

NRL Report 5888

UNCLASSIFIED

A Brief Survey of Elementary Thermoelectric Theory

WILLIAM H. LUCKE

*Energy Conversion Branch
Electronics Division*

May 3, 1963



U.S. NAVAL RESEARCH LABORATORY
Washington, D.C.

CONTENTS

Abstract	1
Problem Status	1
Authorization	1
INTRODUCTION	1
THERMOELECTRIC DEFINITIONS	1
THE KELVIN RELATIONS	2
Classical Thermodynamics	2
Irreversible Thermodynamics	5
TRANSPORT THEORY	10
Fundamental Concepts	10
Transport Integrals	13
FIGURE OF MERIT	13
CONCLUSION	15
REFERENCES	15

Copies available from Office of Technical Services
Department of Commerce - \$.50

A Brief Survey of Elementary Thermoelectric Theory

WILLIAM H. LUCKE

*Energy Conversion Branch
Electronics Division*

The three fundamental effects of thermoelectricity—Seebeck, Peltier, and Thomson—are presented and the corresponding coefficients are defined. Following a simplified derivation of the Kelvin relations, a more sophisticated and detailed presentation of the irreversible thermodynamic approach to these relations is made. It is demonstrated that these relations are valid in spite of the arbitrary assumptions as to the separation of thermal and electrical conduction made in the former derivation.

Next, a brief summary is given of a derivation of the expression for the Seebeck coefficient based on a quasi-free electron model and simple energy dependence of an isotropic relation time.

In closing, the attempt to apply band theory to the search for better thermoelectric materials is very briefly described and some interesting ideas for improvement are mentioned along with criticisms concerning predictions of the maximum possible figure of merit.

INTRODUCTION

It has long been recognized that the concept of the electrons in a conductor as a gas has a certain validity. In terms of this model, the electrons may be thought of as a gas of charged particles which undergoes mass transport in the presence of an electric field, necessarily transporting charge and kinetic energy; this last is due to the fact that the electrons possess thermal energy in the same sense as an ordinary gas. Obviously, this gives us a means of transporting heat by the application of an electric field and in fact, we can by these means, transport heat *against* a thermal gradient! The converse is also true: Because electrons carry charge *and* can be caused to diffuse by the presence of a thermal gradient, they can be made to transport charge *against* an electric field. Thus electrons afford a very direct means for the conversion of electrical to thermal energy and vice versa.

While this naive model gives the broad outline of thermoelectricity correctly, it fails completely where any attempt is made to extend it to a quantitative description of charge or heat transport. If, for example, heat is to be transported, the electrons must be capable of exchanging heat with both the source and the sink; this obviously demands that the electron gas must interact with its container, which is of course the lattice, and the details of this interaction form, as is well-known, one of the great problems of solid-state theory.

However, before we undertake a rudimentary discussion of this formidable subject, it would be well to review the classical history of thermoelectricity and to see what thermodynamics has to tell us.

THERMOELECTRIC DEFINITIONS

The particulars of how Seebeck described the effect named for him or how Peltier discovered his effect a few years later will be omitted here. More will be said about the Thomson effect because of its prominence in the derivation of the very important Kelvin relations. To define these terms:

The Seebeck effect refers to the appearance of a voltage in an electric circuit composed of two dissimilar conductors with the two junctions held at different temperatures (Fig. 1). For a given couple, the voltage depends on the difference in temperature between the junctions. The effect is described quantitatively by the Seebeck coefficient α (formerly the thermoelectric power), which is defined as

$$\alpha = \lim_{\Delta T \rightarrow 0} \frac{\Delta V}{\Delta T}$$

where V is the emf between the two junctions. The Seebeck effect is reversible in that if the temperatures of the junctions are reversed, the polarity of the voltage reverses, but the magnitude is unchanged.

The Peltier effect occurs when a current flows through the junction between two dissimilar conductors. Heat is either rejected or absorbed depending on the direction of the current. Suppose that a conventional current I is flowing from material 2 to material 1, the junction being maintained at a fixed temperature T_h , and that heat is absorbed (i.e., the junction is cooled) at a certain rate dQ/dt . The flow of heat is found to be proportional to the current I , the constant of proportionality being the Peltier coefficient. Thus,

$$\frac{dQ}{dt} = \Pi I.$$

NRI Problem 52E02-02; Project SR 007-12-01.

This is an interim report; work on this problem is continuing.

NOTE: This report is an expanded version of a paper presented at the 7th Annual Seminar of the Chicago-Western Chapter of the American Society for Metals, Chicago, May 10, 1962.

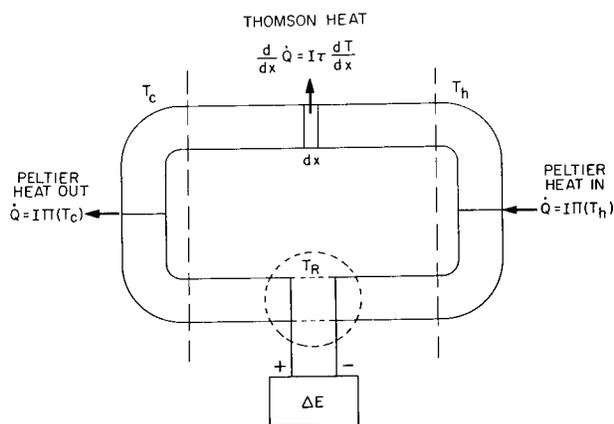


Fig. 1 — Schematic thermocouple circuit showing absorption of Peltier heat at the hot junction, rejection of Peltier heat at the cold junction, and Thomson heat being rejected along one leg and a potentiometer placed to measure the Seebeck voltage. The dashed lines indicate the boundaries of isothermal regions.

If the direction of current flow is reversed but its magnitude is unchanged, heat will now be evolved at the junction at precisely the same rate as it was previously absorbed. Thus, the Peltier effect is also reversible.

The Thomson effect was predicted by Lord Kelvin (William Thomson) on thermodynamic grounds before he demonstrated it experimentally in 1853. Unlike the other two effects, it may be observed where one conductor only is present and refers to the fact that a current flowing up a temperature gradient will absorb (or evolve) heat. Given an elemental length dx of a conductor in a thermal gradient, dT/dx , the time rate of absorption or evolution of heat is (at a constant temperature) proportional to the product of the current I and the gradient. The constant of proportionality will be written as τ and is called the Thomson coefficient.

$$\frac{d}{dx} \left(\frac{dQ}{dt} \right) = \tau I \frac{dT}{dx}$$

If either the direction of the current or the direction of the gradient is reversed, heat which was formerly absorbed will be rejected and vice versa. If both the gradient and current are reversed, the heat continues to be absorbed as before. To calculate the total Thomson contribution to the electrical energy, we integrate over the length of the conductor from, say, $x = 0$, at which end the temperature is T_c , to $x = l$, at which end the temperature is T_h . Thus,

$$\frac{dQ}{dt} = \int_0^l \tau I \frac{dT}{dx} dx = \int_{T_c}^{T_h} \tau I dT$$

Hence, the Thomson contribution depends only on the junction temperatures and is independent of the details of the temperature gradient. This is known as the Law of Magnus and explains the usefulness of thermocouples for temperature measurement.

Lord Kelvin's prediction of the effect was based on a sound application of the second law of thermodynamics to thermocouples: If Peltier heat is reversibly absorbed at the hot junction, and reversibly evolved at the cold junction, then it follows that the heat energy absorbed at the hot junction furnishes the electrical energy both to drive current through the couple and to actuate the heat rejecting mechanism of the cold junction. In other words, one can speak of a forward emf at the hot junction which should be balanced by the IR drop in the legs of the couple and an opposing emf at the cold junction. A potentiometric measurement of the Seebeck voltage would measure the difference between the hot junction emf and the cold junction emf. However, Lord Kelvin observed that for certain couples, the Seebeck voltage dropped to zero when a particular hot junction temperature was reached, and further, that if he increased the hot junction temperature beyond this value, the voltage and current reversed. This would appear to indicate that heat was being absorbed at the cold junction, causing current flow and the rejection of heat at the hot junction, clearly a violation of the second law of thermodynamics! Therefore, he postulated that there must be another mechanism present which could absorb heat and generate electrical energy, and as a result of some very sound and original thinking was led to predict the effect now named for him.

THE KELVIN RELATIONS

Classical Thermodynamics

Continuing his application of the second law of thermodynamics, Lord Kelvin was led to consider a quantitative calculation of the change of entropy for thermocouple operation. Since all three thermoelectric effects are reversible, a couple should, at first sight, offer an ideal application of the second law (i.e., zero net change of entropy); however, it is an experimental fact that the irreversible phenomena of heat and electrical conduction are always present. Kelvin recognized this difficulty but postulated that one could separate the irreversible phenomena from the reversible ones and consider the latter separately. Proceeding on this basis he was able to derive relations between the three effects:

On a per unit time basis we apply the first law:

$$V_{12}I = I\Pi_{12}(T_h) - I\Pi_{12}(T_c) + I \int_{T_c}^{T_h} \tau_1 dT - I \int_{T_c}^{T_h} \tau_2 dT$$

or

$$V_{12} = \Pi_{12}(T_h) - \Pi_{12}(T_c) + \int_{T_c}^{T_h} (\tau_1 - \tau_2) dT$$

For incremental differences between cold and hot junctions we have

$$\Delta V_{12} = \Pi_{12}(T + \Delta T) - \Pi_{12}(T) + \int_T^{T+\Delta T} (\tau_1 - \tau_2) dT.$$

By usual first order expansion this gives

$$\Delta V_{12} = \Pi_{12}(T) + \frac{d\Pi_{12}}{dT} \Delta T - \Pi_{12}(T) + (\tau_1 - \tau_2) \Delta T$$

or

$$\frac{\Delta V_{12}}{\Delta T} = \frac{d\Pi_{12}}{dT} + (\tau_1 - \tau_2).$$

In the limit

$$\left(\text{since } \alpha_{12} = \lim_{\Delta T \rightarrow 0} \frac{\Delta V_{12}}{\Delta T} \right)$$

$$\alpha_{12} = \frac{d\Pi_{12}}{dT} + (\tau_1 - \tau_2)$$

or

$$\frac{d\Pi_{12}}{dT} = \alpha_{12} - (\tau_1 - \tau_2). \quad (1)$$

Now apply the second law for an incremental temperature difference and a reversible system:

$$\begin{aligned} \frac{\Delta Q}{T} = \Delta S = 0 &= \frac{I\Pi_{12}(T + \Delta T)}{T + \Delta T} - \frac{I\Pi_{12}(T)}{T} \\ &+ \int_T^{T+\Delta T} \frac{\tau_1 - \tau_2}{T} dT. \end{aligned}$$

The current can be divided out, and we note that the first two terms are the incremental calculus equivalent of

$$\frac{d}{dT} \left(\frac{\Pi_{12}}{T} \right) \Delta T.$$

Therefore

$$0 = \frac{d}{dT} \left(\frac{\Pi_{12}}{T} \right) \Delta T + \frac{\tau_1 - \tau_2}{T} \Delta T$$

from which we get

$$0 = \frac{1}{T} \frac{d\Pi_{12}}{dT} - \frac{\Pi_{12}}{T^2} + \frac{\tau_1 - \tau_2}{T}$$

or, multiplying through by T and rearranging,

$$\frac{d\Pi_{12}}{dT} = \frac{\Pi_{12}}{T} - (\tau_1 - \tau_2). \quad (2)$$

Comparing Eqs. (1) and (2) we see that

$$\alpha_{12} = \frac{\Pi_{12}}{T}. \quad (\text{Second Kelvin Relation})$$

Appropriate substitution of this result, the second Kelvin relation, in Eqs. (1) or (2) yields the first Kelvin relation, i.e.,

$$\tau_1 - \tau_2 = -T \frac{d\alpha_{12}}{dT}. \quad (\text{First Kelvin Relation})$$

For some reason there seems to be a popular feeling that the greatest value of the Kelvin relations is the establishment of a relation between the Seebeck coefficient α and the Thomson coefficient τ . Actually, when one accepts the concept of an absolute Seebeck coefficient, the form of the relationship between α and τ can be derived by considering the electrons as a working substance that goes through a cycle and returns to its original thermodynamic condition. From the conservation of energy standpoint, it should not be surprising to find Π , α , and τ related; it is significant that their relationships involve the absolute temperature.

The first Kelvin relation leads rather naturally to the concept of absolute Seebeck coefficient. We have

$$\tau_1 - \tau_2 = -T \frac{d\alpha_{12}}{dT}.$$

As we have seen the τ is characteristic of one material only, whereas the α involves the difference of two materials. We come naturally then to write

$$\tau_1 - \tau_2 = -T \frac{d}{dT} (\alpha_1 - \alpha_2)$$

or in general

$$\tau = -T \frac{d\alpha}{dT}.$$

Thus the integration of τ/T of a substance from close to absolute zero up to more common temperatures would give reasonably accurate values for its absolute α . Alternatively, if one member of a couple is in the superconducting state, the measured Seebeck voltage is that of the nonsuperconducting member. Since in the superconducting state there is no transfer of energy between current carriers and lattice, one expects no thermoelectric effects, and, indeed, none have been found. Combinations of these approaches have been applied to the problem and values of absolute α can be found in the literature.

The concept of an absolute α (which we shall see later is associated with energy of transport) enables us to construct a diagram showing the value of α at every point in a thermocouple and making clear that its operation is a cyclic one with the charge carriers in the role of the working substance. Suppose that the absolute α 's of two homogenous, isotropic materials have been determined and lie along the arcs FC and HM (Fig. 2). A couple is now made up of these materials and operated with the junctions held at the absolute temperatures T_c

we account for all the heat absorbed and rejected, and find that we have rejected more than we have absorbed by the amount MHFCM. Evidently this is the work that must be done to absorb heat at a cold sink and reject it to a hot one! This is aside from the Joule losses. Again a reversible coefficient of performance could be written and again this will not give a Carnot expression unless the Thomson coefficient is zero.

At this point some practical comments might be made. The diagram is directly applicable to the graphical analysis of the reversible aspects of a real couple. In practice such a couple would be constructed with the upper leg a p-type semiconductor and the lower leg an n-type one. In this case the hot junction heat absorbed would correspond to the energy necessary for the formation of hole-electron pairs plus the energy for them to diffuse away. The cold junction heat evolved would correspond to the energy given up on recombination plus the energy of transport left after traversing the gradients. The Thomson heats, as before, correspond to changes in the energy of transport, which largely come from the shift of Fermi level with temperature. We have, of course, made the usual simplification that transport of holes is the same as transport of positive charge. There is one change which we have to consider and that is the Thomson heat of the electrons. Originally we spoke of Thomson heat evolved upon moving the unit of positive charge up the gradient in the lower leg, as due to moving to lower values of α . Actually, of course, the electrons move down the gradient from M to H; from the algebraic point of view we would say that this is moving toward increasing α ; however, from the electron's point of view it is moving toward a decreasing value of α . Recall that, in an n-type semiconductor, the Fermi level is below the conduction band edge and so is negative.

Irreversible Thermodynamics

In the preceding section the Kelvin relations were derived on the assumption that the irreversible processes of Joule heating and thermal conduction could be considered separately from the reversible phenomena of thermoelectricity. However, the fact that Thomson and Peltier heat absorption and rejection depend on the flow of current is inescapable; and current flow is— with the exception of superconductivity, where thermoelectric phenomena are absent—always accompanied by irreversible Joule heating. Correspondingly, the Wiedemann-Franz relation shows that at least part of the thermally conducted heat comes from the transport of charge carriers. Hence, both reversible and irreversible effects are intimately connected with the mechanics of carrier transport and so with each other. From this point of view, the correctness of the Kelvin relations would appear to be fortuitous.

In recent years, however, the development of that branch of thermodynamics called irreversible thermodynamics has put the Kelvin relations on a much firmer footing. In this approach the irreversible effects, which arise because of the departure from equilibrium, are calculated to a first order (corresponding to a linear perturbation of the equilibrium state) and are included in the equations. Clearly the results will be valid only for small departures from equilibrium; however, the success of the method shows that it is quite valid for thermocouple analysis under conditions ordinarily encountered.

In general, in a nonstatic situation there will be thermodynamic fluxes, or flows, which originate in thermodynamic forces. These forces are expressible as gradients of thermodynamic variables. For example, a thermal gradient (force) gives rise to a heat flow; similarly a potential gradient gives rise to charge flow, and a concentration gradient to a particle flow. In each of these cases the relation between the flux (or current density) is a linear one:

$$\begin{aligned} J_Q &= -K\nabla T \\ J_e &= -\sigma\nabla\phi \\ J_p &= -D\nabla n \end{aligned}$$

where

$$\begin{aligned} J_Q &= \text{heat flux} \\ J_e &= \text{charge flux} \\ J_p &= \text{particle flux} \\ K &= \text{thermal conductivity} \\ \sigma &= \text{electrical conductivity} \\ D &= \text{diffusion coefficient} \\ T &= \text{temperature} \\ \phi &= \text{electrical potential} \\ n &= \text{particle density.} \end{aligned}$$

Now the experimental facts of thermoelectricity show that there is a crosscoupling between forces and flows. A thermal gradient gives rise to an electric flux, and a potential gradient to a heat flux. If we make the assumption (a) of superposition of effects, and (b) that the crosscoupling terms are linear in the forces, we can write

$$J_e = M_{11}\nabla\phi + M_{12}\nabla T$$

$$J_Q = M_{21}\nabla\phi + M_{22}\nabla T$$

As to the assumption of superposition, there is little to say except that it seems justified in a first order description. The assumption of linear crosscoupling terms seems only natural in view of the linear relation of the primary forces and fluxes; however, it, too, is justifiable in terms of first-order perturbation.

These equations are given to illustrate the form of relations to be used. In addition to the two forces and fluxes we have four coefficients which must be determined through experimentally determinable relations between forces and fluxes. Without going into detail it turns out that four independent coefficients are one too many and the equations are useless as they stand.

Onsager, who laid the foundations of irreversible thermodynamics, showed that if the proper pairs of forces and fluxes are chosen, there are only three independent coefficients to determine: in the absence of a magnetic field the crosscoupling coefficients are equal, i.e., $M_{12} = M_{21}$.

We shall not undertake here a discussion of the various ways in which proper pairs are found. However they are found, they must meet the criterion that the sum of the products of each flux by its force must account for the total rate of entropy production (per unit volume) in the system. In short, we are looking for pairs of forces and fluxes which are linearly related:

$$\mathbf{J}_1 = L_{11}\mathbf{F}_1 + L_{12}\mathbf{F}_2 \quad (3a)$$

$$\mathbf{J}_2 = L_{21}\mathbf{F}_1 + L_{22}\mathbf{F}_2 \quad (3b)$$

where, under the condition that

$$\frac{ds}{dt} = \mathbf{J}_1 \cdot \mathbf{F}_1 + \mathbf{J}_2 \cdot \mathbf{F}_2,$$

$$L_{21} = L_{12}.$$

Perhaps the most physically satisfying method is due to Callen (1). Using what may be called "the method of local equilibrium" we consider the system to be made up of a collection of small volume elements, each of which is in local equilibrium and having its own particular values of the thermodynamic variables. If we neglect thermal expansion, we may write for the i th volume element:

$$T_i \delta S_i = \delta U_i - \bar{\mu}_i \delta N_i$$

where S is the entropy, U the internal energy, $\bar{\mu}$ the electrochemical potential, and N the number of particles. Now the fact that the extensive variables in this equation have different values for the different elements suggests that there must be fluxes of these quantities present in the system, and moreover these fluxes must be related by the same equation as that for the elements, since it holds at every point in the system. Thus if \mathbf{J}_s is the entropy flux, \mathbf{J}_h is the internal energy flux, and \mathbf{J}_p is the particle flux, we must have

$$T\mathbf{J}_s = \mathbf{J}_h - \bar{\mu}\mathbf{J}_p. \quad (4)$$

In the steady state with no heat flow to or from the surroundings, we obviously have

$$\nabla \cdot \mathbf{J}_h = \nabla \cdot \mathbf{J}_p = 0$$

$$\nabla \cdot \mathbf{J}_s = \frac{ds}{dt}.$$

It is convenient to define a fourth flux \mathbf{J}_q , the heat flux, as

$$\mathbf{J}_q = T\mathbf{J}_s$$

following the usual relation between heat and entropy. Substituting this into Eq. (4) we get

$$\mathbf{J}_q = \mathbf{J}_h - \bar{\mu}\mathbf{J}_p$$

or

$$\mathbf{J}_h = \mathbf{J}_q + \bar{\mu}\mathbf{J}_p. \quad (5)$$

We shall use this relation presently to eliminate the internal energy current in favor of the experimentally observable heat current.

Our procedure is to obtain the expression for the time rate of entropy production in terms of experimentally observable quantities, and by comparing the resultant expression with Eq. (3) to identify the proper forces and fluxes. Putting these into a pair of equations of the form of Eq. (3) we are certain of obtaining the Onsager relation between the crosscoupling coefficients. We have

$$T\mathbf{J}_s = \mathbf{J}_h - \bar{\mu}\mathbf{J}_p$$

$$\mathbf{J}_s = \frac{1}{T}\mathbf{J}_h - \frac{1}{T}\bar{\mu}\mathbf{J}_p$$

and taking the divergence of both sides

$$\begin{aligned} \nabla \cdot \mathbf{J}_s &= \frac{ds}{dt} = \nabla \cdot \left(\frac{1}{T} \right) \cdot \mathbf{J}_h + \frac{1}{T} \nabla \cdot \mathbf{J}_h - \nabla \cdot \left(\frac{1}{T} \right) \cdot \bar{\mu}\mathbf{J}_p \\ &\quad - \frac{1}{T} \nabla \bar{\mu} \cdot \mathbf{J}_p - \frac{\bar{\mu}}{T} \nabla \cdot \mathbf{J}_p. \end{aligned}$$

In general $\nabla \cdot \mathbf{J}_p = 0$ and for the case of no heat exchange with the environment $\nabla \cdot \mathbf{J}_h = 0$, so that

$$\frac{ds}{dt} = \nabla \cdot \left(\frac{1}{T} \right) \cdot \mathbf{J}_h - \nabla \cdot \left(\frac{1}{T} \right) \cdot \bar{\mu}\mathbf{J}_p - \frac{1}{T} \nabla \bar{\mu} \cdot \mathbf{J}_p.$$

Now by Eq. (5)

$$\mathbf{J}_h = \mathbf{J}_q + \bar{\mu} \mathbf{J}_p .$$

Therefore

$$\begin{aligned} \frac{ds}{dt} &= \nabla \left(\frac{1}{T} \right) \cdot \mathbf{J}_q + \nabla \left(\frac{1}{T} \right) \cdot \bar{\mu} \mathbf{J}_p - \nabla \left(\frac{1}{T} \right) \cdot \bar{\mu} \mathbf{J}_p \\ &\quad - \frac{1}{T} \nabla \bar{\mu} \cdot \mathbf{J}_p = \nabla \left(\frac{1}{T} \right) \cdot \mathbf{J}_q - \frac{1}{T} \nabla \bar{\mu} \cdot \mathbf{J}_p . \end{aligned}$$

Comparing this with Eq. (3) we may make the following identifications:

$$\begin{aligned} \mathbf{J}_1 &= -\mathbf{J}_p & \mathbf{J}_2 &= \mathbf{J}_q \\ \mathbf{F}_1 &= \frac{1}{T} \nabla \bar{\mu} & \mathbf{F}_2 &= \nabla \left(\frac{1}{T} \right) . \end{aligned}$$

Hence we may write

$$-\mathbf{J}_p = L_{11} \frac{1}{T} \nabla \bar{\mu} + L_{12} \nabla \left(\frac{1}{T} \right) \quad (6a)$$

$$\mathbf{J}_q = L_{12} \frac{1}{T} \nabla \bar{\mu} + L_{22} \nabla \left(\frac{1}{T} \right) \quad (6b)$$

which Callen calls the kinetic equations.

We may now identify the kinetic coefficients (the L_{ij}) with experimentally observable quantities. For example, the electric conductivity σ is defined as the negative ratio of the electric flux $\mathbf{J}_e = e \mathbf{J}_p$ to the potential gradient $\nabla \phi$ with zero thermal gradient, i.e.,

$$\sigma = - \frac{\mathbf{J}_e}{\nabla \phi} \quad \text{for} \quad \nabla T = 0 .$$

For $\nabla T = 0$ the first kinetic equation becomes

$$-\mathbf{J}_p = L_{11} \frac{1}{T} \nabla \bar{\mu}$$

and the electric current is, since $\mathbf{J}_e = e \mathbf{J}_p$,

$$\mathbf{J}_e = \frac{L_{11} e}{T} \nabla \bar{\mu}$$

so that the electric conductivity is

$$\sigma = \frac{L_{11} e}{T} \frac{\nabla \bar{\mu}}{\nabla \phi} .$$

Now $\bar{\mu} = \mu + e\phi$ and $\nabla \bar{\mu} = \nabla \mu + e \nabla \phi$, but in the isothermal case $\nabla \mu = 0$; hence $\nabla \bar{\mu} = e \nabla \phi$ and

$$\sigma = \frac{L_{11} e}{T} \times \frac{e \nabla \phi}{\nabla \phi} = \frac{e^2 L_{11}}{T} .$$

Similarly the thermal conductivity is defined as the ratio of heat flux to the negative of the thermal gradient with the particle current zero:

$$K = - \frac{\mathbf{J}_q}{\nabla T} \quad \text{for} \quad \mathbf{J}_p = 0 .$$

We set $\mathbf{J}_p = 0$ in the first kinetic equation and divide through by L_{11} , getting

$$0 = \frac{1}{T} \nabla \bar{\mu} + \frac{L_{12}}{L_{11}} \nabla \left(\frac{1}{T} \right) .$$

Dividing the second equation by L_{12} gives

$$\frac{\mathbf{J}_q}{L_{12}} = \frac{1}{T} \nabla \bar{\mu} + \frac{L_{22}}{L_{12}} \nabla \left(\frac{1}{T} \right) .$$

Subtracting the second from the first,

$$- \frac{\mathbf{J}_q}{L_{12}} = \left(\frac{L_{12}}{L_{11}} - \frac{L_{22}}{L_{12}} \right) \nabla \left(\frac{1}{T} \right) .$$

Multiplying by L_{12} and recalling that

$$\nabla \left(\frac{1}{T} \right) = - \frac{1}{T^2} \nabla T$$

we have

$$-\mathbf{J}_q = \left(\frac{L_{12}^2}{L_{11}} - L_{22} \right) \left(- \frac{1}{T^2} \nabla T \right)$$

so that

$$K = - \frac{\mathbf{J}_q}{\nabla T} = \frac{L_{11} L_{22} - L_{12}^2}{T^2 L_{11}} .$$

We now have two relations involving the kinetic coefficients. If we can find a third, then all three will be expressible in terms of experimentally observable quantities.

To find the third we turn to the Seebeck effect. Figure 3 is a diagram of a thermocouple made of materials A and B with the left junction at the temperature T_c and the right at the temperature T_h . The terminals of the voltmeter are at the same temperature T_r . Recalling that the Seebeck coefficient is measured under the

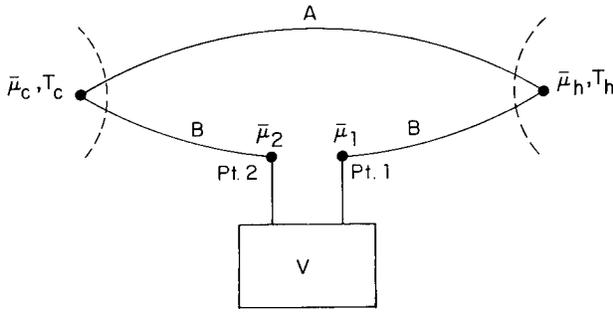


Fig. 3 - Thermocouple composed of materials A and B. Points 1 and 2, the terminals of the potentiometer V , are at the same temperature T_r .

condition of no net electrical current flow (i.e., $\mathbf{J}_e = 0$, hence $\mathbf{J}_p = 0$) we again set $\mathbf{J}_p = 0$ in the first kinetic equation:

$$0 = L_{11} \frac{1}{T} \nabla \bar{\mu} + L_{12} \nabla \left(\frac{1}{T} \right).$$

Solving for $\nabla \bar{\mu}$ we get

$$\nabla \bar{\mu} = -T \frac{L_{12}}{L_{11}} \nabla \left(\frac{1}{T} \right) = \frac{L_{12}}{TL_{11}} \nabla T$$

or

$$d\bar{\mu} = \frac{L_{12}}{TL_{11}} dT.$$

To find the voltage measured we simply sum $\bar{\mu}$ around the circuit. Thus, starting at Pt. 1, Fig. 3, we get

$$\bar{\mu}_h - \bar{\mu}_1 = \int_{T_r}^{T_h} \left(\frac{L_{12}}{TL_{11}} \right)_B dT$$

$$\bar{\mu}_c - \bar{\mu}_h = - \int_{T_c}^{T_h} \left(\frac{L_{12}}{TL_{11}} \right)_A dT$$

$$\bar{\mu}_2 - \bar{\mu}_c = \int_{T_c}^{T_r} \left(\frac{L_{12}}{TL_{11}} \right)_B dT.$$

Adding gives

$$\bar{\mu}_2 - \bar{\mu}_1 = \int_{T_c}^{T_h} \left[\left(\frac{L_{12}}{TL_{11}} \right)_B - \left(\frac{L_{12}}{TL_{11}} \right)_A \right] dT.$$

Now since $\bar{\mu}$ is an energy, the voltage measured by the voltmeter will be related to the difference of the electro-

chemical potentials by

$$V = \frac{1}{e} (\bar{\mu}_2 - \bar{\mu}_1)$$

or

$$V_{AB} = \frac{1}{e} \int_{T_c}^{T_h} \left[\left(\frac{L_{12}}{TL_{11}} \right)_B - \left(\frac{L_{12}}{TL_{11}} \right)_A \right] dT.$$

Since $\alpha_{AB} = dV_{AB}/dT$,

$$\frac{dV_{AB}}{dT} = \alpha_{AB} = \frac{1}{e} \left[\left(\frac{L_{12}}{TL_{11}} \right)_B - \left(\frac{L_{12}}{TL_{11}} \right)_A \right].$$

Recalling the concept of absolute thermoelectric power, i.e.,

$$\alpha_{AB} = \alpha_A - \alpha_B$$

it follows that

$$\alpha_A = - \left(\frac{L_{12}}{eTL_{11}} \right)_A$$

$$\alpha_B = - \left(\frac{L_{12}}{eTL_{11}} \right)_B$$

or in general

$$\alpha = - \frac{L_{12}}{eTL_{11}}$$

and

$$- \frac{L_{12}}{TL_{11}} = e\alpha = S_p.$$

Thus we find the third relationship between the kinetic coefficients and the experimental observables. We shall not go through the algebra here, but straightforward manipulation shows that

$$L_{11} = \frac{T}{e^2} \sigma$$

$$L_{12} = - \frac{T^2}{e} \sigma \alpha$$

$$L_{22} = T^3 \sigma \alpha^2 + T^2 K.$$

It might be well at this point to derive an equation which will give some thermodynamic insight into the nature of α and which will also be useful in developing the Kelvin relations.

The second kinetic equation is

$$\mathbf{J}_q = L_{12} \frac{1}{T} \nabla \bar{\mu} + L_{22} \nabla \left(\frac{1}{T} \right).$$

Now the entropy flux is related to \mathbf{J}_q by $\mathbf{J}_q = T\mathbf{J}_s$; hence

$$\mathbf{J}_s = \frac{\mathbf{J}_q}{T} = \frac{L_{12}}{T} \times \frac{1}{T} \nabla \bar{\mu} + \frac{1}{T} L_{22} \nabla \left(\frac{1}{T} \right).$$

We solve the first kinetic equation (Eq. 6a) for $\nabla \bar{\mu}$ and substitute this in the above expression for \mathbf{J}_s :

$$L_{11} \frac{1}{T} \nabla \bar{\mu} = -\mathbf{J}_p - L_{12} \nabla \left(\frac{1}{T} \right)$$

$$\frac{1}{T} \nabla \bar{\mu} = -\frac{\mathbf{J}_p}{L_{11}} - \frac{L_{12}}{L_{11}} \nabla \left(\frac{1}{T} \right)$$

$$\begin{aligned} \mathbf{J}_s &= \frac{L_{12}}{T} \left[-\frac{\mathbf{J}_p}{L_{11}} - \frac{L_{12}}{L_{11}} \nabla \left(\frac{1}{T} \right) \right] + \frac{1}{T} L_{22} \nabla \left(\frac{1}{T} \right) \\ &= -\frac{L_{12}}{TL_{11}} \mathbf{J}_p + \left(\frac{L_{22}}{T} - \frac{L_{12}^2}{TL_{11}} \right) \nabla \left(\frac{1}{T} \right) \\ &= -\frac{L_{12}}{TL_{11}} \mathbf{J}_p + \left(\frac{L_{11}L_{22} - L_{12}^2}{TL_{11}} \right) \nabla \left(\frac{1}{T} \right) \\ &= S_p \mathbf{J}_p - \left(\frac{L_{11}L_{22} - L_{12}^2}{T^3 L_{11}} \right) \nabla T. \end{aligned}$$

But

$$\frac{L_{11}L_{22} - L_{12}^2}{T^2 L_{11}} = K.$$

Therefore

$$\mathbf{J}_s = S_p \mathbf{J}_p - \frac{K}{T} \nabla T.$$

From this it is apparent that the entropy flux is made up of two terms. The second is the flow of entropy resulting from the flow of heat; the first gives the entropy flow resulting from the particle flow. In other words

$$-\frac{L_{12}}{TL_{11}} (= S_p)$$

is the entropy flux per particle.

For the isothermal case $\nabla T = 0$ and

$$\begin{aligned} \mathbf{J}_s &= -\frac{L_{12}}{TL_{11}} \mathbf{J}_p \\ &= e\alpha \mathbf{J}_p \\ &= \alpha \mathbf{J}_e. \end{aligned}$$

Hence α is the entropy per unit charge flux.

To complete our derivation of the Kelvin relations we must obtain expressions for the Peltier and Thomson coefficients. This is most easily done by returning to the basic equation:

$$T\mathbf{J}_s = \mathbf{J}_h - \bar{\mu}\mathbf{J}_p$$

or

$$\mathbf{J}_h = T\mathbf{J}_s + \bar{\mu}\mathbf{J}_p.$$

We have just seen that

$$\mathbf{J}_s = S_p \mathbf{J}_p - \frac{K}{T} \nabla T.$$

Hence

$$\begin{aligned} \mathbf{J}_h &= TS_p \mathbf{J}_p - K \nabla T + \bar{\mu} \mathbf{J}_p \\ &= (TS_p + \bar{\mu}) \mathbf{J}_p - K \nabla T. \end{aligned}$$

Now taking the divergence (and recalling that $\text{div } \mathbf{J}_p = 0$) gives

$$\nabla \cdot \mathbf{J}_h = S_p \nabla T \cdot \mathbf{J}_p + T \nabla S_p \cdot \mathbf{J}_p + \nabla \bar{\mu} \cdot \mathbf{J}_p - \nabla (K \nabla T).$$

We wish to eliminate $\nabla \bar{\mu}$. From the first kinetic equation (Eq. 6a) we have

$$L_{11} \frac{1}{T} \nabla \bar{\mu} = -\mathbf{J}_p - L_{12} \nabla \left(\frac{1}{T} \right)$$

$$\begin{aligned} \nabla \bar{\mu} &= -\frac{T\mathbf{J}_p}{L_{11}} - \frac{TL_{12}}{L_{11}} \nabla \left(\frac{1}{T} \right) \\ &= -\frac{T\mathbf{J}_p}{L_{11}} + \frac{L_{12}}{TL_{11}} \nabla T. \end{aligned}$$

Recall that

$$\sigma = \frac{e^2 L_{11}}{T}$$

and

$$\frac{L_{12}}{TL_{11}} = -S_p;$$

hence

$$\nabla \bar{\mu} = -\frac{e^2 \mathbf{J}_p}{\sigma} - S_p \nabla T.$$

Substituting,

$$\begin{aligned} \nabla \cdot \mathbf{J}_h &= S_p \nabla T \cdot \mathbf{J}_p + T \nabla S_p \cdot \mathbf{J}_p \\ &\quad - \left(\frac{e^2 \mathbf{J}_p}{\sigma} + S_p \nabla T \right) \cdot \mathbf{J}_p - \nabla (K \nabla T) \\ &= S_p \nabla T \cdot \mathbf{J}_p + T \nabla S_p \cdot \mathbf{J}_p \\ &\quad - \frac{e^2 \mathbf{J}_p^2}{\sigma} - S_p \nabla T \cdot \mathbf{J}_p - \nabla (K \nabla T). \end{aligned}$$

Canceling and rearranging terms,

$$\nabla \cdot \mathbf{J}_h = -\frac{e^2 \mathbf{J}_p^2}{\sigma} + T \nabla S_p \cdot \mathbf{J}_p - \nabla (K \nabla T).$$

Recall that

$$e \mathbf{J}_p = \mathbf{J}_e$$

and

$$S_p = e \alpha;$$

hence

$$\nabla \cdot \mathbf{J}_h = -\frac{\mathbf{J}_e^2}{\sigma} + T (\nabla \alpha) \cdot \mathbf{J}_e - \nabla (K \nabla T). \quad (7)$$

Thus the divergence of the energy current is made up of three terms. The first and third of these are obviously the result of Joule and conducted heat; it follows that the second term must account for the Peltier and Thomson effects.

Recall that the Thomson coefficient is defined for a homogeneous conductor with a temperature gradient by

$$\frac{d}{dx} \left(\frac{dQ}{dt} \right) = \tau \nabla T \cdot \mathbf{J}_e.$$

We may for such a case write the second term in Eq. (7) as

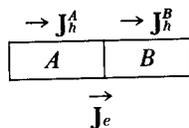
$$T \frac{d\alpha}{dT} \nabla T \cdot \mathbf{J}_e.$$

Comparing this with the above we see that

$$\tau = T \frac{d\alpha}{dT}$$

and the first Kelvin relation has been demonstrated.

The Peltier effect is found at the isothermal junction of two conductors, say A and B:



The Peltier coefficient is defined by

$$\frac{dQ_{\Pi}}{dt} = \Pi \mathbf{J}_e$$

where Π is understood to refer to the heat absorbed (or rejected) at the junction per unit current flux. Since no junction is infinitely sharp, the Peltier effect must refer

to the total heat absorbed (or rejected) in the region where α changes from α_A , its value in material A at temperature T to α_B , its value in material B at the same temperature, i.e.,

$$\begin{aligned} \Pi_{AB} \mathbf{J}_e &= \left(\int_A^B T \frac{d\alpha}{dx} dx \right) \mathbf{J}_e \\ &= T \int_{\alpha_A}^{\alpha_B} d\alpha \\ &= T (\alpha_B - \alpha_A) \end{aligned}$$

or in general

$$\Pi = \alpha T$$

which is the second Kelvin relation.

We shall leave the subject of irreversible thermodynamics here with the statement that it has been applied to the thermomagnetic and galvanomagnetic effects by Callen and others (1-3).

TRANSPORT THEORY

Fundamental Concepts

Thermodynamics, whether reversible or irreversible, can give information on the interrelations between various thermoelectric quantities, but can give us no information on the relation of those quantities and structural details of the materials which exhibit them. To do this, it is necessary to resort to quantum and statistical calculations based on particular models.

We begin by imagining a homogeneous, isotropic bar of conducting or semiconducting material which may carry electrical and thermal currents in the x direction. The rate of generation of heat per unit time per unit volume in the bar is the difference between the electrical energy transported and the divergence of the heat which flows in by thermal conduction, i.e.,

$$\frac{dH}{dt} = J_e - \frac{\partial J_Q}{\partial x}.$$

Our procedure will be to calculate the electrical and thermal currents from first principles and then identify the terms in the result with known experimental quantities.

The electrical current transported by an electron moving with drift velocity v_x is given by the product of v_x and its charge, $-e$. To find the total current J_e it is necessary to sum over all the electrons in a unit volume. Thus,

$$J_e = -2e \int v_x f(m) dV_m$$

where V_m gives the number of modes available in a unit volume, $f(m)$ is the probability that a particular mode is active, and the factor 2 arises because of spin degeneracy.

In more familiar terms

$$J_e = -\frac{2e}{8\pi^3} \int v_x f(\mathbf{k}) dV_{\mathbf{k}}$$

where now $dV_{\mathbf{k}}$ refers to the incremental density of states in \mathbf{k} space. Similarly, the heat current carried by a single electron is given by the product of its kinetic energy ϵ and its velocity v_x ; summing over all electrons in a unit volume gives

$$J_q = \frac{1}{4\pi^3} \int \epsilon(\mathbf{k}) v_x(\mathbf{k}) f(\mathbf{k}) dV_{\mathbf{k}}$$

The most general procedure from this point on may be summarized in three steps:

1. Find the expression for the distribution function f .
2. Pick the proper relation between ϵ and k .
3. Carry out the integrations.

As the reader is undoubtedly aware, any attempt to adequately describe the problems and the approaches used in the first two steps would cover wide areas of the solid state theory. The second step, in particular, which is based on selecting the proper model of band structure to account for particular experimental results, could well have a library written on it. Accordingly, we shall restrict ourselves to the simplest approaches which still yield results of some general validity. We will assume the standard quasi-free-electron model and the existence of an isotropic relaxation time to describe the interaction between electrons and lattice.

However, even these assumptions involve a considerable amount of algebraic detail which would be out of place here. Hence we shall confine ourselves to a qualitative description of them. An excellent review by Frank Blatt appears in the Solid State Physics Series, Vol. 4, for 1957.

Now the distribution function which holds when electrons are in static equilibrium is just the Fermi-Dirac expression:

$$f_0 = \frac{1}{e^{(\epsilon - \epsilon_f)/kT} + 1}$$

The calculation of either electrical or heat currents from this distribution function would yield zero, since it is spherically symmetrical in the velocities. It is apparent that the application of a field, or a gradient sets up a steady-state condition in which a new dynamic equilibrium holds, and correspondingly, there is a new distribution. This new distribution must contain, more

or less explicitly, the mechanism by which the electrons interact with the applied fields. The well-known Boltzmann equation describes the effects of these interactions in general terms, so one begins with this equation and puts in the details pertinent to the model he is using. Of course the model is selected to give agreement with experimental results, and one finds paper after paper in the literature dealing with models of sometimes frightful complexity.

As we have said the simplest model which still has some degree of generality assumes a relaxation time τ^{\dagger} and a spherical dependence of ϵ on k . The relaxation time τ^* is defined in the following way: If f is the dynamic distribution function resulting from the application of external forces, then if these forces are instantaneously removed, the electronic system relaxes from f to the static distribution f_0 exponentially with a characteristic time τ^* , i.e.,

$$-\frac{\partial}{\partial t} (f - f_0) \propto (f - f_0)$$

$$\frac{\partial}{\partial t} (f - f_0) = -\frac{1}{\tau^*} (f - f_0)$$

Focusing our attention on electrical fields for the moment, we know that the power which the electrons gain from the field, namely, the product EJ , instantly and completely appears as heat in the lattice. In other words, the electrons are essentially in thermal equilibrium with the lattice and immediately pass on to the lattice any extra energy obtained. The reciprocal of τ^* is thus a measure of the strength of electron lattice interaction; it also follows that the stronger this interaction, the less will the dynamic distribution differ from the static one. In most cases, this interaction is so strong that the dynamic distribution is adequately described as a first-order perturbation of the static one. In fact, we can write

$$f = f_0 - \tau^* \left(\frac{\partial f_0}{\partial k_x} \frac{eE}{h} + v_x \frac{\partial f_0}{\partial t} \frac{dT}{dx} \right)$$

where the field E and thermal gradient appear explicitly. Substitution of this expression for f in the equations yields, since integration over the static distribution gives zero

$$J_e = -e \left(eE + a \frac{dT}{dx} \right) K_1 - e \frac{1}{T} \frac{dT}{dx} K_2$$

$$J_q = \left(eE + a \frac{dT}{dx} \right) K_2 + \frac{1}{T} \frac{dT}{dx} K_3$$

$\dagger \tau^*$ is used here to designate relaxation time. The similarity of this τ for the Thomson coefficient is unfortunate; the connection between the two is quite indirect.

where

$$a = kT \frac{d}{dT} \left(\frac{\epsilon_f}{kT} \right)$$

and the K 's represent the transport integrals, which are given by

$$K_i = -\frac{1}{\pi h^3} \int \epsilon^{i-1} \left(\frac{\partial \epsilon}{\partial k_x} \right)^2 \frac{\partial f_0}{\partial \epsilon} \tau^*(k) dV_k.$$

We shall have a little to say later concerning the evaluation of the transport integrals. For the present we proceed with our derivation of the Seebeck coefficient, continuing to denote them as above.

We note that in our expression for the accumulation of heat the three terms J_e , J_q , and E all appear. It is convenient to eliminate one of them between the derived expressions for J_e and J_q . So we eliminate E , perform the indicated differentiation on J_q and form the expression for dH/dt . This yields

$$\begin{aligned} \frac{dH}{dt} = & -\frac{J_e^2}{e^2 K_1} - \frac{J_e}{e} \left[a - T \frac{d}{dT} \left(\frac{K_2}{TK_1} \right) \right] \frac{dT}{dx} \\ & - \frac{d}{dx} \left[\left(K_3 - \frac{K_2^2}{K_1} \right) \frac{1}{T} \frac{dT}{dx} \right]. \end{aligned}$$

From comparison of this with Eq. (7), the first term evidently refers to the Joule heating, so that

$$\sigma = -e^2 K_1.$$

The third term is independent of the current but involves dT/dx , so it refers to the conducted heat:

$$\kappa = -\frac{K_3 K_1 - K_2^2}{TK_1}$$

Finally we note the second term, which involves the product of the current J_e and the temperature gradient dT/dx ; recalling that the Thomson heat is given by $-J\tau dT/dx$, we see that the Thomson coefficient is

$$\tau = \frac{1}{e} \left[a - T \frac{d}{dT} \left(\frac{K_2}{TK_1} \right) \right].$$

Recalling that

$$a = kT \frac{d}{dT} \left(\frac{\epsilon_f}{kT} \right),$$

we have

$$\tau = -\frac{T}{e} \times \frac{d}{dT} \left[\frac{K_2}{TK_1} - k \frac{\epsilon_f}{kT} \right].$$

Recalling the first Kelvin relation $\tau = -T d\alpha/dT$,

$$\begin{aligned} \alpha &= \frac{1}{e} \left[\frac{K_2}{TK_1} - k \frac{\epsilon_f}{kT} \right] \\ &= \frac{k}{e} \left[\frac{K_2}{kT K_1} - \frac{\epsilon_f}{kT} \right]. \end{aligned} \quad (8)$$

At this point it might be well to show a figure (Fig. 4) which indicates in more physical terms the meaning of the expression we have just derived.

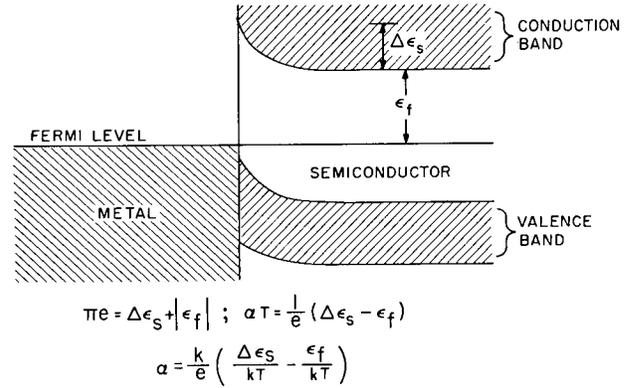


Fig. 4 - Energy diagram of the junction between a metal and an n -type semiconductor. In the metal an electron is at the Fermi level; in the semiconductor it is a distance $\Delta\epsilon_s$ above the edge of the conduction band which in turn is ϵ_f above the Fermi level.

Electrons traveling in the metal will have an average transport energy $\Delta\epsilon_m$ negligibly larger than that of the Fermi level. Upon crossing the junction into the semiconductor, their energy must increase considerably, first because of the height of the conduction band above the Fermi level, and second because they generally will have a higher average energy of transport. Thus the Peltier heat absorbed by the average electron is given by

$$\Pi_{m-semi} e = (\Delta\epsilon_{semi} + |\epsilon_f|)$$

if $\Delta\epsilon_m$ is negligibly small (i.e., set the thermoelectric activity of the metal equal to zero).

Thus

$$\Pi = \frac{1}{e} (\Delta\epsilon_{semi} - \epsilon_f)$$

and since $\Pi = \alpha T$,

$$\alpha_{semi} = \frac{1}{eT} [\Delta \epsilon_{semi} - \epsilon_f]$$

$$= \frac{k}{e} \left[\frac{\Delta \epsilon_{semi}}{kT} - \frac{\epsilon_f}{kT} \right].$$

Comparing this with equation (8), K_2/K_1 gives the average energy of transport referred to the band edge, and the absolute Seebeck coefficient is directly related to the average energy of the electrons participating in the conduction process.

Transport Integrals

Returning now to the transport integrals: We have stated them in the most general form consistent with our solution of the Boltzman equation. To go further we must put in details of the ϵ and k relationship, the density of states, and the functional dependence of the relaxation time. Again we warn that very complex and heinous models may be required, but we shall stick to the simplest model, a single isotropic band of standard form, i.e.,

$$\epsilon = \frac{\hbar^2}{2m^*} k^2.$$

Putting in the quadratic dependence of ϵ we get

$$K_i = -A' \int_0^\infty \tau^* \epsilon^{i+1/2} \frac{df_0}{d\epsilon} d\epsilon$$

where we have replaced the garbage in front of the integral sign with A' . (Since we need only calculate the ratio of K_2 to K_1 , there is no need to worry about constants.) The next step is to make the assumption that τ^* may be written as the product of a function $\phi(T)$ and ϵ to some power s .

$$\tau^* = \phi(T) \epsilon^s.$$

In this context this assumption appears rather arbitrary. Since τ^* is related to the strength of electron-lattice interaction, it is physically reasonable that it should depend on the electron energy. Theoretically if the electrons are scattered primarily by the low-frequency thermal vibrations of the atoms in the lattice, s would be $-1/2$; on the other hand for ionized impurity scattering s should be $3/2$.

Since T is independent of ϵ , we may remove it from the integral and lump it with the other constant, obtaining

$$K_i = -A \int_0^\infty \epsilon^{i+s+1/2} \frac{df_0}{d\epsilon} d\epsilon.$$

Integrating once by parts yields

$$K_i = A (i + s + 1/2) \int_0^\infty \epsilon^{i+s+1/2} f_0(\epsilon) d\epsilon$$

which is easily transformed to

$$K_i = A (i + s + 1/2) (kT)^{i+s+1/2} F_{i+s-1/2} \left(\frac{\epsilon_f}{kT} \right)$$

where

$$F_{i+3-1/2} \left(\frac{\epsilon_f}{kT} \right)$$

is the standard Fermi integral.

Thus, in general

$$\frac{K_2}{K_1} = \frac{(kT) (s+5/2) F_{s+3/2}}{(s+3/2) F_{s+1/2}}$$

illustrating how this term in the Seebeck coefficient depends on the type of scattering experienced by the carriers.

While we have taken a general case, it should be stated that it is not always necessary to evaluate Fermi integrals. If one is dealing with semiconductors of low carrier concentration, Boltzman statistics may be used. In this case K_2/K_1 is simply $kT (s + 5/2)$. If the conductor is fully degenerate, the usual series expansion about the Fermi level may be used.

We have indicated that ϵ_f is measured downward from the edge of the conduction band, so that its magnitude is to be added in calculating the Seebeck coefficient of semiconductors. In the case of metals, it lies at or above the band edge and is subtracted from the K_2/K_1 term. The constant k/e has the approximate value of $86 \mu v/\text{degree}$, so that one sees that in metals the average energy of transport is quite small indeed, being roughly $0.1 kT$.

FIGURE OF MERIT

In any discussion of thermoelectric materials and devices the figure of merit, denoted by Z , inevitably crops up. One should be warned that the Z which turns up in efficiency and coefficient of performance calculations is not the Z which is used in evaluating the merit of a material, though the two are connected. The Z appearing in efficiency calculations can be simply derived from a quite elementary discussion of efficiency. Its value lies primarily in the fact that it relates the inherent characteristics of the materials used to the efficiency of the device.

The Z for a thermocouple is

$$Z = \frac{(\alpha_1 - \alpha_2)^2}{(\sqrt{\rho_1 K_1} + \sqrt{\rho_2 K_2})^2}.$$

One cannot rigorously separate the contributions of the legs (they are usually of different materials), but this is done anyway and the Z for a material is written

$$Z = \frac{\alpha^2}{\rho K} \text{ or } \frac{\alpha^2 \sigma}{K}.$$

In the vicinity of room temperature, Z values up to 3×10^{-3} are common.

One of the earliest attempts to put solid-state theory to use in the thermoelectric field was an analysis of the Z factors: α , σ , and K . These are all functions of the Fermi level, which depends on carrier concentration as well as band structure and temperature. Largely ignoring all save the carrier concentration and using nondegenerate statistics, the Russian Ioffe (4) came up with a now famous diagram (see Fig. 5), which shows a broad maximum in the $\alpha^2 \sigma$ curve in the $10^{19}/\text{cm}^3$ range of carrier concentration. Inclusion of the K curve did not change the position of the maximum.

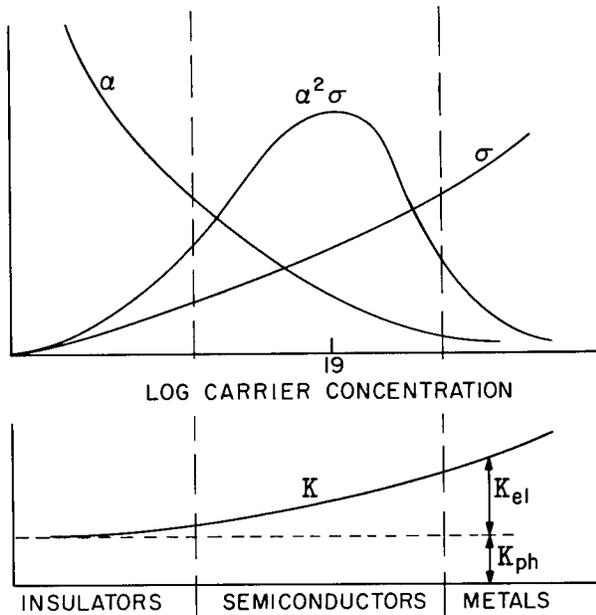


Fig. 5 - Qualitative plot of α , σ , K , and $\alpha^2 \sigma$ against the logarithm of carrier density. The total thermal conductivity K is composed of the sum of the lattice conductivity K_{ph} and the electronic contribution K_{el} . By the Wiedemann-Franz law the latter has the same form as the σ curve.

It will be noticed at once that this is in the region of partial degeneracy, so that Ioffe's calculation was immediately called into question. However, a far more rigorous approach to the problem has not changed his conclusions except to indicate that the variations in effective mass from material to material may cause a

shift of the optimum carrier density to as low as 10^{18} or as high as 10^{21} . The corresponding optimum Seebeck coefficient lies between 200 and 300 $\mu\text{v}/\text{degree}$ or approximately 3 times k/e .

Boiled down to its essentials the attainment of a high figure of merit depends on the discovery, or development, of a material for which the expression

$$\frac{\mu m^{*3/2}}{K_{ph}}$$

is as large as possible. In this expression μ is the carrier mobility,† m^* is the effective mass, and K_{ph} is the lattice thermal conductivity.

Possible approaches to this problem are legion, and a comprehensive survey of them will not be made here. There are, however, certain ideas that are interesting to mention.

It has been suggested that materials having nonspherical energy surfaces may, through the increase in effective mass in a particular direction, offer a means of improving Z . If, for example, a material having ellipsoidal energy surfaces, but otherwise comparable with an isotropic one, is properly oriented, one should realize a high effective mass in the direction of the main axis. However, the concomitant anisotropy of lattice structure could give a very large increase in K_{ph} in the same direction, so that a higher Z value may or may not be found. In general, nonspherical energy surfaces are also associated with materials of low crystal symmetry. Here additional scattering of electrons by transverse lattice modes is to be expected, with a concomitant decrease in the mobility. Whether the decrease in μ will offset the increase in m^* is hard to say.

Another rule-of-thumb calculation indicates that the presence of multivalley energy surfaces may well lead to larger Z values. However, the presence of intervalley scattering can be expected to reduce both the Seebeck coefficient and the mobility, so that an increase in Z is not certain. Present theory does not give any information on how intravalley scattering may be maximized and intervalley and transverse mode scattering minimized.

We have already mentioned that large values of s are realized for ionized impurity scattering, giving an increased value for α . However, it turns out that the resulting reduction in mobility outweighs the increase in α . In fact, ionized impurity scattering must be kept small in thermoelectric materials; a high dielectric constant has this effect, and it has been suggested by Ure that this is a major reason for the success of Bi_2Te_3 .

A narrow band gap means that a material will go intrinsic at relatively low temperatures. For refrigeration this may not pose a problem, but it puts a distinct

† μ shall stand for carrier mobility for the remainder of this paper.

limitation on the use of the material as a generator since the appearance of minority carriers sets up an opposing Seebeck voltage. If the mobilities of the carriers are about equal, the Seebeck coefficient falls drastically with the onset of intrinsic conduction: in addition the ambipolar diffusion of hole-electron pairs increases the thermal conductivity enormously.

Ioffe (4) suggested a means of increasing the ratio μ/K_{ph} by appropriate alloying of semiconductors. Generally speaking, his idea was this: If a given compound A is alloyed with another compound B in which the atoms are of different mass but have a similar valence electron structure, then we might expect phonon propagation to be reduced with carrier mobility only slightly affected. This is particularly reasonable since the phonons have a wavelength of about a lattice constant and should be strongly scattered by the presence of a foreign atom in the lattice. On the other hand, if the electronic structure of the foreign atom is not too different from the indigenous ones, the electrical disturbance produced by it may well be small compared to the electronic wavelength, which may be several lattice constants in length. By and large, his idea has been proven correct. Improvement in the Z value has been achieved, though it is not as great as Ioffe hoped.

Two theories are extant on the mechanisms involved. The Russians have suggested that electrons in a semiconducting compound move primarily on the metal sublattice while holes travel on the nonmetal one. For example, alloying n-type Bi_2Te_3 and Bi_2Se_3 should improve the Z of the material, because electron mobility is little affected but phonon conduction is reduced. Correspondingly, alloying Bi_2Te_3 with Sb_2Te_3 should have little effect on hole mobility. They claim experimental verification for this, but the mobilities given for the pure compounds are so much lower than those found by American workers that the validity of their results has been questioned.

On the other hand Cornish (5) has proposed that if the ratio of the ionic radii of the metallic ions to the nonmetallic ions, R_m/R_{non} is greater for the alloying compound B than for the original compound A, the ratio of the electron to hole mobilities increases. Thus for n-type material, R_m/R_{non} of B should be larger than for A, while for p-type material, R_m/R_{non} for B should be less than for A. In compounds studied so far, the two theories make the same predictions, so at present there is little reason to choose between them.

Attempts have been made to predict a theoretical limit to the figure of merit, but the fact is that in its

present condition the theory of the solid state is incapable of such a prediction. However, Donahoe (6), for example, has predicted an upper limit of 17×10^{-3} at room temperature. This is typically based on favorable but reasonable values of the pertinent parameters. (By reasonable we mean reasonable in the light of present experimental knowledge.) If his prediction is correct, the future of thermoelectric power generation may be limited, but the attainment of only half this upper limit would open up large areas to commercial thermoelectric applications.

It should be noted that claims that the product ZT can never be greater than 1 to 1.5 belong in the same category as the attempts to predict an upper limit to Z .

CONCLUSION

Because of the inability of the present theory to make quantitative predictions concerning the best we can hope for or even to indicate where to look for it, a large effort has gone into a search for better materials. That the effort has been largely Edisonian in character is not the fault of the investigators. Progress has been made, particularly in the area of high temperature materials, but there have been no breakthroughs. The feeling is becoming more and more widespread that any real advance in the application of thermoelectric effects will have to be based on a solid theoretical foundation that is presently lacking. This calls for nothing less than an extension and perfection of the present theory. While the carrier concentration of interest appears to fall midway between the metals and present semiconducting devices, it is too early to concentrate in this area.

REFERENCES

1. Callen, H. B., Phys. Rev. 73:1349 (June 1, 1948)
2. DeGroot, S. R., "Thermodynamics of Irreversible Processes," Amsterdam:North Holland Pub. Co. (New York: Interscience),
3. Domenicali, C. A., Rev. Mod. Phys. 26:237 (Apr. 1954)
4. Ioffe, A. F., "Semiconductor Thermoelements and Thermoelectric Cooling," London:Infosearch Ltd., 1957
5. Cornish, A. J., unpublished
6. Donahoe, F. J., Elect. Engr. 79:488 (1960)

* * *