New Method for Computing the Weak-Field Hall Coefficient

R. S. ALLGAIER
United States Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland
(Received 9 August 1967)

The factor r in the expression $R_0=r/ne$ for the weak-field Hall coefficient is computed for a wide variety of metallic band models. The technique employs a planar-faced approximation to the true Fermi surface, and this greatly simplifies the mathematics. Strongly distorted Fermi surfaces, multivalley models, anisotropic scattering, and band-filling effects are studied, and the results are used to develop a general theory of the behavior of the factor r. The effects of shape and scattering anisotropies, the question of their equivalence, and the reason why a Hall coefficient changes sign are discussed. Finally, the results are applied to the model of Cooper and Raimes, to lead-pipe models, and to experimental Hall data on the noble metals.

1. INTRODUCTION

THIS paper describes a new way to compute the ordinary weak-field Hall coefficient which can be applied to almost any kind of band model.[1] The essence of the method is to approximate an actual constant-energy surface of the band structure by a planar-faced energy surface (PFES). The main advantage of the approach is its simplicity, which persists for the variety of models studied in this paper.

The weak-field Hall coefficient R_0 may be written in the form

$$R_0=r/ne, \tag{1}$$

where n is the number of carriers per unit volume, e is the carrier charge, and r is a *dimensionless* factor which depends on the details of the model.[2] For reasons to be explained shortly, we call r the *mixing* factor.

All results in this paper will be expressed in terms of r. The charge e in Eq. (1) always means the negative electronic charge, but we will never explicitly use this fact. We need only note that a positive r is the normal case for which the signs of R_0 and e are the same, and a negative r is the anomalous case which is conventionally described as p-type behavior.

It will sometimes be convenient to redefine r by substituting n_h for n in Eq. (1), where n_h is the density of missing carriers, i.e., holes. This affects the magnitude, but not the sign, of r.

2. WHAT DETERMINES THE VALUE OF r?

The Hall angle θ for an individual carrier is defined by $\tan\theta=\mu H$, where μ is the drift mobility and H is the magnetic field intensity. We have $\mu=e\tau/m^*$, where τ is the scattering time and m^* is the effective mass. Hence the Hall angle always depends on these two fundamental carrier properties.

If all contributing carriers in a given model have the same Hall angle, $r=1$; otherwise, different Hall angles mix, and $r\neq 1$, except when opposing effects accidentally cancel.

The effective mass and scattering time may depend on carrier energy; for a given energy, they may depend on direction (i.e., on the carrier's position in momentum space); and even at a given point in momentum space there may be two or more carriers with different properties (a multiband model in the reduced Brillouin zone representation).

In this paper, we will concentrate on single-band, metallic models. In realistic examples of such models, the main source of Hall-angle mixing will almost always be the variation of τ and m^* over the Fermi surface.

The variation in the response of *different* carriers to the *same* force determines r in any particular case. This is one kind of anisotropy. A second kind is the anisotropic response of a given carrier to forces in different directions. For noncubic models, this may lead to a value of r which depends on the direction of applied force.

The factor r is sometimes called an anisotropy factor in cases of metallic models. This can be confusing, because of the two distinct types of anisotropy just mentioned. It can also be misleading, since mixing can occur without anisotropy (e.g., $r\neq 1$ for a simple metallic model with two spherical, parabolic bands) while an anisotropic Fermi surface does not always produce mixing ($r=1$ for a single, ellipsoidal energy surface).[3]

For these reasons it seems preferable to treat r as a mixing factor. It is always the mixing of different Hall angles which makes $r\neq 1$.

3. PROBLEM OF THE WEAK-FIELD HALL COEFFICIENT

The first problem which arose in connection with the Hall coefficient was the occurrence of both negative and positive Hall coefficients in various common metals.[4] Nearly 50 years passed before the quantum-mechanical band theory of solids finally provided the solution to this problem. The same theory, however, made it

[1] Some examples have already been described: R. S. Allgaier, Bull. Am. Phys. Soc. **12**, 398 (1967); Phys. Rev. **158**, 699 (1967).

[2] Hall coefficient behavior in numerous models is discussed in A. C. Beer, *Galvanomagnetic Effects in Semiconductors* (Academic Press Inc., New York, 1963).

[3] N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Clarendon Press, Oxford, England, 1936), p. 283.

[4] Early measurements are summarized in L. L. Campbell, *Galvanomagnetic and Thermomagnetic Effects: The Hall and Allied Phenomena* (Longmans, Green and Company, Inc., New York, 1923), pp. 120–124.

possible to predict complicated band models undreamed of in earlier days.

All of the *simple* calculations of r which have been carried out to date have been based on spherical or ellipsoidal energy surfaces. Even small deviations from these nonmixing shapes lead to enormous complications. For example, Cooper and Raimes[5] considered a metallic model with slight anisotropies in the Fermi surface and scattering time. They found that these modifications introduced about 100 terms into each of two integrals that had to be evaluated.

Another long computation[6] was based on Cohen's nonellipsoidal, nonparabolic model.[7] It is one of very few examples of a strongly distorted Fermi surface for which the integration over the Fermi surface may be carried out exactly. The result was very complicated.

The present-day problem of the Hall coefficient, therefore, is to compute reasonably accurate values of r for the many kinds of band models now known to occur in metals. The factor r has become a rather nebulous quantity in today's world of monstrous Fermi surfaces, and the ordinary weak-field Hall coefficient does not enjoy a very high reputation as a useful solid-state parameter.

4. GENERAL DESCRIPTION OF THE PFES METHOD

A. Why a PFES?

It might seem at first that substituting a PFES for the actual curved Fermi surface is a large step in the wrong direction, since the conventional Jones-Zener weak-field solution to the Boltzmann equation[8] breaks down for a PFES. This occurs because carriers drifting across an edge of a PFES undergo a large change in direction, and this conflects with the assumption of weak magnetic fields.

The kinetic method of Shockley,[9] McClure,[10] and Chambers[11] can be applied to a PFES, but it is rather long and mathematical. So far as we know, this alternative method has been used to compute r for just two PFES models, a cube[12] and an octahedron,[13] assuming degenerate statistics and an isotropic scattering time. The results, surprisingly enough, were simply $r=\frac{1}{2}$ and $\frac{2}{3}$ for the cube and octahedron, respectively.

We showed recently that these same two results could be obtained using a very short and elementary geometrical approach to the problem.[1] It turns out that the same approach may easily be extended to a much larger class of models, without losing its simplicity.

Substituting a PFES for the actual Fermi surface imposes only a slight limitation on the utility of the method. We will present examples which suggest that rather crude PFES approximations can lead to values of r within 10%, and sometimes within 5%, of the correct result. Furthermore, we will demonstrate several advantages which are specifically associated with the PFES approach.

B. Outline of the Method

The procedure described below applies to all of the models investigated in this paper. The electric and magnetic fields, E and H, are oriented in the positive x and z directions, respectively. The sample is assumed to have no transverse boundaries so that the electric field may be held fixed in magnitude and direction. The applied fields and resulting components of the current density i are shown in Fig. 1. Under these conditions, the weak-field Hall coefficient is given by[1]

$$R_0=E_H/i_xH=E(-i_y)/i_x^2H. \qquad (2)$$

This formula is appropriate for all of the cases considered in this paper, viz., models with cubic symmetry, or models in which the x and y directions are equivalent (i.e., models with isotropic conductivity in the x-y plane).

The formula is also correct for cases in which the conductivity in the x-y plane is not isotropic, but R_0 is also anisotropic, and the value given by Eq. (2) will differ from that obtained when the current, rather than the electric field, is parallel to the x axis.

The figures in this paper which illustrate the response of the carriers to the applied forces actually correspond to the behavior of positive charges. We find it less confusing to consider the response of positive charge to applied fields. This does not affect the results, which are expressed in terms of r, since r never contains the charge e.

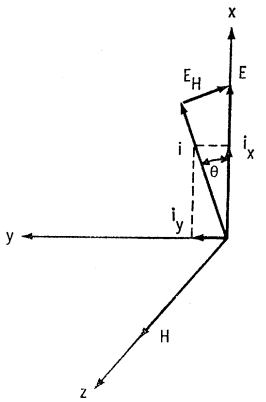

FIG. 1. Configuration of fields (E,H) and currents (i_x,i_y) used to compute the Hall coefficient. θ is the Hall angle.

[5] J. R. A. Cooper and S. Raimes, Phil. Mag. 4, 145 (1959).
[6] R. S. Allgaier, Phys. Rev. **152**, 808 (1966).
[7] M. H. Cohen, Phys. Rev. **121**, 387 (1961).
[8] H. Jones and C. Zener, Proc. Roy. Soc. (London) **A145**, 268 (1934).
[9] W. Shockley, Phys. Rev. **79**, 191 (1950).
[10] J. W. McClure, Phys. Rev. **101**, 1642 (1956).
[11] R. G. Chambers, Proc. Roy. Soc. (London) **A238**, 344 (1957).
[12] C. Goldberg, E. Adams, and R. Davis, Phys. Rev. **105**, 865 (1957).
[13] H. Miyazawa, in *Proceedings of the International Conference on the Physics of Semiconductors, Exeter* (The Institute of Physics and the Physical Society, London, 1962), p. 636.

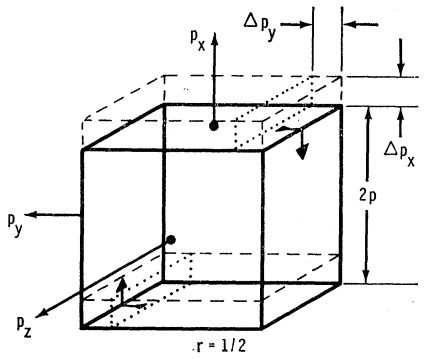

FIG. 2. A cubic Fermi surface in momentum space. The displaced distribution (height $=\Delta p_x$) which produces i_x is shown by dashed lines. The edge-crossing part (width $=\Delta p_y$) of the displaced distribution is separated from the rest by dotted lines. Carrier-drift arrows identify the edges across which carriers flow to produce i_y.

We illustrate the procedure by reviewing the case of a cube. The result for the cube was derived earlier,[1] but the present version eliminates some unnecessary steps.

Figure 2 shows a cubic Fermi surface of edge length $2p$, centered on the origin of momentum space. The electric field displaces the occupied region of momentum space along the p_x axis, as indicated by the dashed lines in the figure. The magnitude of the shift, for this and all other cases to follow, is $\Delta p_x = eE\tau$. The longitudinal current produced by the displaced carriers is the product of (a) the carrier density in the newly occupied and newly vacated states, (b) the charge e, and (c) the velocity v_x on the top face of the Fermi surface. The result is

$$i_x = (2/h^3)(8p^2 eE\tau)(ev_x). \tag{3}$$

The Lorentz force causes part of the displaced distribution (separated from the rest by dotted lines in Fig. 2) to drift across the upper right edge of the Fermi surface and this produces a transverse current. There is an equal contribution from carriers no longer drifting across the lower left edge. The width of the edge-crossing volume of carriers is $\Delta p_y = ev_x H\tau$, and the transverse current is therefore

$$i_y = (2/h^3)(4peE\tau)(ev_x H\tau)(ev_y). \tag{4}$$

Equations (3) and (4) are written in what might be called a $(\Delta n)ev$ notation. The result becomes simpler if we change to a $ne(\Delta v)$ form. To do this we use the relation

$$n = (2/h^3)(2p)^3 \tag{5}$$

between the total carrier concentration and the occupied volume of momentum space. Combining Eqs. (1) through (5) and using the symmetry-imposed condition $v_x = v_y$ leads immediately to

$$r = \tfrac{1}{2}. \tag{6}$$

We did not have to know the magnitude of the carrier velocity on the Fermi surface, nor specify the relation between carrier velocity and energy (i.e., the dispersion relation), to obtain this result. We did assume implicitly that "nearby" energy surfaces (within a few kT of the Fermi energy) also had planar faces which were parallel to the corresponding faces of the Fermi surface. This makes the velocity constant over any given face.

The method outlined above neglects the change in the longitudinal current which occurs when the magnetic field is turned on. The fractional change in i_x is of order $(\mu H)^2$. This is a higher-order correction, negligible for R_0, but it prevents the results from being used to compute the weak-field magnetoresistance.

In Appendix A, we show that the computation of the above result may be extended, without complication, to the case of classical statistics.

C. Procedure for Oblique Faces and Edges

In the more general case, the PFES will involve obliquely oriented faces and edges, such as those shown in Fig. 3. The contribution to i_x from a given face is determined from the x component of the velocity on that face, and from the volume swept out by the face as it is displaced in the p_x direction. That volume is simply $eE\tau$ times the *projected* area A' of the face on a p_y-p_z plane. This projected area is usually easier to determine than the area of the oblique face itself.

Similarly, the contribution to i_y from the volume of carriers crossing a particular edge is computed from three projected lengths: the p_z projection of the edge length, the p_x displacement $eE\tau$, and the p_y component of the displacement along the face due to the Lorentz force. This force generally involves both v_x and v_y, but only v_x appears in the p_y projection. Also needed is the *change* in v_y in going from the initial to the final face (in the case of the cube, the initial value of v_y was zero).

Hence for the general case shown in Fig. 3, the contribution to i_y from the carriers passing from face 1

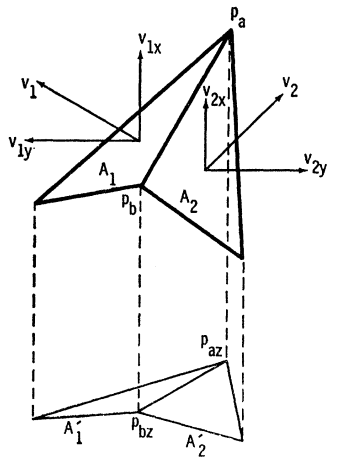

FIG. 3. Two adjacent, oblique faces of a PFES, intersecting in an oblique edge, and the projections needed in the computation.

to face 2 across edge $p_a p_b$ (and from the two faces related to them by inversion) is

$$i_y(1 \rightarrow 2) = (2/h^3)(p_{bz} - p_{az})(2eE\tau) \times (ev_{1x}H\tau)(e)(v_{2y} - v_{1y}). \quad (7)$$

D. Constant-Shape Approximation

In the case of the cube, the plane-parallel (v=const) condition automatically means that the over-all shape of the constant-energy surface does not change with energy. This follows because all faces are equivalent by symmetry.

When the PFES is composed of two or more kinds of nonequivalent faces, the plane-parallel condition may be imposed, but the overall shape of the PFES may still change with energy. To simplify the results, we will always make a constant-shape approximation. Again, it need only apply within a few kT of the Fermi level.

This approximation means that the separation of each face of the Fermi surface from the corresponding face of a nearby surface will be proportional to the distance of that face from the origin of momentum space. As a result, the velocities on nonequivalent faces of the Fermi surface will be inversely proportional to the corresponding distances. Since r is a dimensionless quantity, this relation among velocities is sufficient to eliminate all velocities from any given result for r.

We note that specifying a constant shape does not require specifying the dispersion relation. We also note that the constant-shape approximation is very realistic for almost all metals having a Fermi energy much larger than kT, and this is the case in which we are most interested.

5. RESULTS AND THEIR DISCUSSION

A. PFES Approximations to a Sphere

Figure 4 presents a series of PFES approximations to a sphere. The starting point, shown in part (a) of the figure, is indeed a crude approximation—a cube—which, as we already know, gives $r=\frac{1}{2}$.

In Fig. 4(b) we have cut off the corners of the cube by connecting the midpoints of adjacent edges, thus creating a PFES with eight {111} faces and six {100}

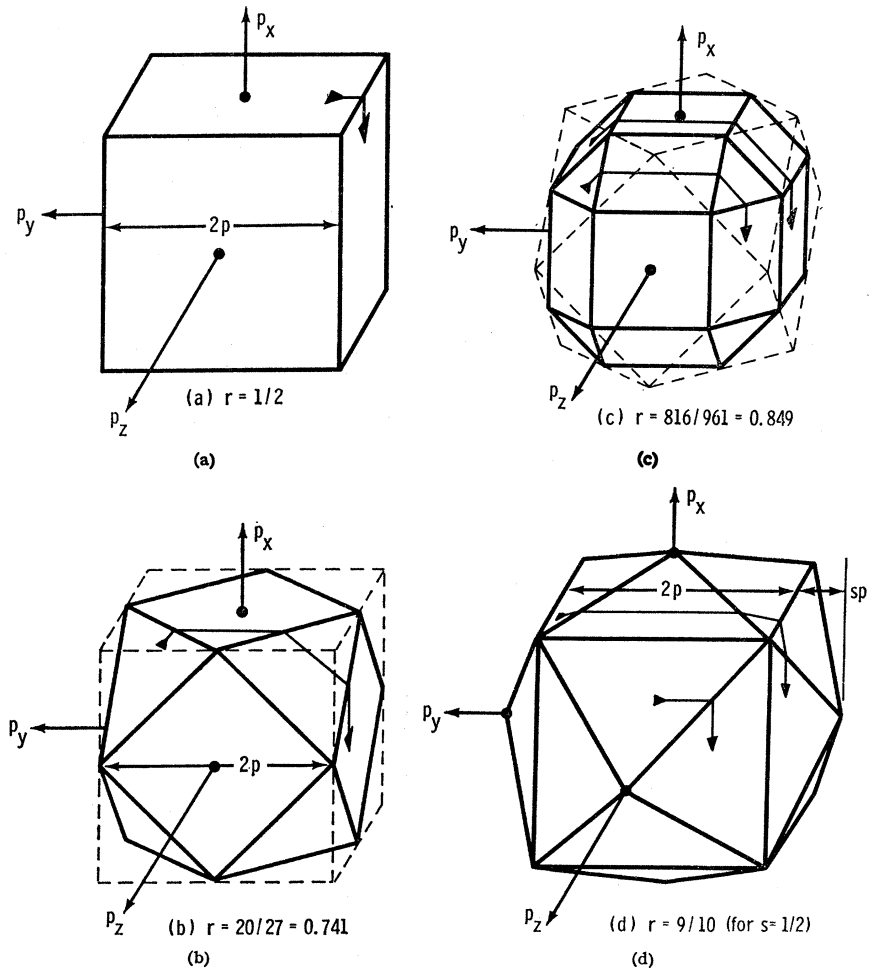

FIG. 4. A series of PFES approximations to a spherical Fermi surface. The edges contributing to the transverse current are marked by arrows (if they lie on the upper front quarter of each surface).

faces. The result is

$$r=20/27=0.741. \tag{8}$$

Cutting off the corners has taken us approximately half way from the cubic value $\frac{1}{2}$ to the spherical value of unity.

In Fig. 4(c), we again connect midpoints of adjacent edges, thus creating twelve {110} faces, in addition to the 14 faces already present. We now find

$$r=816/961=0.849. \tag{9}$$

Again, we have gone nearly halfway from the previous value toward unity.

Carrying this edge-bisecting procedure a step further becomes rather tedious; instead, we examine the PFES

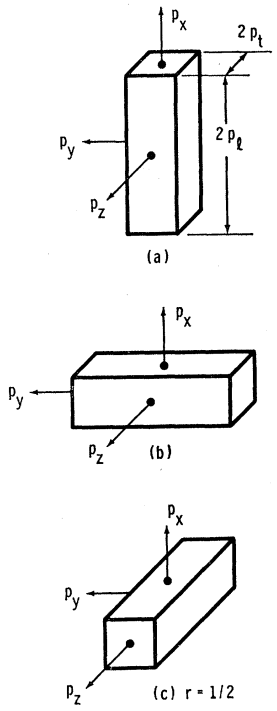

FIG. 5. A ⟨100⟩ multivalley model.

shown in Fig. 4(d). It has only one kind of face. The occupied volume consists of a central cube (of edge length $2p$) on each face of which a square-based pyramid of height sp has been erected. The figure corresponds to a case for which $s<1$. When $s>1$, the edges of the cube become concave (i.e., p type). The result for this model is

$$r=(1+s)(s^3-s^2+2s+1)/2(1+s^2)^2. \tag{10}$$

When $s=0$, the PFES becomes a cube and $r=\frac{1}{2}$, as it should. For $s=\frac{1}{2}$, $r=9/10$, its maximum value. This is rather close to unity for such a simple model. When $s=1$, the cube edges disappear, the PFES acquires the dodecahedral shape of the bcc Brillouin zone, and $r=\frac{3}{4}$. With further extension of the pyramids, r falls very slightly below $\frac{1}{2}$, and then approaches $\frac{1}{2}$ as $s\to\infty$.

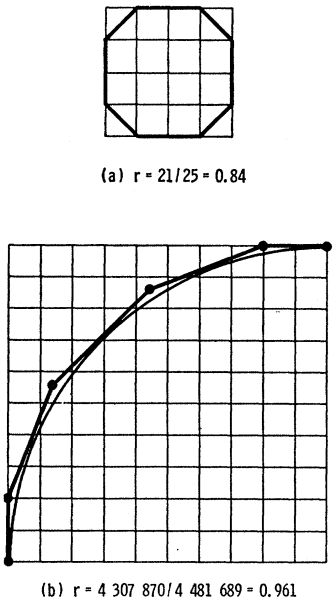

FIG. 6. Cross sections for two cylindrical models. Part (b) shows one quadrant of a 16-sided cross section.

B. Multivalley Models

Figure 5 presents another group of models which use the cube as a point of departure. In parts (a), (b), and (c) of the figure, the cube edges parallel to the p_x, p_y, and p_z axes have in turn been lengthened relative to the other two sets of edges. The longer and shorter edges are now labelled p_l and p_t. We let $K=p_l^2/p_t^2$; this reduces to the usual definition $K=m_l/m_t$ when the bands become parabolic.

According to the discussion following Eq. (2), we may use that equation to compute r only for the elongated surface in Fig. 5(c). The result is

$$r=\tfrac{1}{2}, \tag{11}$$

as for the cube. This is not surprising, since the longitudinal and transverse currents involve only v_t. The longitudinal dimension p_l cancels. The result remains $r=\frac{1}{2}$ when the surface is made endless by letting $K\to\infty$, or by allowing it to intercept opposite faces of a Brillouin zone. Ziman[14] has shown that the corresponding case of a cylinder with a circular cross section gives $r=1$.

We have investigated some intermediate cases as a further assessment of the potential accuracy of the PFES approach. For a cylinder with a *regular* octagonal cross section,

$$r=(\sqrt{2}+1)/2\sqrt{2}=0.854. \tag{12}$$

For the octagonal cross section shown in Fig. 6(a),

$$r=84/100. \tag{13}$$

The results for the two octagonal cross sections differ

[14] J. M. Ziman, Phil. Mag. 3, 1117 (1958).

only slightly, but the latter case was easier to derive because the vertices were located at simpler positions.

Figure 6(b) shows one quarter of a 16-sided approximation to a circular cross section. This case gives

$$r=4,\ 307,\ 870/4,\ 481,\ 689=0.961. \qquad (14)$$

The large numbers are not due to mathematical complexities, but are merely the result of finding the least common denominator of a sum of fractions.

This PFES approximation has come within 4% of the result for the circular cylinder. Actually, the latter case itself may be treated by a simple extension of the present method. This is shown in Appendix B.

We return to the main purpose of this section, and arrange the surfaces of Fig. 5 in a cubically symmetric, ⟨100⟩ multivalley model. The current components from the combination of surfaces lead to

$$r=\tfrac{1}{2}[3K(K+2)/(2K+1)^2], \qquad (15)$$

which differs from the result for a ⟨100⟩, ellipsoid-of-revolution multivalley model only by the presence of the factor $\frac{1}{2}$.

In Eq. (15), $f(K)\rightarrow\frac{3}{4}$ as $K\rightarrow\infty$. This result may be obtained directly from Fig. 5, keeping in mind that the ends of the three surfaces have vanished. For orientation (c) i_x and i_y are the same as for the cube, and hence by themselves would give $r=\frac{1}{2}$. For orientation (b), the contribution to i_x is the same as for (c), but there is no i_y. There is no i_x for orientation (a), and hence no i_y either. Since the new result will be expressed relative to the total carrier concentration $N=3n$, and since the combination i_y/i_x^2 appears in R_0, the effect is to multiply the result for (c) alone by 3 and divide it by 4.

C. Band-Filling Effects

We consider next the model shown in Fig. 7. The carriers lie within the volume formed by a cube surrounded by six arms with square cross sections. Pertinent "length" and "width" dimensions, p_b and p_a, are identified in the figure. The two situations shown in Fig. 7 correspond to $p_a\ll p_b$ and $(p_b-p_a)\ll p_b$. The result for this model is

$$r=(p_b-2p_a+2p_a^2/p_b)(3p_b-2p_a)/2(2p_b-2p_a+p_a^2/p_b)^2. \qquad (16)$$

When $p_a/p_b\rightarrow 0$, $r\rightarrow\frac{3}{8}$. This is the same limit that occurs for the multivalley model of the previous section as $K\rightarrow\infty$. This shows that when the arms become long and thin, it no longer matters whether their ends are connected or not. When $p_a/p_b=1$, the surface becomes a cube, and $r=\frac{1}{2}$ as it should. For some intermediate values, r falls slightly below $\frac{3}{8}$. We have not determined the precise value of the minimum, but it is approximately 0.31 and occurs near $p_a/p_b=\frac{1}{3}$.

The main purpose of this section, however, is to investigate band-filling effects, so we place the PFES of Fig. 7 in a simple cubic Brillouin zone of edge length $2p_b$, as suggested in part (a) of the figure. The outer ends of the arms vanish, and the occupied region of momentum space (in the repeated zone scheme) becomes a mutually orthogonal array of intersecting tubes with square cross sections. The two parts of Fig. 7 now correspond to a nearly empty and a nearly full band. The result is simpler, because now there is only one kind of face instead of two:

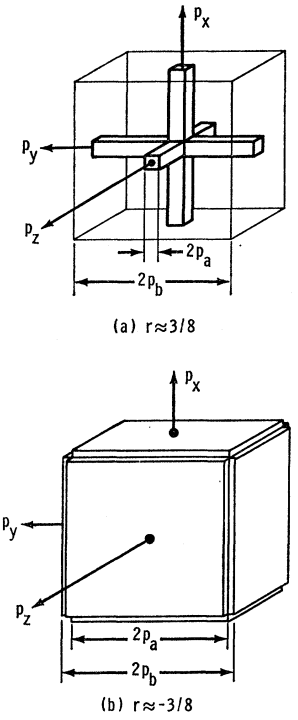

FIG. 7. A model to illustrate band-filling effects. The Fermi surface (arm thickness $=2p_a$) is placed in a simple cubic Brillouin zone (edge length $2p_b$). Parts (a) and (b) correspond to nearly-empty and to nearly-filled bands, respectively.

$$r=(p_b-2p_a)(3p_b-2p_a)/8(p_b-p_a)^2. \qquad (17)$$

The behavior of r is summarized in Fig. 8, plotted as a function of n/n_{fb}, where n_{fb} is the carrier density for the

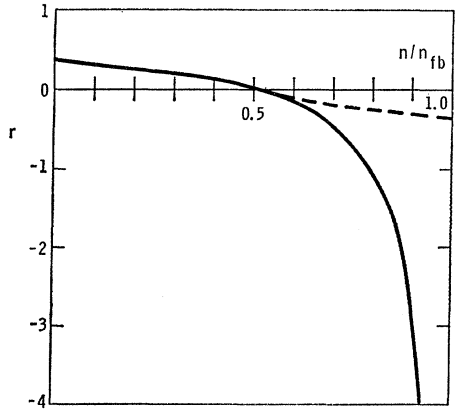

FIG. 8. The behavior of r for the model of Fig. 7, plotted as a function of carrier concentration n relative to the filled-band concentration n_{fb}. The dashed curve is r as computed from the equation $R_0=r/n_h e$, where n_h is the hole density.

completely filled band. As before, $r \to \frac{3}{8}$ as $n/n_{fb} \to 0$. For $n/n_{fb}=\frac{1}{2}$, $r=0$, and r becomes negative (the Hall coefficient becomes p type) for all higher carrier densities. If we recompute r relative to n_h (the density of missing carriers) the dashed curve results; it approaches $-\frac{3}{8}$ as $n_h \to 0$. In fact, the dashed curve for $n/n_{fb}>\frac{1}{2}$ and the solid curve for $n/n_{fb}<\frac{1}{2}$ are exact antisymmetrical reflections of each other. This occurs because the multiply connected arrays of occupied and empty states are identical for equal values of n and n_h.

We believe this is the first time that a simple expression for a Hall coefficient has been written down which (a) is valid for all energies between the bottom and top of a band, (b) changes from n to p type as the band fills, and (c) reduces to simpler limiting forms near the bottom and the top of the band. We mention once again that these results were obtained without specifying the dispersion relation.

One unrealistic feature of the above model is that both the bottom and top of the band consist of an array of lines, rather than a finite number of points. This defect is eliminated by a model consisting of a cubic Fermi surface in the dodecahedral bcc Brillouin zone, as shown in Fig. 9. In this case, both the bottom and top of the band collapse to a single point. The trouble with this model is that $r=\frac{1}{2}$ in the bottom half of the band and $r=-\frac{1}{2}$ in the top half. A discontinuous jump from the one value to the other occurs when $n/n_{fb}=\frac{1}{2}$; at this point, the entire lengths of all 12 edges of the cubic Fermi surface simultaneously contact the Brillouin zone, and suddenly the Fermi surface becomes a cube enclosing empty states rather than full ones. The solid and dashed lines in Fig. 9 suggest a filled cube, and part of an empty cube, respectively.

We consider one more band-filling model, an octahedral Fermi surface enclosed in a simple cubic Brillouin zone of edge length $2p$. Three stages in the band filling are shown in Fig. 10. For $n/n_{fb}<\frac{1}{6}$, the octahedron doe

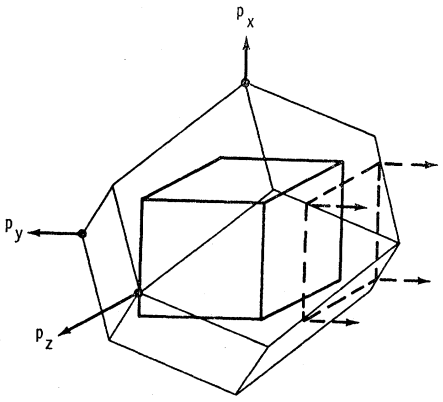

FIG. 9. Another band-filling model: a cubic Fermi surface in the bcc Brillouin zone. When the zone is more than half-filled, the Fermi surface becomes a cube enclosing empty states, as suggested by the dashed lines.

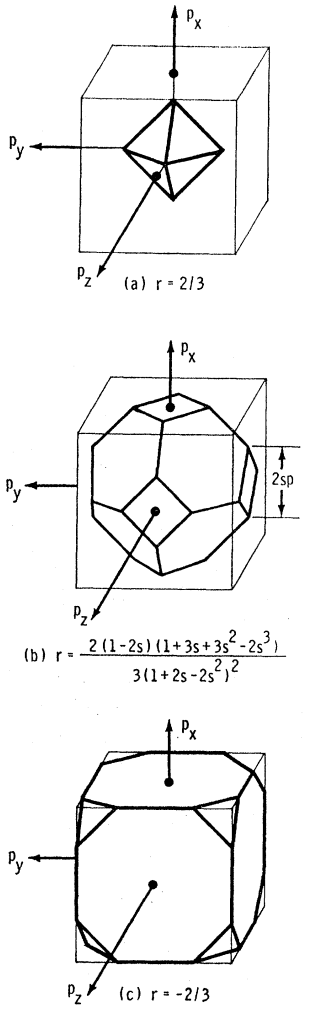

FIG. 10. Three stages in the filling of a simple cubic Brillouin zone by an octahedral Fermi surface.

not touch the zone faces. The result is $r=\frac{2}{3}$, as derived earlier.[1,13] When n/n_{fb} lies between $\frac{1}{6}$ and $\frac{5}{6}$, the octahedron is truncated by the faces of the Brillouin zone, and gradually changes its character with rising Fermi level. The result in this range is

$$r=2(1-2s)(1+3s+3s^2-2s^3)/3(1+2s-2s^2)^2, \quad (18)$$

where $2sp$ is the distance between opposite corners of a square face of the Fermi surface. The formula changes sign at $s=\frac{1}{2}$, when the band is half-filled. For $n/n_{fb}>\frac{5}{6}$, $r=-\frac{2}{3}$, if the result is expressed relative to n_h. In this final range, the Fermi surface has again become a simple octahedron, now enclosing empty states. The behavior of r is summarized in Fig. 11.

D. Anisotropic Scattering

To show how anisotropic scattering affects the PFES approach, we once again start with the cube as an example. At first, we assume that the scattering time has

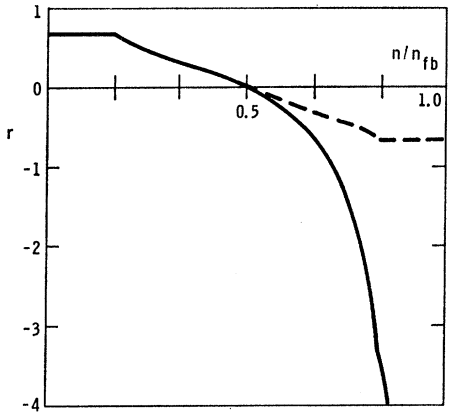

FIG. 11. Behavior of r for the model of Fig. 10. The dashed line is r relative to the hole density.

a constant value τ_e along the edges of a face, but varies in some fashion as we move from an edge toward the center of the face. Then the average scattering time $\bar{\tau}$ will generally differ from τ_e.

Parts (a) and (b) of Fig. 12 suggest how the cube might be displaced by the applied electric field when τ increases and decreases, respectively, in going from the edge to the center of a face. Equation (3) for i_x will now contain $\bar{\tau}$ instead of τ, but τ will be replaced by τ_e in Eq. (4) for i_y, since the latter formula refers only to carriers lying within an infinitesimal distance Δp_y of the edge. If we let $\bar{\tau}=f\tau_e$, Eq. (6) becomes

$$r=1/2f^2. \tag{19}$$

This result may be interpreted using the notion that every PFES consists of n- or p-type edges (i.e., convex or concave) and planar faces which are "neutral." The condition $f<1$ emphasizes the edges; this makes the surface more strongly n-type, and r increases. There are no p-type regions, so r cannot change sign, and in fact r can become zero only in the limit of infinite discrimination against the edges ($f=\infty$).

In the most general case, τ could vary along the cube edge, as well as between the edge and center of a face. We would then replace τ_e by $\bar{\tau}_e$; but we may still write $\bar{\tau}=f\bar{\tau}_e$, so the result retains the simple form of Eq. (19).

For a second example of τ anisotropy, we consider again the PFES shown in Fig. 4(c). It consists of six {100}, eight {111}, and 12 {110} faces. The area of an individual face is no more than about 5% of the total surface area. A simple way of introducing scattering time anisotropy into such a model is to assign a constant, but different, τ to each of the three kinds of nonequivalent faces. We do this and obtain

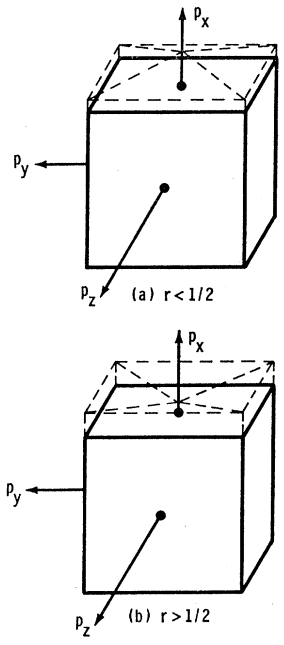

FIG. 12. The effect of scattering anisotropy on the displacement of a cubic Fermi surface. Parts (a) and (b) correspond to a τ which increases and decreases, respectively, in going from the edge to the center of a face.

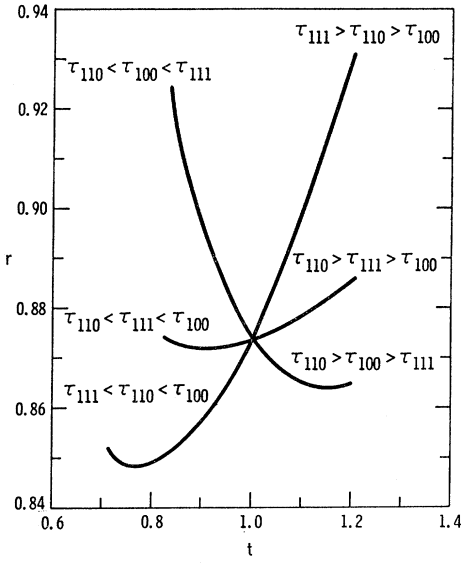

FIG. 13. The factor r for the model of Fig. 4(c) when each kind of face has a different scattering time. Plotted as a function of $t=\tau_a/\tau_b=\tau_b/\tau_c$ where a, b, c identifies the inequality sequence associated with each curve.

$$r=35(24+36t_{110}^2+12t_{111}^2)/3(12+16t_{110}+3t_{111})^2, \tag{20}$$

where $t_{110}=\tau_{110}/\tau_{100}$ and $t_{111}=\tau_{111}/\tau_{100}$. The behavior of Eq. (20) is plotted in Fig. 13 under the assumption that

$$t=\tau_a/\tau_b=\tau_b/\tau_c, \tag{21}$$

where a, b, and c identify a particular sequence of scattering times. Making the two ratios equal is a way of saying that there is a smooth variation of τ over the three types of faces. The curves in Fig. 13 correspond to the three possible ways of substituting the scattering times into Eq. (21).

Figure 13 shows that r is more sensitive to some kinds of scattering anisotropy than to others. It also shows

that all three kinds of anisotropy can cause r to drop below its value for isotropic scattering (i.e., its value due to shape anisotropy alone), but in each case, r rises again when the scattering anisotropy becomes too strong.

6. INTERPRETING THE BEHAVIOR OF r

A. Effect of Shape Anisotropy

For isotropic scattering, all of the PFES models considered in this paper make $r<1$. When a surface is distorted, the "n-typeness" is strengthened at some points and weakened at others. The latter effect outweighs the former, and r decreases.

The ultimate case of weakness is $r=0$, a completely neutral surface. This limit is attained in the infinitely oblate multivalley model of Fig. 5. In this limit, the edges (the only place where the n-typeness asserts itself) have receded to infinity, and their effect becomes negligible. Another case of $r=0$ occurred for the cube with τ-anisotropy in the limit of infinite discrimination against the edges (i.e., the edges vanish). Again, the surface has become completely neutral.

The factor r is not less than one for all models with a distorted Fermi surface and isotropic scattering. The Cooper and Raimes model[5] predicts $r<1$ and $r>1$ from shape anisotropy alone, and the calculation[6] based on Cohen's nonellipsoidal, nonparabolic model[7] leads to $r>1$ for certain ranges of the shape parameters. But the variety of shapes considered in this paper does suggest that the usual effect of shape anisotropy alone is to make $r<1$.

The values we have obtained suggest that r will most likely lie between $\frac{1}{2}$ and 1 for reasonably shaped Fermi surfaces (we exclude band-filling effects from the present discussion). Many actual materials, both solid and liquid, do have r values in this range. We did find $r<\frac{1}{2}$ in some cases, but these surfaces are usually very distorted, and are probably the least realistic models.

Sometimes the formula

$$n^*=1/Re \tag{22}$$

is used to compute a nominal, or effective, carrier density. This number seems rather misleading when $r\neq 1$ is due mainly to shape anisotropy, since n^* increases when a fixed carrier density takes on a less strongly n-type configuration.

When dealing with the evolution of a surface from n to p type, it seems more useful to regard the weak-field Hall coefficient as the product of a charge-density factor and a "strength" factor. The characterization of a Fermi surface as strongly n type, weakly n type, neutral, weakly p type, and strongly p type would then correspond to a continuous change in the value of r from approximately $+1$ to approximately -1 (assuming r is computed relative to the density of holes in the p-type range). The corresponding variation of n^* does not seem as meaningful.

B. Effect of Scattering Anisotropy

One of the earliest computations of r was made by Gans.[15] He assumed a spherical, parabolic band, classical statistics, and a constant mean-free path, and obtained $r=3\pi/8=1.18$. In his model, $r\neq 1$ solely because τ was a function of energy. The result for this kind of τ mixing may be written in the more general form

$$r=\langle\tau^2\rangle/\langle\tau\rangle^2, \tag{23}$$

where the angular brackets indicate a certain type of average over energy. It is well known that $r\geq 1$ in such a case.

We have not located any corresponding expression for the general case of a metallic model with a spherical Fermi surface and anisotropic scattering. Cooper and Raimes,[5] who used a particular kind of scattering anisotropy, did show that $r>1$ when the Fermi surface was spherical.

This behavior seems reasonable since the effect of a variation in τ is to focus attention on less than the total number of carriers present. Fewer carriers means a larger Hall coefficient, and hence r should increase. From this point of view, the use of an effective carrier density n^* seems much more appropriate for scattering-time anisotropy than for shape anisotropy.

C. Combined Effect of Shape and Scattering-Time Anisotropies

A distorted Fermi surface has some regions which are more, and others which are less, strongly n type. If we introduce the kind of scattering anisotropy which emphasizes the weaker n-type regions, then the over-all n-typeness will be weakened, and r should decrease below the value characteristic of the shape anisotropy alone. But if the scattering anisotropy becomes too strong, the model begins to behave as though fewer carriers are present, and r rises again. This is just the result found in all three curves of Fig. 13. However, if τ anisotropy should emphasize a p-type region on an otherwise n-type Fermi surface, r could ultimately become zero and change sign.

It is also possible to understand the differences among the three curves of Fig. 13, but first we must determine the nature of the curved Fermi surface to which the PFES of Fig. 4(c) is an approximation. We note that the relative distances of the {111}, {110}, and {100} faces from the origin are 1.16, 1.06, and 1.00 for this PFES. We conclude therefore that the true surface will be most sharply curved and most nearly flat near the ⟨111⟩ and ⟨100⟩ directions, respectively. We also note that the {110} distance is closer to the {100} than to the {111}, and that the area of the {110} faces is relatively large. Thus the curvature of the true surface near the ⟨110⟩ directions will be nearly as weak as that in the ⟨100⟩ directions.

[15] R. Gans, Ann. Physik **20**, 293 (1906).

Consequently, the most effective way of changing r would be to make τ_{111} the largest, and τ_{100} the smallest scattering time, or vice versa. It would be slightly less effective to interchange the τ's corresponding to the two weakest n-type regions. The least effective way of changing r would be to emphasize one of the weaker n-type regions while de-emphasizing the other, i.e., to assign an intermediate value to τ_{111}. Because the {110} regions are slightly more n type than the {100}, the net effect of emphasizing τ_{110} and de-emphasizing τ_{100} should be to increase r slightly. All of these predictions are confirmed by the curves of Fig. 13.

D. Are Mass and Scattering Anisotropies Equivalent Effects?

The value of r for an ellipsoidal multivalley model may be written in terms of the diagonal components of the effective-mass tensor. Herring and Vogt[16] pointed out that if both m^* and τ are diagonal tensors in the same coordinate system the only effect is to replace m_{ii} by m_{ii}/τ_{ii} in the expressions for r and other transport coefficients. Hence any given value of r can be obtained from mass or scattering anisotropy alone, and the two effects are completely equivalent.

Davis[17] and Cooper and Raimes[5] discussed the "one-way-only" equivalence of m^* and τ in their model. The results showed that a distorted Fermi surface with isotropic scattering can have $r>1$ and $r<1$, while $r>1$ for anisotropic scattering on a spherical energy surface. Thus it appeared that a suitably anisotropic shape could reproduce the effect of any given scattering anisotropy but not vice versa.

In contrast to these earlier results, we find that shape distortion alone usually makes $r<1$ while scattering anisotropy makes $r>1$. Thus for most of the cases we considered, the two effects do not appear to be equivalent.

The ellipsoidal multivalley model is a very special model. As we noted in the introduction, there is no mixing of Hall angles on a single ellipsoidal surface. This means that the response of all carriers in a valley to the applied forces may be expressed in terms of a mobility tensor. Clearly, this remains true if both m^* and τ are diagonal tensors in the same coordinate system, and τ and m^* must therefore be equivalent.

A still simpler case of equivalence is the isotropic metallic two-band model. Here the response of the carriers in each band may be described by a scalar mobility, and a given result depends only on the quotient τ/m in each band.

The Davis-Cooper-Raimes model was restricted to *slightly* anisotropic surfaces and scattering, and the one-way equivalence is meaningful only for $r\approx 1$. In all probability, r in their model would drop below unity when the Fermi surface became more strongly distorted.

Because of the special nature of these earlier models, the conclusions regarding the equivalence of τ and m^* anisotropies may not be extended to other models, and we believe that our conclusion that the two kinds of anisotropy usually produce opposite, nonequivalent effects is of more general validity.

E. Why Does a Hall Coefficient Change Sign?

The textbook explanation for a p-type Hall coefficient is usually accompanied by a figure showing an s-shaped energy-momentum curve. The p-type region is identified with the negative second-derivative portion of the curve, i.e., with a negative effective mass. By contrast, we derived expressions for r in Sec. 5 D which changed sign as the band was filled, and the results were independent of the energy-momentum relation.

To clarify the distinction between these two approaches, we discuss the simple case in which the same dispersion relation applies in all directions. The constant energy surfaces will then be spheres. The spheres will enclose filled states or empty states according to whether the family of dispersion curves has a common point at their bottom or top ends. Then r will be precisely $+1$ or -1 for *all* Fermi levels within the band.

The shape of the one-dimensional energy-momentum curve specifies the derivative $\partial^2\mathcal{E}/\partial p_i^2$; the Hall coefficient, however, is determined by the curvature of the Fermi surface, i.e., by the values of the cross derivatives $\partial^2\mathcal{E}/\partial p_i\partial p_j$.

Of course constant-energy surfaces in real bands collapse to a point, or to a finite set of points, at the top of the band, and their energy-momentum curves do acquire negative second derivatives near the band maxima. Unfortunately, the textbook explanations for the p-type Hall coefficient behavior emphasize this second characteristic, and this can be misleading, since only the first characteristic is essential to the phenomenon.

7. SOME APPLICATIONS

The models worked out in this paper were chosen because they illustrated the potentialities of the PFES approach, and because they answered questions concerning the general behavior of r. We would like to present a few brief examples showing how these ideas about r can be put to use.

A. "Lead-Pipe" Models

The free-electron model for the third zone in lead consists of an array of interconnected tubes lying along all the edges of the fcc Brillouin zone.[18] The empty states in the heavy-mass d-band of Pd also form a cubically symmetric array of connected tubes.[19]

We suggested in Sec. 5 D that the value of r in such an array will be nearly the same as that found in a strongly prolate multivalley model. Consequently, r should be

[16] C. Herring and E. Vogt, Phys. Rev. **101**, 944 (1956).
[17] L. Davis, Phys. Rev. **56**, 93 (1939).
[18] J. R. Anderson and A. V. Gold, Phys. Rev. **139**, A1459 (1965).
[19] J. J. Vuilleman, Phys. Rev. **144**, 396 (1966).

about $\frac{3}{4}$ of the cross-section factor of the tubes, and the latter is more likely to be nearer to unity (circular) than to $\frac{1}{2}$ (square cross section).

We would expect therefore that r for lead-pipe models would often have values in the range between $\frac{1}{2}$ and $\frac{3}{4}$, so long as the tubes do not become too short and thick, and so long as scattering anisotropy does not emphasize the p-like connecting regions too strongly.

B. Application of the Cooper-Raimes Model to the Noble Metals

Cooper and Raimes[5] made the Fermi surface and the scattering time anisotropic by incorporating fourth-order, cubically symmetric surface harmonics into the functions describing the surface and the scattering. Their most general result included the effects of alloying; for the pure metal, it became

$$r=\{1+(4/21)[9A^2-18A(C-B)-(C-B)^2]\}^{-1}, \quad (24)$$

to second order in A, B, and C. The parameters A and B specify the shape, and C the scattering-time anisotropies (zero values correspond to spherical energy surfaces and isotropic scattering).

Cooper and Raimes wanted to apply their model to silver, so they chose negative values for A and B, which corresponded to the kind of $\langle 111\rangle$ bulges that were thought then to be present in silver. This by itself reduces r to 0.92. A positive value for the scattering parameter C was required to bring r down to the experimental value of 0.81, and this makes the scattering time longest in the $\langle 100\rangle$, and shortest in the $\langle 111\rangle$ directions, opposite to present-day views on the scattering anisotropies in the noble metals.[20]

We see that Cooper and Raimes needed a scattering anisotropy which emphasized the least n-type part of their Fermi surface, so that r would be less than the value associated with shape alone. In a general sense, their Fermi surface resembles the actual Fermi surface of silver, but the bulges in the model do not have the necks of the real surface. Hence the model surface is most *strongly n* type in the $\langle 111\rangle$ directions, while the partially concave necks in the real surface make those directions the most *weakly n* type.

It is this interchange of the locations of the strongest and weakest n-type regions of the Fermi surface which causes the Cooper and Raimes model to predict the wrong kind of scattering anisotropy.

C. Further Comment on r in the Noble Metals

The Fermi surface now known to be appropriate for the noble metals, combined with scattering anisotropy which emphasizes the neck regions, actually is most useful for explaining the very low values of r (as low as $\frac{1}{3}$) which occur in concentrated alloys of the noble metals at low temperatures. But do we have to invoke scattering anisotropy to explain the room-temperature values of r in the pure noble metals,[21] which lie between 0.66 and 0.80? Ziman's computed values of r, using his eight-cone model with isotropic scattering,[21] lie between 0.49 and 0.63. Hence his model overemphasizes the importance of the necks.

The cross section of the necks constitutes only a few percent of the total area of the Fermi surface. Thermal scattering is not expected to be strongly anisotropic in the noble metals,[22] and therefore the partially p-type nature of the neck regions can only have a very minor effect on the value of r. We believe that a more important effect of the necks is to distort the main body of the Fermi surface in a way that makes it resemble the dodecahedral bcc Brillouin zone. This can be seen rather clearly from Figs. 7 and 8 of Burdick's band calculation on copper.[23] The dodecahedron was treated in Sec. 5 A; it is a particular case ($s=\frac{1}{2}$) of Eq. (10) and Fig. 4(d). The result was $r=\frac{3}{4}$, and hence it appears that we can obtain semiquantitative agreement with room-temperature experimental values in the pure noble metals without taking into account the nature of the necks themselves, and without having to introduce scattering anisotropy.

8. CONCLUSIONS

We have described an approximate method for computing the mixing factor r in the weak-field Hall coefficient, and presented results for a wide variety of models.

We discussed the procedure in detail only for the simple case of a cube, but none of the results obtained required a long or complicated computation. Only a few terms had to be evaluated, no complicated functions appeared, and no integration over the constant-energy surface was required. These simplifications were possible because we used planar-faced constant-energy surfaces with only a few kinds of nonequivalent faces.

Another advantage of the PFES approach is that anisotropic scattering can be incorporated into any given model without introducing any complications. New factors may appear in the result for r, but there are no new terms to compute.

Using a PFES makes it possible to demonstrate band-filling effects in a simple, direct way. This occurs because a PFES, placed in a Brillouin zone, can "turn itself inside out" without losing its mathematically simple representation.

We want to make it clear that we have not developed a new theory for solving transport problems. The method is just a pictorial approach to the solving of the conventional Boltzmann equation. But the approach puts the emphasis where it belongs—on geometry—and makes it practical to obtain answers for complicated Fermi surfaces.

Having done this for a variety of models, it became possible to explain why r behaves the way it does. We

[20] V. Heine, Phil. Mag. **12**, 53 (1965).

[21] J. M. Ziman, Advan. Phys. **10**, 1 (1961).

[22] J. M. Ziman, Phys. Rev. **121**, 1320 (1961).

[23] G. A. Burdick, Phys. Rev. **129**, 138 (1963).

were able to discuss how shape and scattering-time anisotropies affect r, separately and together, and we were able to apply these ideas to obtain a clearer understanding of previously published results.

We have discussed in some detail how the mixing factor r is affected by shape and scattering-time anisotropies. If information about one of these characteristics is available, the weak-field Hall coefficient might be used to deduce the nature of the other. But efforts to do this have rarely succeeded because of the mathematical complexities which have always accompanied the introduction of nonsimple Fermi surfaces.

There have been discussions about the appropriate procedure for computing transport in alloys and in liquids. But how can one argue about the point beyond which conventional transport theories must be replaced by alternative approaches, when the potentialities of conventional theories have never been evaluated properly because of mathematical complexities?

The method described in this paper gets around this mathematical roadblock, at least as far as the weak-field Hall coefficient is concerned. We believe that therein lies its greatest potential usefulness.

Note added in proof. Professor R. G. Chambers has pointed out that the change in i_x is *linear* in H for a PFES. Miyazawa[13] has also noted the "unphysical" nature of higher-order corrections to the weak-field Hall coefficient.

ACKNOWLEDGMENTS

The simple method for treating the cube, upon which the present generalization was based, was worked out at the Cavendish Laboratory of Cambridge University. My stay there during the 1965–66 academic year was made possible by the hospitality of Professor Sir Nevill Mott and by the financial support of the U. S. Naval Ordnance Laboratory. I am indebted to T. Reglein who checked the computations for me, and found what I hope is the only error. I also want to express my gratitude to Stanley Rogers and Betty Jew for the excellent illustrations which they prepared, and which are a very important part of this paper.

APPENDIX A: EXTENSION TO CLASSICAL STATISTICS

We treat the cube as an example. The volume element dV at the top of a cube at energy $\mathcal{E}$, no longer the Fermi energy, is composed of the area $A=4p^2$ and length dp_x, which is not related to $\Delta p_x=eE\tau$.

The contribution to i_x from the volume element is $(\Delta n)ev_x(\mathcal{E})$. We have

$$\Delta n/(2/h^3)=\Delta f dV=\left(-\frac{\partial f_0}{\partial\mathcal{E}}\frac{\partial\mathcal{E}}{\partial p_x}\frac{\partial p_x}{\partial t}\tau\right)dV$$

$$=\left(-\frac{\partial f_0}{\partial\mathcal{E}}\frac{\partial\mathcal{E}}{\partial p_x}eE\tau\right)dV. \quad \text{(A1)}$$

We must specify the dispersion relation, so we assume parabolic bands: $\mathcal{E}=p_i^2/2m$ normal to each face of the cube. Then

$$i_x=\int_0^\infty\left(\frac{4}{h^3}\right)(4p^2)\left(-\frac{\partial f_0}{\partial\mathcal{E}}\frac{p_x}{m}eE\tau\right)(e)\left(\frac{p_x}{m}\right)dp_x.$$

The edge crossing volume at energy $\mathcal{E}$ is $dV=dAdp_x=(2p)\Delta p_y dp_x$, where Δp_y is the same as in the metallic case. Therefore

$$i_y=\int_0^\infty\left(\frac{4}{h^3}\right)(2p)\left(e\frac{p_x}{m}H\tau\right)\times\left(-\frac{\partial f_0}{\partial\mathcal{E}}\frac{p_x}{m}eE\tau\right)(e)\left(\frac{p_y}{m}\right)dp_x. \quad \text{(A2)}$$

The relation which introduces the total carrier concentration becomes

$$n=\int_0^\infty\left(\frac{2}{h^3}\right)(24p^2)f_0dp. \quad \text{(A3)}$$

For classical statistics, $f_0=ae^{-\mathcal{E}/kT}$. We work out the familiar case $\tau=b\mathcal{E}^{-1/2}$. Carrying out the integration, using $p_x=p_y$, and substituting the results into Eq. (2) for the Hall coefficient gives

$$R_0=\tfrac{1}{2}(3\pi/8)(1/ne). \quad \text{(A4)}$$

In other words, the factor r is simply the classical τ-mixing factor $3\pi/8$ times the shape factor $\frac{1}{2}$.

APPENDIX B: EXTENSION TO CURVED FERMI SURFACES

We treat the case of a right circular cylinder in orientation (*c*) (i.e., with E and H perpendicular to and parallel to the cylinder axis, respectively).

Let the cylindrical Fermi surface have a radius p_t and length p_l. Consider the element of surface area $dA=p_t d\theta dp_l$ lying at an angle θ with respect to the p_x direction. The projected area $dA_x=dA\cos\theta$, and $v_x=v\cos\theta$. Hence

$$i_x=\int_{-\pi/2}^{\pi/2}\left(\frac{4}{h^3}\right)(p_t p_l eE\tau)(ev)\cos^2\theta d\theta. \quad \text{(B1)}$$

The change in v_y at θ is $v\cos\theta\Delta\theta$, where $\Delta\theta=evH\tau/p_t$. Thus

$$i_y=\int_{-\pi/2}^{\pi/2}\left(\frac{4}{h^3}\right)(p_l eE\tau)(evH\tau)(ev)\cos^2\theta d\theta. \quad \text{(B2)}$$

The total carrier concentration is given by

$$n=(2/h^3)(\pi p_t^2 p_l). \quad \text{(B3)}$$

Combining the equations in the usual fashion leads to

$$r=1, \quad \text{(B4)}$$

agreeing with Ziman's calculation.[14]