (higher q) modes, though small, is far from negligible. The exponent η in $Q_p \propto T^{-\eta}$ over the range 80°–100° is reduced by boundary scattering from 3.50 to 3.03. Note that the reduction of η would have been twice as great if we had assumed simple additivity of $1/\tau(\mathbf{q})$ and $1/\tau_b$, instead of using the treatment of Eqs. (43) and (44).

At low T the observed points fall below the theoretical curves, whereas ideally they should pass above as the contribution Q_{pt} of the transverse modes starts to be appreciable. This may be at least partly due to experimental errors, which become larger in this range, or to inhomogenity of the sample. If the effect turns out to be real it will dictate choice of a smaller value of $f_l L/L_0$; however, too small a value would worsen the fit at higher T. According to the calculations presented below in Fig. 11, the saturation effect of Sec. VII should not have made a major reduction in the Q_p of this specimen by 25°.

Figure 7 of reference 2 shows the effect on Q of a moderate reduction in the dimensions of specimen 7. As the diameter reduction was not absolutely uniform, it is hard to estimate its effect precisely; however, the data given suggest that the cut down specimen had an effective L about 0.78 times the original value. Figure 11 gives a comparison of the observed change in Q produced by cutting down with the change in Q_{pl} computed from (46) and Fig. 6. At 40° and 30° the agreement is all that could be desired, and favors the belief that the numerical assumptions of Figures 10 and 11 are not far from the truth. However, the uncertainties in the

experimental data are considerable, and as data covering a much wider range of sizes will probably be available soon, it is hardly worthwhile to speculate about such things as the departures of the points from the curve at high and low T.

2. Nondegenerate p-Type Germanium, Larger n

Figure 4 of the paper of Geballe and Hull[2] shows that as the carrier concentration n increases, Q_p changes little at first, but eventually decreases markedly. This is undoubtedly the saturation effect of Sec. VII, and we shall discuss it in detail presently. The slight rise in the plotted Q_p occurring just before the fall may or may not be real. It is of the same order of magnitude as the correction which would be produced in Q_p by taking impurity scattering into account in the calculation of $\Delta\epsilon_T$ (see Sec. II). If any part of the rise is real, it may conceivably be due to breakdown of the assumption that impurity scattering and lattice scattering act independently, an assumption involved in our use of Eq.

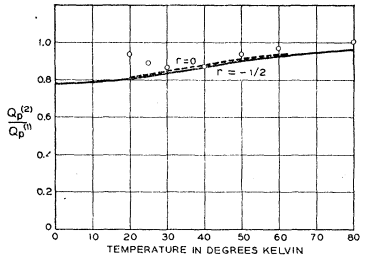

FIG. 11. Sample comparison of observed and calculated values of the change in Q_p due to reducing diameter of sample. Curves: calculated from Fig. 6 with numerical assumptions of Fig. 10. Points, Geballe and Hull, specimen 7, $L_0^{(2)}/L_0^{(1)}\approx 0.78$.

(21). At sufficiently large impurity density this speculation may be broadened into the statement that impurity band conduction is affecting both Q_e and Q_p.

The saturation effect can be predicted quantitatively from the theory of Sec. VII, provided we can adopt a value for the critical density n_1 defined by Eq. (57) and for the boundary scattering reduction factor ρ_l in Fig. 7. If T_0 is taken in the liquid air range, so that $Q_{pl}(T_0)$ is known, the only unknown quantity in Eq. (57) is the effective mass m^* entering into λ. For p germanium this cannot be far from the value $0.3m$, which is close to the inertial average mass[15] or to the larger of the two cyclotron-resonance masses.[43,44] As for ρ_l, which in Fig. 10 is

[44] It is clear that for the two-band model (see reference 33) the larger mass determines the effective average wave number q of the phonons which scatter the holes. The reason is that only transitions in which both initial and final hole states are of the small-mass type can give q's significantly smaller than those going with large-mass states of the same energy. But such transitions are rare, because only a small fraction of the holes are in small-mass states, and these are scattered predominantly into large-mass states, because the latter states are so much more plentiful than the former.

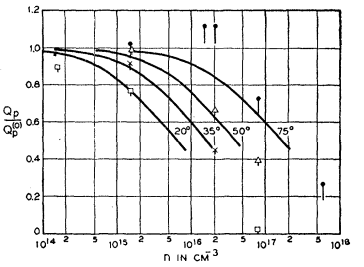

FIG. 12. Comparison of observed saturation ratios for p type germanium with the theory of Sec. VII. Curves: theoretical values obtained from Fig. 7 as described in the text. Points: observed values, from Fig. 4 of reference 2. The vertical line on each point gives a roughly estimated upper limit to the amount by which the "observed" ratios should be reduced to take account of the effect of impurity scattering in increasing the theoretical Q_e and thus reducing the "observed" Q_p. Ratios have been fitted to unity for $n=1.9\times10^{14}$ cm^{-3} except for the 20° curve, where theory and experiment were made to coincide at 1.5×10^{15}.

the ratio of the value given by the curve to that on the "infinite specimen" line, Fig. 10 gives the values 0.24–0.27 at 35°, 0.49–0.56 at 50°, and 0.77–0.82 at 75°, the first figure of each pair referring to $r=-\frac{1}{2}$, the second to $r=0$. Calculations similar to those of Fig. 10 but with other reasonable choices of the adjustable parameters have usually given values within a few hundredths of a unit of these, so it will be reasonable, for a rough analysis, to take the ρ_l values at 35° and 50° to be the same respectively as those used in Fig. 7, *viz.*, 0.28 and 0.54, and to interpolate Fig. 7 to $\rho_l=0.8$ for 75°. The theoretical curves so obtained are compared with experiment in Fig. 12, assuming $Q_p\approx Q_{pl}$.

Except for the 75° data, the agreement is very good indeed—a little too good, in fact, since practically all the minor effects which have not been incorporated into the theory ought to cause the present theoretical curves to be too far to the left. These effects include the contribution of transverse modes; the effect of saturation on the mobility of the electrons, i.e., $\beta_{ex}\neq\beta_{ex}^{(0)}$ in Eq. (53); departures from the assumption $r=0$; possible failure of the phonon scattering law to assume its asymptotic form in (33), resulting in an effective $s<0$ in Eq. (36); and the previously mentioned possible increase in the electron-lattice interaction in the presence of strong impurity scattering. The low ratio shown by the 20° point at the extreme left is probably due to some systematic error in the data for this specimen, as the points for this specimen at other temperatures (not shown) are out of line with the general trend of the other specimens.

While the present calculation with $m^*=0.3m$ gives reasonable agreement for the saturation effect, an m^* say twice as large would, by Eq. (57), require a shift of the calculated curves to the right by a factor $2^{\frac{3}{2}}$, and so would not fit. In view of the one-sided nature of most of the corrections which the theory is likely to need, it is probably safe to say that an $m^*\gtrsim 0.6m$ would be hard to reconcile with the data.

3. Degenerate p-Type Germanium

The most highly degenerate of the p-type specimens studied by Geballe and Hull was their No. 1, with $n=7.1\times10^{18}$ cm^{-3}. Their measured value, $Q=296$ μv/deg at 300°K, can be fitted to the conventional theory of Q_e, as presented in Fig. 8, by suitable choice of the density-of-states mass $m^{(N)}$. The required value is roughly $0.8m$ if τ_e is assumed independent of energy ($r=0$), $0.4m$ if $\tau_e \propto |\epsilon-\epsilon_b|$ ($r=1$). According to the discussion given in Sec. VIII, an effective r near 1 would be reasonable, but a value nearer to 0 is not excluded. It is interesting to see that the range of $m^{(N)}$ values suggested by the thermoelectric data includes the value found for high-resistivity material; however, the band structure for the degenerate material has been formed by the growth and coalescence of impurity bands, and might easily have a higher or lower $m^{(N)}$.

If we take this result as evidence that the conventional theory of Q_e is a fair approximation to the truth, we may compute Q_p by comparing Q with theoretical Q_e values at lower temperatures, fitting Q_e to Q at room temperature. This procedure gives, for the specimen mentioned, values from 160 to 220 μv/deg for $Q_p(60°)$, assuming the effective r to be between 0 and 1 and to increase or remain constant as the temperature is lowered. Such values accord with the predictions of Sec. VIII in being far below the value for high-resistivity material (1900–2000 μv/deg), but they are rather larger than expected. If at 60° we take $m^{(N)}=0.8m$, the more favorable of the possible room temperature values, we find from (A10) that in the absence of saturation the ratio of Q_p for specimen No. 1 to that for a nondegenerate specimen should be about 0.3 at 60°. The ratio of the Q_p values just quoted is at least a quarter of this, so the saturation ratio should be at least $\frac{1}{4}$. The curves of Fig. 9, though rough, suggest that the saturation ratio for the contribution of the longitudinal phonons alone should be a little less, of the order of 0.1 to 0.2. More complete data will be needed to decide whether this discrepancy is trivial or whether it arises from the transverse phonon contribution, from inadequacy of the conventional theory of Q_e, from incorrectness of the assumed band structure, or from inappropriateness of any band model for the calculation of Q_p. Concerning the underlying band structure, it is to be noted that for a given m^* a change in the number of valleys, e.g., from 6 to 1, has no effect on the predicted saturation effect in the nondegenerate range, but affects the saturation considerably in the degenerate range.

4. N-Type Germanium

The data of reference 2 on n-type germanium show the same qualitative features as for p-type, the only appreciable differences being a more gradual bending over at low temperatures and a smaller exponent η in the high-temperature law $Q_p \propto T^{-\eta}$. There are several conceivable causes for the small value 2.4 found for the apparent η:

(i) As was suggested above, the band structure may be of the degenerate many-valley type, with a fairly small spin-orbit splitting. This would be in line with the fact that Geballe and Hull[2] had to take $m^{(N)}\approx0.75m$ to fit their n-type Q_e around room temperature, whereas with a nondegenerate four-valley model one would expect $m^{(N)}/m=4^{\frac{2}{3}}(0.08^2\times1.3)^{\frac{1}{3}}=0.51$, with the cyclotron masses,[40] or $m^{(N)}/m>4^{\frac{2}{3}}\times0.6=1.5$, with the Benedict-Shockley mass.[15] If the suggested spin-orbit splitting were present, the theory of all these effects above liquid air temperature would require revision. In particular, the extrapolation of Q_e to the liquid air region would become uncertain, since both the position of the Fermi level and the term $\Delta\epsilon_T$ in Eq. (2) might depend on T in a way different from that assumed in Eqs. (4) and (7). In such a case, the resulting modification of Q_e might be in part responsible for the small exponent η. In addition, the temperature variation of Q_p might well be altered, since both the effective m^* and the effective f_l in our formulas could be changing with temperature.

(ii) As we have seen for p germanium, boundary scattering is capable of reducing the effective exponent η appreciably, even in the range 80°–100°K. This effect could considerably alleviate the η discrepancy for n germanium, even with a nondegenerate many-valley model, if the average effective mass m^* for n germanium were considerably smaller than that for p germanium. In such case the ξ_l of Eq. (46) would be considerably larger, and so the boundary scattering ratio ρ_l would depart more from unity at any given temperature. A value $\Delta\eta=d\ln\rho_l/d\ln T$ close to unity would not be inconceivable.

(iii) It is possible, of course, that the exponents s and γ in Eq. (36) depart appreciably from their ideal zero values, i.e., that the limiting form of the phonon-phonon scattering law for small T and $q\rightarrow0$ is not realized. In such case the larger η for p germanium would probably have to be attributed to something like spin-orbit complications.

Particularly interesting are the data on the reduction of Q_p due to cutting down the thickness of the specimen. From Fig. 7 of reference 2 it appears that when a specimen initially 0.16×0.148 cm was reduced to 0.16×0.025 cm over most of its length, Q_p was reduced in the ratio 0.83 at 60°K, 0.65 at 25°K. At first sight it seems surprising that the reduction in Q_p is not more pronounced: we might intuitively suppose that at 25° it should be comparable with the ratio of thicknesses. However, a study of the formulas of Casimir,[35] which determine the boundary scattering length L, shows that for a slab-shaped specimen L is much larger than the thickness, and in fact increases logarithmically without limit as the width increases, for fixed thickness. Although no exact expression for L is available, it seems reasonable that L for the cut-down specimen may have been large enough (0.3 to 0.4 of the original value) to

make possible a fit of the 60° result to the theory of Sec. VI, with $Lf_l/L_0 \lesssim 1$. The 25° result could be accounted for by such an L if the contribution Q_{pt} of the transverse phonons to Q_p is an appreciable fraction of the total at this temperature. However, there are other factors which should be looked into, among them the effect of the highly anisotropic effective mass, which will cause the function $R_x(\mathbf{q})$ to have a much broader peak that the function used in Sec. VI, hence to differ in its boundary scattering properties.

The data on the saturation effect for n germanium, Fig. 5 of reference 2, are less detailed than the data for p type, but are sufficiently similar in their behavior to justify the belief that they too can be roughly accounted for by the theory of Sec. VII.

X. PREDICTED BEHAVIOR OF OTHER SEMICONDUCTORS

The comparison of theory and experiment just given for germanium indicates that many interesting things can be learned about a semiconductor from a study of its thermoelectric power in the range where Q_p is appreciable. The planning of future experiments in this field may be facilitated by a rough prediction of the order of magnitude of the Q_p to be expected for a new material. For cubic materials, to which Eq. (38) applies, Q_p at any given T should be proportional to $c^2 f_l/\mu A_l m^*$, where c is the velocity of sound, μ the mobility, m^* the effective inertial mass defined in Appendix A, A_l the phonon-phonon scattering coefficient defined by (33), and f_l the fraction of the scattering of the carriers which is due to low-energy longitudinal phonons. The variation of A_l with the material can be estimated by dimensional means, if we are willing to assume the different materials to have the same ratio of anharmonic to harmonic elastic constants. Under this assumption A_l must depend only on $\hbar$, k, c, and the density ρ. Dimensional analysis gives

$$A_l \propto k^3/\hbar^2 c^3 \rho, \tag{64}$$

whence

$$Q_p \propto \rho c^5 f_l/\mu m^*. \tag{65}$$

Semiconductors belonging to the trigonal system, if suitable ones can be found, would be of especial interest, since according to reference 32 these should have an asymptotic phonon-phonon scattering law of the type in Eq. (36) with $s=1$, $\gamma=0$, and correspondingly should have rather large phonon-phonon relaxation times, hence large Q_p.

ACKNOWLEDGMENTS

I am much indebted to T. H. Geballe for frequent discussions of the experimental material. T. S. Benedict, B. Goodman, C. Kittel, and F. J. Morin have also provided useful material and suggestions.

I am grateful to the staff of the Institute for Advanced Study for their hospitality during the year 1952–1953.

APPENDIX A: CALCULATION OF Q_p FOR THE MANY-VALLEY MODEL, WITH FERMI STATISTICS

We wish to calculate the quantity $\mathbf{R}(\mathbf{q})$, defined at the start of Sec. IV, for the many-valley model of Fig. 1(b), and for Fermi statistics. To make the calculation tractable we assume that the combined effect of all scattering processes acting on the charge carriers can be described by a relaxation time $\tau_e(\epsilon)$ which is a function of energy only. This can be shown to be justified for any intervalley scattering and for intra-valley lattice scattering provided the phonon energy can be neglected and provided only the volume dilatation contributes to the deformation potential.[45] However, use of a $\tau_e(\epsilon)$ is only roughly correct if shearing distortions contribute, as they usually will.

The total $\mathbf{R}(\mathbf{q})$ which we wish to calculate is a sum of contributions from the different valleys surrounding the different band edge points $\mathbf{K}_\nu$. For a cubic crystal, however, the average of $\mathbf{E}\cdot\mathbf{R}(\mathbf{q})$ over the different valleys and over directions of $\mathbf{q}$ is the same as the average, for a single valley, over all directions of $\mathbf{q}$ and the electric field $\mathbf{E}$. Let us therefore fix attention on one of the valleys, say that around $\mathbf{K}_1$. Let the coordinate axes be chosen parallel to the principal axes of the ellipsoidal energy surfaces in this valley. The variation of energy with $\mathbf{K}$ in this valley is given by

$$|\epsilon_K - \epsilon_b| = (\hbar^2/2m)(\alpha_x \Delta K_x^2 + \alpha_y \Delta K_y^2 + \alpha_z \Delta K_z^2), \tag{A1}$$

where ϵ_b is the band edge energy, $\alpha_i = m/m_i^*$ are effective mass ratios for different directions, and $\Delta\mathbf{K} = \mathbf{K} - \mathbf{K}_1$. With this notation, the first-order distribution function in this one valley is given by the analog of Eq. (20),

$$f^{(1)}(\mathbf{K}) \propto \tau_e(\epsilon_K)(\sum E_i \alpha_i \Delta K_i) df^{(0)}(\epsilon_K)/d\epsilon_K, \tag{A2}$$

where $f^{(0)}$ is the Fermi function.

We shall consider only the case in which the phonon energies are negligible compared with the electron energies and with kT. For this case one gets the same result by neglecting the exclusion principle in the transition probabilities as by including it through factors $(1-f)$. Therefore the first two lines of Eq. (22) remain valid for the contribution of each valley in the present problem. When we come to insert Eq. (21) in Eq. (22), however, we meet the complication that the matrix element M is in general dependent on the direction of $\mathbf{q}$, since the deformation potential for the valley at $\mathbf{K}_1$ in general depends on a particular component of the shearing strain as well as on the volume dilatation. While it is not quite consistent to take this anisotropy of M into account while assuming a relaxation time dependent only on energy, we shall do this here, partly because the analysis for this case is fairly easy, and

[45] These arguments will be given in a forthcoming publication (reference 19). It can also be shown that for a many-valley model the shortcomings of the assumption of existence of a $\tau(\epsilon)$ can be corrected, to a fair approximation, by use of three relaxation times, a procedure roughly equivalent to letting τ depend on position over an energy surface, as we did in (5) and (6) of the text.

partly because anisotropy of M is likely to be more important than failure of the relaxation time assumption.

It is convenient to associate with any vector $\mathbf{V}$ a primed vector $\mathbf{V}'$, defined by $V_i'=\alpha_i^{\frac{1}{2}}V_i$. Then (A2) can be written

$$f^{(1)}\propto\tau_\epsilon(\Delta K')(\mathbf{E}'\cdot\Delta\mathbf{K}')[df^{(0)}/d\epsilon]_{\epsilon=\hbar^2\Delta K'^2/2m}. \quad \text{(A3)}$$

By steps like those used in (22) to (25) of the text we find

$$\mathbf{E}\cdot\mathbf{R}_1(\mathbf{q})=\frac{G(\mathbf{E}\cdot\mathbf{q})(\mathbf{E}'\cdot\mathbf{q}')}{q'}\times\int_{q'/2}^{\infty}\tau_\epsilon(\Delta K')\frac{df^{(0)}}{d\epsilon}\Delta K'd\Delta K', \quad \text{(A4)}$$

where G is a function of the direction of $\mathbf{q}$, proportional to $|M|^2$. Averaging of (A4) over directions of $\mathbf{E}$ replaces the factor in front of the integral by $Gq'E^2/3$.

The averaging over directions of $\mathbf{q}$ is not easy to carry out explicitly on Eq. (A4). However, if we assume $\tau(\mathbf{q})$ to obey Eq. (36), we may evaluate the left sides of Eqs. (26) and (19), as in the text, by interchanging the order of integration on $\mathbf{K}$ and $\mathbf{q}$. With

$$d\mathbf{q}=(\alpha_x\alpha_y\alpha_z)^{-\frac{1}{2}}q'^2dq'd\Omega, \quad \text{(A5)}$$

we find

$$(2\pi)^{-3}\int\mathbf{E}\cdot\mathbf{R}(\mathbf{q})d\mathbf{q}=\frac{2}{3\pi^2}\frac{N_VE^2\langle G\rangle'}{(\alpha_x\alpha_y\alpha_z)^{\frac{1}{2}}}\times\int_0^{\infty}\Delta K'^5\langle\tau_\epsilon(\Delta K')\frac{df^{(0)}}{d\epsilon}d\Delta K' \quad \text{(A6)}$$

$$(2\pi)^{-3}\int\tau(\mathbf{q})\mathbf{E}\cdot\mathbf{R}(\mathbf{q})d\mathbf{q}=\frac{2^{2-s}}{6\pi^2(2-s)}\frac{N_VE^2\langle G(q'/q)^{2+s}\rangle'}{AT^{3-s-\gamma}(\alpha_x\alpha_y\alpha_z)^{\frac{1}{2}}}\times\int_0^{\infty}\Delta K'^{3-s}\tau_\epsilon(\Delta K')\frac{df^{(0)}}{d\epsilon}d\Delta K', \quad \text{(A7)}$$

where N_V is the number of band edge points $\mathbf{K}_\nu$, and $\langle\ \rangle'$ denotes an average over directions of $\mathbf{q}'$. For phonons of branch α ($\alpha=l$ or t), Eq. (A6) may be set equal to $\mathbf{E}\cdot\mathbf{J}f_\alpha/\mu$, where f_α is the fraction of the crystal momentum given up by the carriers which is delivered to low energy modes of this branch. Combining this with Eqs. (A7), (19), and (11) we find for the contribution of phonons of this branch to the thermoelectric power

$$Q_{p\alpha}=\frac{P_\alpha\bar{c}_\alpha^2}{JT}=\frac{f_\alpha\bar{c}_\alpha^2\langle G_\alpha(q'/q)^{2+s}\rangle'}{2^s(2-s)\mu AT^{4-s-\gamma}\langle G_\alpha\rangle'}\times\frac{\int_0^{\infty}\Delta K'^{3-s}\tau_\epsilon(\Delta K')(df^{(0)}/d\epsilon)d\Delta K'}{\int_0^{\infty}\Delta K'^5\tau_\epsilon(\Delta K')(df^{(0)}/d\epsilon)d\Delta K'}. \quad \text{(A8)}$$

Let us specialize first to Maxwellian statistics. For each direction of $\mathbf{q}$ the expression (A4) is of the same form as Eq. (25), so the present $R_x(\mathbf{q})$ will be an average of curves like the full ones in Fig. 5, the average being taken over a family of curves of the same shape but different horizontal scales. Thus the peak of the final $R_x(\mathbf{q})$ curve will be broader and flatter than in Fig. 5; this difference in shape will disappear as α_x, α_y, α_z approach equality. Now in applications like those of Secs. V and VI, we wish to approximate $\mathbf{R}(\mathbf{q})$ by the equations of the text, which involve m^* or λ. A good choice for this parameter is the one which makes E times Eq. (37) equal Eq. (A7). With this choice the approximating $\mathbf{R}(\mathbf{q})$ has the same integral and the same $(-2-s)$th moment as the correct one. Equating the ratio of Eq. (37) to Eq. (26) to the ratio of Eq. (A7) to Eq. (A6), we find

$$m^*=m\left[\frac{\langle G'\rangle}{\langle G(q'/q)^{2+s}\rangle'}\right]^{2/(2+s)}. \quad \text{(A9)}$$

This m^* always lies between the minimum and the maximum of the effective masses $m_i^*=m/\alpha_i$ in the three principal directions. Numerical evaluation of Eq. (A9) has shown that if G is isotropic m^* is usually quite close to the geometric mean of the m_i^*, i.e., to $m^{(N)}/N_P^{\frac{2}{3}}$. For the maximum possible anisotropy of G, on the other hand, m^* may be closer to one of the extremes m_i^* than to the geometrical mean. With the identification in Eq. (A9), Eq. (A8) coincides with Eq. (38). For the general many-valley model, f_α/μ is not rigorously given by an expression like Eq. (31), but a study of the analogs to Eqs. (27) and (29) shows that the variation of f_α/μ with the scattering exponent r must be very similar to that used in Eq. (39).

In the limit of extreme degeneracy $df^{(0)}/d\epsilon$ becomes a delta function and Eq. (A8) becomes, with Eqs. (A9) and (40),

$$Q_{p\alpha}=\frac{f_\alpha\bar{c}_\alpha^2\lambda^{2+s}}{2^s(2-s)\mu A_\alpha T^{4-s-\gamma}}\left(\frac{kT}{\epsilon_F}\right)^{1+\frac{1}{2}s}. \quad \text{(A10)}$$

The significance and limitations of this equation are discussed in Sec. VIII. Note that for the present case f_α/μ exceeds the value in nondegenerate material in the ratio $(16\epsilon_F/9\pi kT)^{\frac{1}{2}}$.

APPENDIX B: EVALUATION OF THE BOUNDARY SCATTERING RATIO ϱ

We would like to evaluate

$$\rho=\int\int\int\Delta P_x dydzd\mathbf{q}\Big/\int\int\int\Delta P_{0x}dydzd\mathbf{q}, \quad \text{(B1)}$$

where y and z run over the cross section of a cylindrical specimen of any specified shape with its axis in the x direction and where $\Delta\mathbf{P}$ is the density of crystal mo-

mentum in $\mathbf{q}$ and coordinate space, as defined in Sec. VI, and ΔP_0 is the same in the absence of boundary scattering.

We shall assume for the present that the boundary scattering is completely diffuse. Then the argument of the text, leading to Eq. (44), can be applied to any shape of cross section provided y is defined as the distance from the back surface, measured in a direction normal to the specimen axis and coplanar with this axis and $\mathbf{q}$, so that $q_z=0$. Let us define $1/\tau_b(\mathbf{q},T)$ to be the ratio of the rate of destruction of $\Delta\mathbf{P}$ by boundary scattering to the integral of $\Delta\mathbf{P}$ over the cross section. If $y_1(z)$ is the maximum value of y, i.e., the thickness of the specimen at z,

$$1/\tau_b(\mathbf{q},T)=(cq_y/q)$$
$$\times\int\Delta P_x(\mathbf{q};y_1,z)dz\Big/\int\Delta P_x(\mathbf{q};y,z)dydz, \quad \text{(B2)}$$

where $\Delta\mathbf{P}$ is given by Eq. (44). This τ_b is, of course, the τ_b which has to be combined with $\tau(\mathbf{q})$ by the method of adding reciprocals, in order to get the effective relaxation time for mode $\mathbf{q}$. The difference between the present theory and the usual one lies in the fact that Eq. (B2) depends on temperature, i.e., on $\tau(\mathbf{q})$. At low temperatures $\tau(\mathbf{q})\to\infty$ and the integrals in Eq. (B2) simplify; at high temperatures $\tau(\mathbf{q})\to 0$ and they simplify again. One finds easily

$$\tau_b(\tau\to 0)/\tau_b(\tau\to\infty)=2\Big(\int y_1dz\Big)^2\Big/\int dz\int y_1^2dz. \quad \text{(B3)}$$

For the slab this has the value 2 mentioned in the text; for a circular cylinder it is $3\pi^2/16=1.85$; for a square it varies from 1.5 to 2, depending on the direction of $\mathbf{q}$. Thus for all these cases the mean value of Eq. (B3) over directions is within 10 percent or so of 2. This suggests the following approximation for the evaluation of Eq. (B1): Let us choose an average $\bar{y}_1$ of $y_1(z)$ as the limit on y in Eq. (B1), in such way that at low T ($\tau\to\infty$) we get the right ρ. Then use of this same $\bar{y}_1$ should give $1-\rho$ correct to 10 percent or so at high $T(\tau\to 0)$. It is reasonable to expect the behavior at intermediate T to be nearly correct also.

To determine $\bar{y}_1$, we note that the correct ρ for $\tau\to\infty$ corresponds to a weighted average $\bar{\tau}_b$, with weight q_x^2/q^2, equal to L/c, where L is the Casimir length introduced in the text. For a circular cylinder we have $q_x/q=\cos\theta$, $q_y/q=\sin\theta$, and if $y_1=\bar{y}_1$ is taken independent of z in Eq. (B2), then for $\tau\to\infty$,

$$\tau_b^{-1}=2c\sin\theta/\bar{y}_1,$$
$$\bar{\tau}_b=3\pi\bar{y}_1/8c=L/c. \quad \text{(B4)}$$

In the present approximation the value of Eq. (B1) for longitudinal modes becomes, by (44), (45), and (33),

$$\rho\approx\frac{\Big\langle 3\cos^2\theta\int_0^\infty\Big\{1-\frac{q_l^2\sin\theta}{q^2}\Big[1-\exp\Big(\frac{-q^2}{q_l^2\sin\theta}\Big)\Big]\Big\}F(q)dq\Big\rangle}{\int_0^\infty F(q)dq}, \quad \text{(B5)}$$

where the angular brackets denote an average over directions of $\mathbf{q}$,

$$q_l^2=3\pi c/8LA_lT^3, \quad \text{(B6)}$$

and

$$F(q)\propto\bar{R}_x(q)=\lambda q[1-\Phi(\lambda q/2)] \quad (r=\tfrac{1}{2}) \quad \text{(B7)}$$
$$=\lambda q\exp(-\lambda^2q^2/4), \quad (r=0) \quad \text{(B8)}$$

etc. (see Fig. 5), Φ being the error function and λ being defined by Eq. (40). For $r=-\frac{1}{2}$ and 0 the integral on q in (B5) can be evaluated analytically:

$$r=-\tfrac{1}{2}:\rho\approx 1-\langle 3\cos^2\theta\cdot 4\xi_l\sin\theta\ln(\tfrac{1}{2}+\tfrac{1}{2}[1+(\xi_l\sin\theta)^{-1}]^{\frac{1}{2}})\rangle, \quad \text{(B9)}$$

$$r=0:\rho\approx 1-\langle 3\cos^2\theta\cdot\xi_l\sin\theta\ln[1+(\xi_l\sin\theta)^{-1}]\rangle, \quad \text{(B10)}$$

where

$$\xi_l=\lambda^2q_l^2/4 \quad \text{(B11)}$$

has the value given by Eq. (46). The full curves for $r=-\frac{1}{2}$ and 0 in Fig. 6 were constructed by evaluating the angular averages in (B9) and (B10) numerically. The curve for $r=1$ was estimated from these two and the rigorously known limiting behavior at large and small ξ_l.

For transverse modes we may evaluate Eq. (B1) similarly, using (34) instead of (33) for $\tau(\mathbf{q})$ in Eqs. (44) and (45). The result is, for $r=0$

$$\rho\approx 1-\langle 3\cos^2\theta\{1-\Phi[(\xi_l\sin\theta)^{-1}]\}\times\exp[(\xi_l\sin\theta)^{-2}]\rangle, \quad \text{(B12)}$$

where ξ_l is given by Eq. (47). The dashed curve in Fig. 6 was again obtained by numerical evaluation of the angular average.