PHYSICAL REVIEW VOLUME 102, NUMBER 5 JUNE 1, 1956

Theory of Transport Effects in Semiconductors: The Nernst Coefficient, and Its Relation to Thermoelectric Power

P. J. Price
Watson Scientific Computing Laboratory, (International Business Machines Corporation), Columbia University, New York, New York

(Received January 13, 1956)

A general expression for the Nernst coefficient, in terms of phenomenological parameters of the band transport equations, is derived for a two-band semiconductor for which Boltzmann statistics may be applied to the carrier distributions. The expression reduces to that given by Putley when his assumptions are made. The information obtainable from experimental values of the Nernst coefficient is discussed in the light of the expression derived in the paper. Also discussed is how this information might be supplemented from experimental values of the thermoelectric power. An expression for the latter, in terms of the phenomenological band transport parameters, is given. It reduces to that given by Johnson and Lark-Horovitz, when their assumptions are made. It is pointed out that the phonon drag effect should cause an anomaly in the Nernst coefficient corresponding to that in the thermoelectric power, and an estimate of the former anomaly is made.

IN a recent paper[1] (to be referred to here as I) concerning ambipolar diffusion of the electrons and holes down a temperature gradient in a semiconductor, an expression was derived for the contribution of the ambipolar flow to the Nernst effect. In the present paper this calculation is extended to give an expression, in terms of phenomenological parameters of the band transport equations, for the complete Nernst coefficient. This expression reduces to that given by Putley[2] when his special assumptions are made. The information obtainable from measurement of the Nernst coefficient is discussed in the light of the expression derived here. Also discussed is how this information might be supplemented by experimental values of the thermoelectric power. An expression for the latter, in terms of the same band transport-equation parameters, is derived. It is a generalization of the result given by Johnson and Lark-Horovitz.[3] It is pointed out that the "phonon drag effect"[4] should cause an anomaly in the low-temperature Nernst coefficient comparable to that which it causes in the thermoelectric power, and a rough estimate of the anomaly in the former (for *p*-type germanium at 50°K) is given.

1. THE TRANSPORT EQUATIONS

We confine ourselves to a cubic semiconductor. For a homogeneous single cubic crystal, the relevant phenomenological bulk transport equation, to terms linear in the transport vectors and magnetic field, is

$$\mathbf{E}=\rho\mathbf{J}+R\mathbf{H}\times\mathbf{J}-\Omega\,\mathrm{grad}T-B\mathbf{H}\times\mathrm{grad}T, \tag{1}$$

where $\mathbf{E}$, $\mathbf{J}$, $\mathbf{H}$, and T have their usual meanings. Of the four coefficients, ρ, R, and B are, respectively, the resistivity and the Hall and Nernst coefficients. The fourth coefficient, Ω, can be shown to be related to the thermoelectric power, Q, by Eq. (27) of Sec. 3. To find expressions for ρ, R, B, and Ω, we calculate the current density $\mathbf{J}$ due to an electric field and temperature gradient acting together, and compare the result with the solution of (1) (again, to linear terms only) for $\mathbf{J}$:

$$\rho\mathbf{J}=\mathbf{E}+\Omega\,\mathrm{grad}T-\mu^H\mathbf{h}\times\mathbf{E}+(cB-\mu^H\Omega)\mathbf{h}\times\mathrm{grad}T, \tag{1'}$$

where $\mu^H\equiv cR/\rho$ is the Hall mobility and $\mathbf{h}\equiv\mathbf{H}/c$. For a two-band semiconductor

$$\mathbf{J}=e(n_2\mathbf{u}_2-n_1\mathbf{u}_1), \tag{2}$$

where the n_s are the carrier concentrations and the $\mathbf{u}_s$ the drift velocities in the bands. The subscript labels the band: $s=1$ for electrons, $s=2$ for holes. We calculate the $\mathbf{u}_s$ subject to the same assumptions[5] as in I (Sec. 1). In the absence of a magnetic field,

$$-\mathbf{u}_s=\pm\mu_s\mathbf{E}+D_s\,\mathrm{grad}(\log n_s)+D_s{}^T\,\mathrm{grad}T,$$

where the upper sign applies for electrons and the lower one for holes. (This meaning is to be given to the symbols $\pm$, $\mp$ wherever they occur in this paper, except that in an equation applying to an extrinsic crystal the meaning might be stated more clearly by saying that the upper sign refers to *n*-type and the lower to *p*-type.) As in I, we substitute the Einstein formula

$$D_s=kT\mu_s/e \tag{3}$$

for the diffusion coefficient, and set

$$D_s{}^T\equiv\gamma_sD_s/T=\gamma_s\mu_sk/e. \tag{4}$$

Then

$$\mathbf{u}_s=-\mu_s[\pm\mathbf{E}+(kT/e)\,\mathrm{grad}(\log n_s)+(k\gamma_s/e)\,\mathrm{grad}T]. \tag{5}$$

The effect of the magnetic field (in the linear limit which we are considering) on each of the three terms

---

[1] P. J. Price, Phil. Mag. 46, 1252 (1955).

[2] E. H. Putley, Proc. Phys. Soc. (London) B68, 35 (1955), Table 2.

[3] V. A. Johnson and K. Lark-Horovitz, Phys. Rev. 92, 226 (1953).

[4] H. P. R. Frederikse, Phys. Rev. 92, 248 (1953); C. Herring, Phys. Rev. 96, 1163 (1954).

[5] It should be observed that for silicon (except for *n*-type extrinsic) assumption (*d*) should be valid only below room temperature, according to the estimate of 0.035 ev [see for example F. Herman, Proc. Inst. Radio Engrs. 43, 1703 (1955) Fig. 9] for the spin-orbit splitting of the valence band at the zone center.

of (5) is to rotate the component normal to **H** by an angle of H/c times the corresponding "mobility." For the first term of (5) the latter is the band Hall mobility, $\mu_s{}^H$. For the second term it is also $\mu_s{}^H$, as was shown in I. For the third term the "mobility" has in general a different value, which we denote by the symbol $\lambda_s{}^H$. Then

$$\begin{aligned}\mathbf{u}_s=-\mu_s\{&\pm\mathbf{E}+\mu_s{}^H\mathbf{h}\times\mathbf{E}\\&+(kT/e)[\mathrm{grad}(\log n_s)\pm\mu_s{}^H\mathbf{h}\times\mathrm{grad}(\log n_s)]\\&+(k\gamma_s/e)[\mathrm{grad}T\pm\lambda_s{}^H\mathbf{h}\times\mathrm{grad}T]\}.\end{aligned}\tag{6}$$

For the concentration n_s in (6) we substitute the equilibrium value $n_s(T)$. Hence

$$\begin{aligned}\mathbf{u}_s=-\mu_s\{&\pm\mathbf{E}+\mu_s{}^H\mathbf{h}\times\mathbf{E}\\&+(k/e)[(Td(\log n_s)/dT+\gamma_s)\ \mathrm{grad}T\\&\pm(\mu_s{}^HTd(\log n_s)/dT+\lambda_s{}^H\gamma_s)\mathbf{h}\times\mathrm{grad}T]\}.\end{aligned}\tag{6'}$$

Setting

$$\alpha_s\equiv Td(\log n_s)/dT+\gamma_s,\quad \beta_s\equiv\gamma_s(\lambda_s{}^H-\mu_s{}^H),$$
$$\sigma_s\equiv e\mu_s n_s,\tag{7}$$

substitution of (6′) into (2) gives

$$\begin{aligned}\mathbf{J}=&(\sigma_1+\sigma_2)\mathbf{E}+(k/e)(\sigma_1\alpha_1-\sigma_2\alpha_2)\ \mathrm{grad}T\\&-(\sigma_2\mu_2{}^H-\sigma_1\mu_1{}^H)\mathbf{h}\times\mathbf{E}\\&+(k/e)[\sigma_1(\mu_1{}^H\alpha_1+\beta_1)+\sigma_2(\mu_2{}^H\alpha_2+\beta_2)]\mathbf{h}\times\mathrm{grad}T.\end{aligned}\tag{8}$$

By comparison of (8) with (1′) we find the standard results

$$1/\rho\equiv\sigma=\sigma_1+\sigma_2,\quad \mu^H=(\sigma_2\mu_2{}^H-\sigma_1\mu_1{}^H)/\sigma,\tag{9}$$

and the further results

$$\Omega=(k/e)(\sigma_1\alpha_1-\sigma_2\alpha_2)/\sigma,\tag{10}$$

and [making use of (9)]

$$\begin{aligned}B=(k/ec)[&(\sigma_1\sigma_2/\sigma^2)(\mu_1{}^H+\mu_2{}^H)(\alpha_1+\alpha_2)\\&+(\sigma_1\beta_1+\sigma_2\beta_2)/\sigma].\end{aligned}\tag{11}$$

2. NERNST COEFFICIENT

The formula (11) is conveniently written

$$\begin{aligned}B=(k/ec)[\beta_{12}\sigma_1\sigma_2/\sigma^2+\beta_1\sigma_1/\sigma+\beta_2\sigma_2/\sigma]\\ \equiv B_{12}+B_1+B_2,\end{aligned}\tag{12}$$

where

$$\begin{aligned}\beta_{12}\equiv&(\mu_1{}^H+\mu_2{}^H)(\alpha_1+\alpha_2)\\=&(\mu_1{}^H+\mu_2{}^H)[Td(\log n_1n_2)/dT+\gamma_1+\gamma_2],\end{aligned}$$

and the β_s are defined in (7). The assumptions on which (10) and (11) depend include the assumption that n_1 and n_2 are small enough for Boltzmann statistics to be applicable.[6] We then have, by Eq. (28) of Sec. 3,

$$n_1n_2=\mathrm{const.}T^3\exp(-\Delta/kT),\tag{13}$$

where $\Delta\equiv\epsilon_{01}-\epsilon_{02}$ is the energy gap between the bands, and hence[7]

$$\beta_{12}=(\mu_1{}^H+\mu_2{}^H)[\Delta/kT+3+\gamma_1+\gamma_2].\tag{12'}$$

The term of (12) proportional to $\sigma_1\sigma_2/\sigma^2$ (i.e. B_{12}), with the expression (12′) for β_{12}, was derived in I (it is $B-B_0$, Eq. (21), there) and a comparison was made with the formula for B put forward by Putley,[2] which was evidently derived by the conventional assumptions of spherical energy surfaces and of scattering of the carriers by Debye longitudinal lattice modes only. (For brevity, we will refer to these assumptions as defining "the Classical Model.") It was noted in I that for the classical model, since $\mu_s{}^H=3\pi\mu_s/8$ and $\gamma_s=\frac{1}{2}$, B_{12} reduced to the corresponding term of Putley's formula [i.e., to the first term of Eq. (24) in I]. The remaining terms, B_1+B_2, of (12) were not then available for comparison with the remaining term of Putley's result. It can be shown that for the classical model $\lambda_s{}^H=0$. (This result follows from Eq. (16) of the text: see the discussion following Eq. (18).) With this further substitution, (12) and (12′) reduce to a formula for B which is identical with Putley's formula. Thus Eq. (12), with β_{12} given by (12′) and the β_s by (7), provides a formula for the Nernst coefficient, in terms of phenomenological transport-equation parameters of the bands, for a model more general and more realistic than the classical one.

To make clearer the relation of (12) to Putley's formula, it may be helpful to give the expressions for the phenomenological parameters μ_s, $\mu_s{}^H$, γ_s, $\lambda_s{}^H$ which result from assuming a model of generality intermediate between that of this paper and that of the "classical model": namely, in which each band has spherical energy surfaces and a carrier scattering kernel $S(\mathbf{p},\mathbf{p}')$ sufficiently symmetrical for the transport parameters to be expressed in terms of a relaxation time $\tau(v)$ (where $\mathbf{v}$ is the carrier velocity), but in which the form of the function $\tau(v)$ is not limited to that for the classical model, $\tau=\mathrm{const}/v$. The theory of the transport parameters for this "intermediate model" is outlined in the appendix. The results are[8]

$$\mu_s=(e/m_s)(2,1)_s/(2,0)_s,\tag{14}$$

$$\mu_s{}^H=(e/m_s)(2,2)_s/(2,1)_s,\tag{15}$$

$$\lambda_s{}^H=\left(\frac{e}{m_s}\right)\frac{(2,0)_s(2,2)_s-(4,2)_s}{(2,0)_s(2,1)_s-(4,1)_s},\tag{16}$$

$$\gamma_s=\frac{3}{2}\left[\frac{(4,1)_s}{(2,0)_s(2,1)_s}-1\right].\tag{17}$$

Here m_s is the band effective mass and $(i,j)_s$ means the

---

[6] This assumption (g) of I is implicit in the use of the Einstein relation (3) between D_s and μ_s, and hence in Eq. (5). For the "soft" semiconductors (those with a gap Δ of a few tenths of an ev or less) especially, this assumption can effectively limit the applicability of the results of the present paper.

[7] We do not take into account here the variation with temperature of the effective energy gap. See note (e) added in proof.

[8] The possible infinite value of $\lambda_s{}^H$ implied by (16) has no physical significance, because $\lambda_s{}^H$ occurs in the transport equations only in the combination $\gamma_s\lambda_s{}^H$, which remains finite when the denominator of (16) vanishes.

average, over the (Maxwell-Boltzmann) distribution of carriers in the band, of $v^i\tau_s{}^j$:

$$(i,j)_s \equiv \langle v^i \tau^j \rangle_s. \tag{18}$$

For the classical model, $(i,j)_s = \langle v^{i-j} \rangle_s (v\tau_s)^j$. Then the numerator of (16) vanishes, and hence $\lambda_s{}^H = 0$. The results quoted for $\mu_s{}^H$ and γ_s for the classical model follow in a similar way. From (16) and (17), we have for the β_s of (12)

$$\beta_s = \frac{3\mu_s{}^H}{2}\left[\frac{(4,2)_s}{(2,0)_s(2,2)_s} - \frac{(4,1)_s}{(2,0)_s(2,1)_s}\right]. \tag{19}$$

It follows from (19) that if $\tau_s = \text{const}\ v^{r_s}$ then the sign of β_s is the sign of r_s, and that if τ_s is independent of v then $\beta_s = 0$.

From (12) it follows (subject to the same general conditions as for the "Dunlap ellipse"[9]: that the relevant band mobilities and the forbidden gap Δ, and for our case the γ_s also, remain independent of doping over the range in question) that if the conductivity is varied at constant temperature by doping then the curve of B *versus* μ^H should be a parabola, with the vertex the maximum of B. This result was obtained in I for $B_{12}(\mu^H)$. For the complete expression (12) it still holds, the parabola being displaced (without rotation) by the extra terms B_1+B_2. (See Fig. 1.) By solving Eq. (9) for σ_1/σ and σ_2/σ and substituting the results in (12), we obtain an expression for B in terms of μ^H:

$$\begin{aligned}(ec/k)(\mu_1{}^H+\mu_2{}^H)^2B = &-\beta_{12}(\mu^H)^2 \\ &+[\beta_{12}(\mu_2{}^H-\mu_1{}^H)+(\beta_2-\beta_1)(\mu_1{}^H+\mu_2{}^H)]\,\mu^H \\ &+[\beta_{12}\mu_1{}^H\mu_2{}^H+(\beta_1\mu_2{}^H+\beta_2\mu_1{}^H)(\mu_1{}^H+\mu_2{}^H)].\end{aligned} \tag{20}$$

The ends of the doping curve are at the points

$$\mu^H = \mp\mu_s{}^H, \quad B = (k/ec)\beta_s. \tag{21}$$

For the parabola $B_{12}(\mu^H)$, the vertex (maximum of B_{12}) is at the conductivity minimum, where $\sigma_1 = \sigma_2$. The effect of the extra terms, B_1+B_2, in (12) is to shift the vertex of the new parabola $B(\mu^H)$ to the point where

$$\beta_{12}(\sigma_1-\sigma_2) = \sigma(\beta_1-\beta_2). \tag{22}$$

For germanium and silicon this displacement should normally be small because the Δ/kT term of (12′) makes β_{12} large compared with β_1 and β_2. As an alternative to (20), we may express (12) as a formula for B in terms of

$$x \equiv (\sigma_0/\sigma)^2, \tag{23}$$

where σ_0 is the minimum value of σ for the given temperature. Then

$$B = (k/ec)[\tfrac{1}{4}\beta_{12}\,x \pm \tfrac{1}{2}(\beta_1-\beta_2)(1-x)^{\frac{1}{2}} + \tfrac{1}{2}(\beta_1+\beta_2)]. \tag{24}$$

By means of (20) or (24) it is possible (for the temperature range in which the purest available crystals are intrinsic or nearly so, and have coefficients β_1, β_2, and β_{12} not appreciably altered from their intrinsic-state values by the levels of impurity concentrations present) to determine the β coefficients of (12): that is, Δ being known, the combinations $\gamma_1(\lambda_1{}^H-\mu_1{}^H)$, $\gamma_2(\lambda_2{}^H-\mu_2{}^H)$, $\gamma_1+\gamma_2$ of the four unknowns γ_1, γ_2, $\lambda_1{}^H$, $\lambda_2{}^H$.[10] To obtain the separate values of the latter, a fourth independent combination must be determined. Equation (10) indicates that the thermoelectric power involves the γ_s, and therefore that there is the possibility of obtaining the values of all four unknowns by combined measurements of the Nernst coefficient and thermoelectric power. This possibility is investigated in the following section.

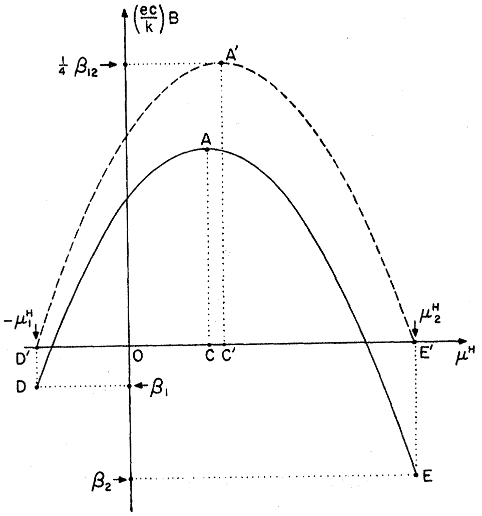

FIG. 1. The Nernst coefficient-Hall mobility parabola. The full curve is for the complete expression (12). (β_1 and β_2 have been assumed to be negative.) The dashed curve is for the first term, B_{12}, alone. The vertex of the latter, A', is at the conductivity minimum. Hence $\mathrm{OC'} = (\mu_2{}^H-\mu_1{}^H)/2$ and $4\mathrm{C'A'}/\mathrm{D'E'} = \Delta/kT + 3 + \gamma_1 + \gamma_2$.

The values of B to be expected, according to (12), when $B \simeq B_{12}$ (that is, when $B_{12} \gg B_1+B_2$) have been discussed in I. For Ge and Si, B_{12} should be a good approximation to B for intrinsic or near-intrinsic specimens,[5] except at the highest temperatures (near the melting points); but for the "soft" semiconductors, such as were studied by Putley, B_1+B_2 is never really negligible compared to B_{12}. For extrinsic specimens, on the other hand, we have

$$\left.\begin{aligned} B \simeq B_1 \simeq (k/ec)\beta_1 \equiv (k/ec)\gamma_1(\lambda_1{}^H-\mu_1{}^H), \quad &(n\text{-type}) \\ B \simeq B_2 \simeq (k/ec)\beta_2 \equiv (k/ec)\gamma_2(\lambda_2{}^H-\mu_2{}^H). \quad &(p\text{-type}) \end{aligned}\right\} \tag{25}$$

At the lowest temperatures, (25) may of course be taken as an equality. To get an idea of what order of magni-

---

[9] W. C. Dunlap, Phys. Rev. **79**, 286 (1950), Sec. 5.

[10] It is assumed that the $\mu_s{}^H$ are known, since they may be obtained by means of the Dunlap ellipse. (If they are not known the β's may still be obtained by means of (24), but then the β's still involve the $\mu_s{}^H$.) Even if β_{12} and the $\mu_s{}^H$ are known accurately, the combination $\gamma_1+\gamma_2$ is thereby determined with an uncertainty equal to the uncertainty in Δ divided by kT: this limitation is likely to be a serious one in practice.

tudes (25) predicts, we express k/ec in practical units:

$$k/ec=0.862\times10^{-12}\ \text{volt}^2\ \text{sec/cm}^2\ \text{gauss degree.} \tag{26}$$

If the combination of mobilities in (25) has a (positive or negative) value of order 10^4 cm²/volt sec, then we expect

$$|B|\sim10^{-8}\binom{\gamma_1}{\gamma_2}\ \text{volt/gauss degree.}$$

Unless the phonon drag effect is important (see Sec. 4), γ_1 and γ_2 can be expected to be of order unity; so $|B|\sim10^{-8}$ in practical units. (In estimating the available voltages from this result, it must be remembered that values of H should be less than those at which effects quadratic in H are appreciable—that is, roughly speaking, those at which the magnetoresistance effect is appreciable.)

3. THERMOELECTRIC POWER

It was concluded in Sec. 2 that the four unknown phenomenological transport coefficients, the γ_s and the λ_s^H, cannot be determined completely from experimental values of the Nernst coefficient, but that there is the possibility of obtaining the required additional information from measurements of the thermoelectric power, which should involve the γ_s. To investigate this possibility, we derive an expression for the absolute thermoelectric power, Q, of a semiconductor by means of the general formula[11]

$$Q=d(\zeta/e)/dT-\Omega(T). \tag{27}$$

Here Ω is the transport coefficient appearing in Eq. (1) and ζ is the chemical potential of the electrons (the Fermi level, on the band model) in the substance when it is electrically neutral and in equilibrium at temperature T. For a two band semiconductor we have, subject to assumptions (d), (f_1, (g) of I,

$$\begin{aligned} n_1&=2(2\pi m_1kT/h^2)^{\frac{3}{2}}\exp[(\zeta-\epsilon_{01})/kT],\\ n_2&=2(2\pi m_2kT/h^2)^{\frac{3}{2}}\exp[(\epsilon_{02}-\zeta)/kT], \end{aligned} \tag{28}$$

where the ϵ_{0s} are the band-edge energies and the m_s are the density-of-states masses. From the product and the quotient, respectively, of the two Eqs. (28), we obtain Eq. (13) and the formula

$$2\zeta=\epsilon_{01}+\epsilon_{02}+kT[\log(n_1/n_2)-\log(m_1/m_2)^{\frac{3}{2}}]. \tag{29}$$

By (10) and (7), the second term of (27) is given by

$$\begin{aligned}\Omega=(k/e)\{&(\sigma_1/\sigma)[Td(\log n_1)/dT+\gamma_1]\\ &-(\sigma_2/\sigma)[Td(\log n_2)/dT+\gamma_2]\}.\end{aligned} \tag{30}$$

On substituting (29) and (30) into (27), and making use of (13), we obtain

$$\begin{aligned}Q=\frac{k}{e}\bigg[&\frac{1}{2}\log\left(\frac{n_1}{n_2}\right)-\frac{3}{4}\log\left(\frac{m_1}{m_2}\right)\\ &-\left(\frac{\sigma_1-\sigma_2}{2\sigma}\right)\left(\frac{\Delta}{kT}+3\right)-\left(\frac{\sigma_1\gamma_1-\sigma_2\gamma_2}{\sigma}\right)\bigg].\end{aligned} \tag{31}$$

An alternative form of this result is

$$\begin{aligned}Q=-\frac{k}{e}\bigg[&\frac{\sigma_1}{\sigma}\left(\frac{\epsilon_{01}-\zeta}{kT}+\frac{3}{2}+\gamma_1\right)\\ &-\frac{\sigma_2}{\sigma}\left(\frac{\zeta-\epsilon_{02}}{kT}+\frac{3}{2}+\gamma_2\right)\bigg].\end{aligned} \tag{32}$$

This last expression for Q reduces for the classical model to that given by Johnson and Lark-Horovitz [Eq. (24) of reference 3]. (One may see this by setting $\gamma_1=\gamma_2=\frac{1}{2}$ in (32), and noting that n_1 and n_2 in their expression are given by (28) here and that c in their notation stands for μ_1/μ_2.)

We are concerned here with the possibility of determining the γ_s empirically. We consider first the conditions where the Dunlap ellipse, and the Nernst-coefficient parabola of Sec. 2, can be obtained—that is, the temperature range where by comparing crystals with different dopings at fixed temperature an appreciable range of conductivities, centered on the minimum value for the temperature, can be covered. The result (31) may be written

$$\begin{aligned}Q=\frac{k}{e}\bigg\{&\left[\frac{1}{2}\log\left(\frac{\sigma_1}{\sigma_2}\right)-\left(\frac{\sigma_1-\sigma_2}{2\sigma}\right)\left(\frac{\Delta}{kT}+3+\gamma_1+\gamma_2\right)\right]\\ &-\left[\frac{1}{2}\log\left(\frac{\mu_1}{\mu_2}\right)+\frac{3}{4}\log\left(\frac{m_1}{m_2}\right)+\frac{(\gamma_1-\gamma_2)}{2}\right]\bigg\}.\end{aligned} \tag{31'}$$

From this formula we see at once that plotting a doping curve for Q would be useless for our purpose, since it could determine only the combination $(\Delta/kT+3+\gamma_1+\gamma_2)$ which is already given by the doping curve for B. To obtain $\gamma_1-\gamma_2$ (and hence, $\gamma_1+\gamma_2$ being known,[10] the γ_s separately) we would have to be able to determine the *absolute* value of Q,[11] as defined by (27). It should be possible to locate the point on the doping curve where the conductivity σ is a minimum: then the absolute value of Q, times $-e/k$, is equal to the second square bracket of (31′). The resulting determination of $\gamma_1-\gamma_2$ is, of course, still only as accurate as the knowledge of the other two terms in the bracket allows. In the extrinsic range we have, by (32),

$$Q=(k/e)[(\zeta-\epsilon_{0s})/kT\mp(\gamma_s+\tfrac{3}{2})] \tag{33}$$

(where $s=1$ or 2, according to which band is conducting). Thus the appropriate γ might be determined

---

[11] Equations (27) and (31) will be derived and discussed in a forthcoming paper (III of the series) on thermoelectricity. It seems worth anticipating this paper with the remark that, although of course only the difference between values of Q for a pair of substances can be measured, the expression (27) gives the "absolute" thermoelectric power of a single substance in the sense that it makes TdQ/dT equal to the Thomson coefficient, which *is* defined for a single substance. In I it was implied erroneously that Ω is the Thomson coefficient: the actual general relation of the latter to Ω will be derived in the forthcoming paper.

if the absolute value of Q could be obtained, and if the distance of the Fermi level from the corresponding band edge were estimated with an error small enough compared with kT. Since Eqs. (31), (32) refer to the *absolute* value of Q, as defined by (27), one can hope at best to determine experimentally the difference between the values of the γ_s in the semiconductor studied and those in reference semiconductors for which the classical model or some other simple model is known to be valid, so that for the latter the γ_s can be assigned theoretical values. (Alternatively, one might hope to use as a reference substance in the thermocouple a metal so well understood theoretically that absolute values of Q could be calculated for it.) In spite of these disappointing conclusions, we shall see in the following section how, in the extrinsic case where (33) applies, the thermoelectric power can provide information of the kind sought.

A method of determining $\gamma_1+\gamma_2$, independent of that of Sec. 2 and possibly more accurate, seems worth noting: for a p-n junction, provided the values of the γ_s are independent of conduction type and of doping,[12] the thermoelectric power of the couple is, by (33),

$$Q_p-Q_n=\frac{k}{e}\left[\frac{(\epsilon_{01}-\zeta_n)+(\zeta_p-\epsilon_{02})}{kT}+3+\gamma_1+\gamma_2\right]. \quad (34)$$

The left-hand side of (34) is a purely experimental quantity (being measured either directly or by comparison of the two semiconductor-reference metal couples). The first term in the square bracket on the right can be estimated if the impurity concentrations and activation energies, and the density-of-states masses of the bands, are known. This method of determining $\gamma_1+\gamma_2$ should be applicable at temperatures between the transition range and the range where the phonon drag effect appears.

4. PHONON DRAG EFFECT

The result (25) indicates an effect which seems not to have been noticed so far. The Nernst coefficient in the extrinsic range is proportional to the appropriate γ (i.e., γ_1 for n-type, γ_2 for p-type), and γ measures the drift velocity due to the Soret effect: the drift caused by a temperature gradient in itself, apart from the diffusion caused by any resulting gradient of carrier concentration. If, as is normally so, the relaxation processes determining the mobility are interactions of the carriers with a substrate (lattice plus impurities and defects) which itself *does not deviate appreciably from local thermal equilibrium*, then the Soret effect arises entirely from the fact that carriers entering a given element of volume from different directions represent Maxwell distributions for different temperatures and hence have different mean speeds. In these circumstances we may expect values of γ of order unity. [See, for example, Eq. (17).] If, however, the substrate equilibrium is appreciably disturbed by the temperature gradient, in that the anisotropy of the distribution of lattice modes with which the electrons interact[13] is appreciable, then the scattering kernel $S(\mathbf{p},\mathbf{p}')$ of the carrier relaxation processes will be anisotropic, in the sense that $S(-\mathbf{p},-\mathbf{p}')\neq S(\mathbf{p},\mathbf{p}')$, and hence there will be a net rate of transfer of momentum (in the direction of $-\mathrm{grad}T$) from the substrate to a Maxwell distribution of carriers. This transfer of momentum will by itself cause a drift of carriers down the temperature gradient, in addition to the drift of the normal origin mentioned above, and hence it will contribute an extra term to γ. The carriers are entrained by the flux of lattice modes. Anomalies observed in the low-temperature thermoelectric power of germanium,[14] and more recently in those of silicon[15] and of indium antimonide,[16] have been interpreted[4,15,16] as being due to this effect, termed the "phonon drag effect." The anomaly occurs at low temperatures because the contribution to γ from the effect is essentially proportional to the flux of thermal energy transported by the lattice vibration modes, and the "lattice-mode thermal conductivity" of the single crystal increases greatly with decreasing temperature in the range in question. Since B is proportional to γ, there should be an anomaly in the former corresponding to the anomaly in Q.[17] From (25) and (33), we have

$$B=\left[\frac{k}{e}\left(\frac{3}{2}\pm\frac{\epsilon_0-\zeta}{kT}\right)\pm Q\right]\left(\frac{\mu^H-\lambda^H}{c}\right) \quad (35)$$

(where ϵ_0, μ^H, and λ^H refer to the appropriate band). Since the contribution to γ due to the phonon drag effect is evidently large compared with the value, of order unity, to be expected normally, it seems that the anomaly in Q and in γ can be adequately estimated in practice, from measurements on a thermocouple of the semiconductor and a reference metal, without one's being able to infer the absolute value of Q from the measurements.

For p-type germanium, enough information is available for an attempt to estimate the anomaly in B. Let the superscript ph denote the anomalous contribution. From Geballe and Hull's data, as analyzed by Herring (reference 4, Fig. 2), it appears that for a high-resistance sample $\gamma_2^{ph}\sim 50$ at 50°K (and increases to about double this value at 25°K). Specifically, for Geballe and Hull's "Sample 7"[14] Herring finds $Q^{ph}\simeq 3.0$ millivolts/degree.

---

[12] This condition is not satisfied at temperatures where the phonon drag effect is appreciable or where the effect of impurity scattering on the mobilities is appreciable.

[13] Scattering of a carrier at an impurity atom, properly described, still involves absorption or emission of lattice-vibration quanta to take up the recoil momentum.

[14] H. P. R. Frederikse, reference 4; T. H. Geballe and G. W. Hull, Phys. Rev. **94**, 1134 (1954).

[15] T. H. Geballe and G. W. Hull, Phys. Rev. **98**, 940 (1955).

[16] H. P. R. Frederikse and E. V. Mielczarek, Phys. Rev. **99**, 1889 (1955).

[17] P. J. Price, Bull. Am. Phys. Soc. Ser. II, **1**, 47 (1956).

From Morin's published data[18] we take $\mu_2^H \simeq 6.3\times 10^4$ cm²/volt sec at 50°K. Corresponding to (35), we have

$$B^{ph} = \pm Q^{ph}(\mu^H - \lambda^{H\,ph})/c, \tag{36}$$

where $\lambda^{H\,ph}$ is the galvanomagnetic mobility appropriate to the anomalous contribution to γ. Since no value of $\lambda_2^{H\,ph}$ is available, some assumption about it or estimate of it is needed. One possibility is to suppose it equal to λ_2^H for the normal Soret effect, and to set the latter, in turn, equal to zero, the value it has on the classical model. In this way we arrive at the estimate, for the p-type "Sample 7" at 50°K, given in reference 17:

$$B^{ph} \simeq -2\times 10^{-6} \text{ volt/gauss degree.} \tag{37}$$

On further consideration, this assumption about $\lambda_2^{H\,ph}$, and the resulting estimate of B^{ph}, seem dubious. It is implausible that λ^H should be the same for the normal and anomalous contributions to γ.[19] Even for the classical model, however, the calculation of $\lambda^{H\,ph}$ in terms of the perturbation of the phonon distribution proves to be complicated and delicate; and at the time of writing no final conclusions have been reached. (The source of the difficulty is that the effect vanishes in the zeroth order of the ratio s/v, where s is the velocity of sound, and derives from the first order of this ratio: that is to say, it originates in what are normally neglected corrections in the classical model, from the difference between the scattering by absorption of phonons and the scattering by emission of phonons.) Since in any case a detailed discussion of this problem would be inappropriate to the plan of this paper, the subject will not be followed any further here. The result (37) must be replaced by the weaker statement

$$|B^{ph}| \sim 10^{-6} \text{ volt/gauss degree} \tag{38}$$

(for Geballe's "Sample 7" at 50°K).

The useful voltages actually available are limited by the restriction, mentioned at the end of Sec. 2, of the values of H to those for which magnetoresistance and other effects quadratic in H are negligible. A rough criterion for the allowable fields is that $(\mu^H H)^2 \ll c^2$. That is, say,

$$\mu^H H/c \leqslant t \gtrsim 1, \tag{39}$$

(where t might be a few tenths). We then have, roughly,

$$|B^{ph}| H < \sim |Q^{ph}| t. \tag{40}$$

This criterion would limit the voltages available from the anomalous Nernst effect to those from the anomalous Seebeck effect times t times the appropriate geometrical factor. In our case, this is of order a millivolt/degree times the geometrical factor. However, actual measurements of the Hall effect as a function of field H show that for p-type Ge the criterion (39) is not stringent enough: the Hall mobility falls off at much lower fields.[20] At 50°K the field is limited to less than 100 gauss, for linearity to be preserved, and this limits $|B^{ph}|H$ to of order 100 microvolts/degree.

ACKNOWLEDGMENTS

I am indebted to Dr. G. R. Gunther-Mohr and Mr. P. B. Linhart for helpful comments on the manuscript, and to Mr. Linhart for other assistance in the preparation of this paper.

APPENDIX: VALUES OF THE TRANSPORT PARAMETERS FOR A GENERAL $\tau(v)$

We need only consider a single band, so we drop the subscript s. The analysis corresponds to that in I from Eq. (16) to Eq. (19) there. Here we adopt the same notation, except that $w(\mathbf{p}) \equiv \pm(\epsilon - \epsilon_0)$ stands for the excitation energy of the particle relative to the band edge. By retaining the terms in ∇T and substituting

$$\mathfrak{D} f_1 = f_1/\tau, \tag{A1}$$

we obtain, in place of Eq. (19) of I,

$$(f_1/\tau) + q(\mathbf{v}\times\mathbf{h})\cdot(\partial f_1/\partial\mathbf{p}) = f_0\mathbf{v}\cdot[\mathbf{X} - w\mathbf{Y}], \tag{A2}$$

where again $\mathbf{h} \equiv \mathbf{H}/c$, and

$$\begin{aligned} \mathbf{X} &\equiv (q/kT)\mathbf{E} - \nabla \log n + (3/2T)\nabla T, \qquad \text{(A3)} \\ \mathbf{Y} &\equiv (1/kT^2)\nabla T. \end{aligned}$$

If we denote the right-hand side of (A2) by ϕ, then the solution of (A2) for f_1, as far as terms linear in $\mathbf{h}$, is

$$f_1 = \tau\phi - \tau q(\mathbf{v}\times\mathbf{h})\cdot[\partial(\tau\phi)/\partial\mathbf{p}]. \tag{A4}$$

In evaluating the second term of (A4), we make use of the facts that $\partial w/\partial\mathbf{p} = \mathbf{v}$, that $\tau = \tau(v)$, and that $w = mv^2/2$ (where m is the band-edge effective mass). Then we find

$$f_1 = \tau f_0\{(\mathbf{X} - w\mathbf{Y})\cdot\mathbf{v} - (q\tau/m)[\mathbf{h}\times(\mathbf{X} - w\mathbf{Y})]\cdot\mathbf{v}\}. \tag{A5}$$

The drift velocity, $\mathbf{u} = \int f_1 \mathbf{v} d^3\mathbf{p}$, according to (A5) is

$$\begin{aligned} \mathbf{u} = {} & \mathbf{X}\cdot\langle\tau\mathbf{vv}\rangle - \mathbf{Y}\cdot\langle w\tau\mathbf{vv}\rangle \\ & - (q/m)[(\mathbf{h}\times\mathbf{X})\cdot\langle\tau^2\mathbf{vv}\rangle - (\mathbf{h}\times\mathbf{Y})\cdot\langle w\tau^2\mathbf{vv}\rangle], \end{aligned} \tag{A6}$$

where the brackets $\langle\ \rangle$ denote averages over the equilibrium distribution. Since we are assuming spherical energy surfaces, the four scalar products $\mathbf{X}\cdot\langle\tau\mathbf{vv}\rangle$, etc., may be replaced by $\mathbf{X}\langle\tau v^2\rangle/3$, etc. Then, in the notation (18),

$$\begin{aligned} 3\mathbf{u} = {} & (2,1)\mathbf{X} - \tfrac{1}{2}m(4,1)\mathbf{Y} \\ & - (q/m)\mathbf{h}\times[(2,2)\mathbf{X} - \tfrac{1}{2}m(4,2)\mathbf{Y}]. \end{aligned} \tag{A7}$$

On the other hand, Eq. (6) becomes, when written in terms of $\mathbf{X}$ and $\mathbf{Y}$,

$$\begin{aligned} \mathbf{u}_s = {} & \mu_s(kT/e)[\mathbf{X} - (\gamma_s + \tfrac{3}{2})kT\mathbf{Y} \pm \mu_s^H \mathbf{h}\times\mathbf{X} \\ & \mp (\gamma_s\lambda_s^H + \tfrac{3}{2}\mu_s^H)kT\mathbf{h}\times\mathbf{Y}]. \end{aligned} \tag{A8}$$

---

[18] F. J. Morin, Phys. Rev. 93, 62 (1954).

[19] In any case the results of the classical model are not even approximately correct for p-type Ge, as can be seen from the fact[18] that μ_2^H/μ_2 varies strongly with temperature, even in the absence of impurity scattering.

[20] Harman, Willardson, and Beer, Phys. Rev. 94, 1065 (1954).

Comparing coefficients of (A7) and (A8), remembering $q = \mp e$, and substituting

$$3kT = m(2,0),$$

we arrive at the results (14)–(17).

Notes added in proof.—(a) The effect predicted in Sec. 4 has been demonstrated by Geballe[21] in his measurements of the low-temperature Nernst coefficient of germanium. The experimental work is now concentrated on further developments, such as the interesting and important question of the effect of *strong* magnetic fields. Dr. Herring is conducting a detailed analysis of the Bell Laboratories results.

(b) The difficulty, mentioned in the passage between Eqs. (37) and (38), over the calculation of the Hall angle for the phonon drag effect has been clarified, in a discussion with Dr. Herring and on further study, as follows: The various phonon drag effects are proportional, for given relaxation rates, to the square of the velocity of the lattice modes involved. The difficulty explained in the paper appears when one tries to calculate the Seebeck and Nernst effects, because this velocity factor then introduces itself through the electron energy change[22] in electron-lattice collisions, which cannot be neglected but rather is the origin of the effects. If, however, one calculates instead the Peltier and Ettinghausen effects (which are equivalent to the two former effects, by the Kelvin and Bridgman formulas respectively), thus considering the energy transported with an electric current rather than the current induced by a temperature gradient, then this velocity-of-sound factor appears in the effect of collisions with electrons on the lattice mode distribution, in a way that does not cause mathematical difficulty. I was trying to work by the former method, while Herring used the latter one. By the latter method, the Hall angle corresponding to the "phonon drag" contribution to the Ettinghausen effect may be readily calculated for the classical model. If the resulting formula for $\lambda^{H\ ph}$ on the classical model is combined with Geballe's results, it appears that the low-energy phonon relaxation time is proportional approximately to the square of the phonon wavelength. It seems not worthwhile to give this calculation and formula, since the real interest is now in the actual departure from the classical model, and since my result is contained in the far more advanced work of Herring on the latter question. It should be mentioned, however, that the formulas for the anomalous contributions to the Seebeck and Nernst effects which I obtained "the hard way" for the classical model do not agree with those for the Peltier and Ettinghausen effects obtained "the easy way." It has been pointed out by Sondheimer[23] that the deviations of the phonon and electron distribution functions from equilibrium are actually given by two *coupled* equations such that it is incorrect to solve for one deviation by assuming the other to be as if the phonon drag effect were absent (which is what has been done so far, except in the calculation by Herring mentioned below). This criticism is quite correct, and bears equally on both kinds of calculation described above: on further examination, however, it turns out that the present disagreement between them is not thereby accounted for. The order of magnitude of the necessary correction to the results of calculations of the phonon drag effects is the proportion of relaxation events, for the low-energy phonons, which are collisions *with an electron*. This is normally negligible, except for high electron concentrations. The existence of this correction was pointed out by Herring in his paper (reference 4, the beginning of Sec. VII), where he makes an estimate of the correction, as a function of electron concentration, for the thermoelectric power of germanium. I am indebted to Geballe and Herring for communicating and discussing their work.

(c) Reference 8 might usefully be amplified. What happens in the circumstance referred to is that the component of the Soret drift parallel to the temperature gradient vanishes, while the component normal to the temperature gradient does not. This is an interesting situation, though not evidently of any deep significance.

(d) A theory in some respects comparable to, and in some respects supplementing, the present one has been published by Bass and Tzidilkovski.[24]

(e) The results given in Secs. 2 and 3, and Eq. (35), depend on the assumption that the band energies are independent of temperature. Generalizations to include a temperature variation are given in a forthcoming paper (III of this series) on thermoelectricity. See the paragraph containing Eqs. (47)–(49), Sec. 2, there.

---

[21] C. Herring and T. H. Geballe, Bull. Am. Phys. Soc. Ser. II. **1**, 117 (1956). This discovery, which was made independently of my work, was briefly communicated to me by Dr. Herring while the manuscript of the present paper was in the final stages of preparation. I thought it best to complete the latter without reference to the experimental work, adding this note instead of modifying the text.

[22] It is convenient here to use the word "electron" to mean either conduction electron or hole.

[23] E. H. Sondheimer, Proc. Roy. Soc. (London) **A234**, 391 (1956).

[24] F. G. Bass and I. M. Tzidilkovski, J. Expertl. Theoret. Phys. U.S.S.R. **28**, 312 (1955) [Soviet Phys.—JETP **1**, 267 (1955)].