Geometric interpretation of the weak-field Hall conductivity in two-dimensional metals with arbitrary Fermi surface

N. P. Ong
Joseph Henry Laboratories of Physics, Princeton University, Princeton, New Jersey 08544
(Received 2 July 1990)

The Hall conductivity σ_{xy} of a two-dimensional metal in the weak-field, semiclassical, limit has a simple geometric representation. σ_{xy} (normalized to e^2/h, where e is the electron charge and h is Planck's constant), is equal to twice the number of flux quanta ϕ_0 threading the area A_l, where A_l is the total "Stokes" area swept out by the scattering path length $l(\mathbf{k})$ as $\mathbf{k}$ circumscribes the Fermi surface (FS). From this perspective, many properties of σ_{xy} become self-evident. The representation provides a powerful way to disentangle the distinct contributions of the three factors, FS area-to-circumference ratio, anisotropy in $l_{\mathbf{k}}$, and negative FS curvature. The analysis is applied to the Hall data on $2H$-$NbSe_2$ and the cuprate perovskites. Previous model calculations of σ_{xy} are critically reexamined using the new representation.

I. INTRODUCTION

In the weak-field semiclassical limit, the Hall conductivity[1,2] σ_{xy} of a metal is very sensitive to the local curvature[3] of the Fermi surface (FS). Since FS shapes may be quite complex, it is usually a difficult task to compare the observed σ_{xy} with predictions of band structure. In two-dimensional (2D) systems, however, there exists an appealing geometric representation of the weak-field Hall conductivity that brings out explicitly its precise relation to FS curvature. The representation clarifies the way in which the local curvature and anisotropy in the scattering length interact to determine the Hall current. It also provides insight into the many attempts to systematize model calculations of σ_{xy} using highly idealized FS models.[4-8]

The geometric interpretation is as follows. Consider moving a point $\mathbf{k}$ around the FS (Fig. 1). The "scattering path length" vector, defined as $l(\mathbf{k})=\mathbf{v}_{\mathbf{k}}\tau_{\mathbf{k}}$, sweeps out, in the plane l_x-l_y (the "l space"), a closed curve that is usually more complicated than the FS curve ($\hbar\mathbf{v}_{\mathbf{k}}=\partial\varepsilon_{\mathbf{k}}/\partial\mathbf{k}$ and $\tau_{\mathbf{k}}$ is the relaxation time). For instance, the l curve may self-intersect at several points. Regardless of the FS shape, however, the weak-field σ_{xy} (normalized to e^2/h) is equal to twice the number of flux quanta ϕ_0 threading the l curve (e is the electron charge and h Planck's constant).

In Sec. II the Jones-Zener solution of the Boltzmann equation is recast as a map from $\mathbf{k}$ space to l space, to derive the representation. The nature of the l curve is discussed in Sec. III with emphasis on self-intersecting segments. The geometric consequences of this representation are described in Secs. IV and V, and the variation of ρ_{xy} in Secs. VI. In Sec. VII, applications to $2H$-$NbSe_2$, and the cuprate perovskites are discussed. In the former, a surprisingly good agreement between the band structure and the high-temperature Hall coefficient is obtained. In the latter, the conventional Boltzmann model fails. The analysis provides a clear statement of what is anomalous about the temperature dependence of R_H. Previous model calculations[4-8] are compared in Sec. VIII. I conclude with a discussion of the physics underlying the geometric representation in Sec. IX.

II. GEOMETRIC REPRESENTATION OF σ_{xy}

I assume that the Boltzmann approach is valid, and anomalous contributions to the Hall scattering (such as magnetic skew scattering[9]) and magnetic breakdown effects[10] are absent. From the Jones-Zener[11] solution to the Boltzmann equation, the weak-field Hall conductivity is

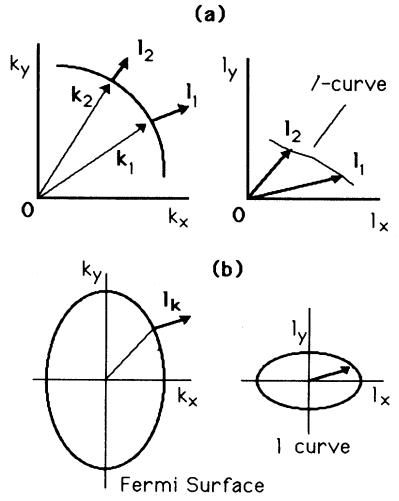

FIG. 1. (a) Mapping from the FS in $\mathbf{k}$ space to the l curve in the space of $l(\mathbf{k})$. (b) Map from an elliptical FS to its l curve, assuming isotropic τ.

$$\sigma_{xy}=2(e^3/\hbar)B\sum_{\mathbf{k}}\left[\frac{-\partial f_{\mathbf{k}}}{\partial\varepsilon}\right](v_y\tau_{\mathbf{k}})\left[v_y\left[\frac{\partial}{\partial k_x}\right]-v_x\left[\frac{\partial}{\partial k_y}\right]\right](v_x\tau_{\mathbf{k}})\ , \tag{1}$$

where $f_{\mathbf{k}}$ is the Fermi-Dirac distribution and ε the electronic energy. The magnetic field $\mathbf{B}$ is taken $\|-\hat{\mathbf{z}}$, and the electric field $\mathbf{E}\|\hat{\mathbf{x}}$. For my purpose it is instructive to derive Eq. (1) as follows. The Hall current J_H arises from the second-order displacement of the FS by $\delta\mathbf{k}_B=(e\tau_{\mathbf{k}}/\hbar)\mathbf{v}\times\mathbf{B}$, i.e., $J_H=2e\sum_{\mathbf{k}}v_y(\delta\mathbf{k}_B\cdot\nabla)\delta f_E$. Here, δf_E, the first-order displacement due to $\mathbf{E}$, is given by $(e\tau_{\mathbf{k}}/\hbar)(\mathbf{E}\cdot\nabla)f_{\mathbf{k}}$. Thus, $J_H=2E(e^3/\hbar)B\sum_{\mathbf{k}}(-\partial f_{\mathbf{k}}/\partial\varepsilon)(v_y\tau_{\mathbf{k}})(\mathbf{v}\times\mathbf{B})\cdot\nabla(v_x\tau_{\mathbf{k}})$, which is equivalent to Eq. (1). For a 2D system (at temperatures $T\ll$ Fermi energy ε_F), Eq. (1) simplifies to

$$\sigma_{xy}=(e^3/2\pi^2\hbar)\int dk_t|\mathbf{v}|^{-1}[v_y\tau_{\mathbf{k}}(\mathbf{v}\times\mathbf{B})\cdot\nabla(v_x\tau_{\mathbf{k}})]\ , \tag{2}$$

where $|\mathbf{v}|^{-1}$ is the density of states factor, k_t the component of $\mathbf{k}$ along $\hat{\mathbf{t}}$ (the unit vector tangential to the FS curve), and the integral, taken around the FS curve, is evaluated at ε_F. Since $\mathbf{v}\times\mathbf{B}/|\mathbf{v}|=B\hat{\mathbf{t}}$ in a 2D system, the integral reduces to $B\int dk_t(\hat{\mathbf{t}}\cdot\nabla l_x)l_y=BA_l$, with A_l defined by

$$A_l=(\mathbf{B}/B)\cdot\int d\mathbf{l}\times\mathbf{l}/2\ . \tag{3}$$

A_l is the area swept out by the vector $\mathbf{l}(\mathbf{k})$ as $\mathbf{k}$ moves around the FS. Because its sign is determined by the "circulation" of $\mathbf{l}(\mathbf{k})$, I will call it the "Stokes" area. In terms of A_l, Eq. (2) reduces to

$$\sigma_{xy}=(e^2/h)A_l/(\pi l_B^2) \tag{4a}$$

$$=(e^2/h)2(\phi/\phi_0)\ , \tag{4b}$$

where $l_B=\sqrt{\hbar/eB}$ is the magnetic length, $\phi=BA_l$ the magnetic flux threading the l curve, and $\phi_0=h/e$ the flux quantum.

The transformation from an integral in $\mathbf{k}$ space [Eq. (1)] to one in $\mathbf{l}$ space [Eq. (3)] is valid for arbitrary $\mathbf{l}(\mathbf{k})$ and an arbitrary 2D FS. The geometric representation immediately implies that (regardless of the symmetry of the 2D crystal) σ_{xy} is a scalar quantity independent of rotation of the x-y axes relative to the crystal's principal axes in the x-y plane. (The shape of the l curve cannot depend on the axis orientation)

Equations (3) and (4) are related to Tsuji's well-known expression[3] for σ_{xy} in terms of FS curvature. However, Tsuji's expression (an integral in $\mathbf{k}$ space) makes no reference to the curve in $\mathbf{l}$ space. The geometric relation between σ_{xy} and the flux threading A_l (especially the Stokesian nature of the l curve), and the many implications for real metals, seem to have gone unnoticed.[1-4]

The three-dimensional (3D) generalization of Eq. (4a) is

$$\sigma_{xy}^{3D}=(e^2/h)\int(dk_z/2\pi)A_l(k_z)/(\pi l_B^2)\ , \tag{5}$$

where the integral is over slices of the FS, of thickness dk_z, parallel to the x-y plane, and A_l is as defined in Eq. (3), with $\mathbf{l}$ replaced by its projection $\mathbf{l}_\perp=(\mathbf{l}-\mathbf{l}\cdot\hat{\mathbf{z}})$.

III. SECONDARY LOOPS IN THE l CURVE

As the point $\mathbf{k}$ moves around the FS curve, the vector $\mathbf{l}(\mathbf{k})$ traces out its own curve. Whether the two vectors $\mathbf{k}$ and $\mathbf{l}$ rotate in the same, or opposite, sense is determined by the local curvature κ of the FS. [$\kappa=d\theta/ds$, where θ is the angle $\mathbf{l}(\mathbf{k})$ makes with $\mathbf{x}$, and s the arc length along the FS. Hereafter, I rewrite dk_t as ds.] A positive κ implies that $\mathbf{k}$ and $\mathbf{l}(\mathbf{k})$ rotate in the same sense. I shall restrict attention to FS curves in which κ is finite or zero everywhere, i.e., cusps are excluded. (Obviously, the FS curve cannot self-intersect.) If the FS is a convex closed curve (κ non-negative everywhere), the l curve is closed and simple (does not self-intersect). For example, for an elliptical FS with effective mass m_x and m_y along $\mathbf{x}$ and $\mathbf{y}$, respectively [Fig. 1(b)] the l curve is also an ellipse with semiradii $\hbar k_{Fx}\tau/m_x$ and $\hbar k_{Fy}\tau/m_y$ (τ is assumed isotropic). Using $A_l=\pi\hbar^2\tau^2(k_{Fx}k_{Fy})/m_xm_y$ in Eq. (4a), I obtain the correct σ_{xy} in the weak-field limit.

The more interesting cases arise when the FS is nonconvex and $l_{\mathbf{k}}[=|\mathbf{l}(k)|]$ is $\mathbf{k}$ dependent. The l curve is then a nonsimple closed curve with the same symmetry as the FS. Suppose we move $\mathbf{k}$ anticlockwise along a segment PQ in which κ first changes from positive to negative, and then back to positive (Fig. 2). In the l_x-l_y plane, $\mathbf{l}(\mathbf{k})$ reverses its sense of rotation each time κ changes sign. Hence, if $l_{\mathbf{k}}$ takes on different values on the negative and positive κ segments, the l curve must intersect itself, describing a closed loop that does not encircle the origin. (To distinguish the closed loops from the single loop that encircles the origin, I refer to them as secondary and primary, respectively.) The circulation of the secondary loop is the same as (opposite to) that of the primary loop when $l_{\mathbf{k}}$ is smaller (larger) on the negative-κ segment, compared with its value on the neighboring segments (Fig. 2).

The elemental area $\delta A_l=\hat{\mathbf{z}}\cdot\mathbf{l}\times\delta\mathbf{l}/2$ shares the same sign as the local FS curvature κ. Hence, as $\mathbf{l}(\mathbf{k})$ sweeps out the secondary loop in the example above, the contri-

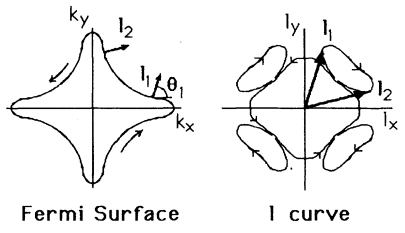

FIG. 2. The l curve generated by a fourfold symmetric FS . The negative-κ segments on the FS along [110] give rise to negative secondary loops in the l curve. l_1 and l_2 are the scattering path-length vectors at the points where κ vanishes.

bution of segment PQ to the total area A_l includes the area A_s enclosed by the secondary loop, so the total area is

$$A_l = A_p(\text{primary}) + \sum_m A_s(\text{secondary loop } m) \ . \tag{6}$$

The sign of A_s is the sign of each secondary loop as determined by its circulation. To summarize, each segment of the FS with negative κ generates a secondary loop that is positive (negative) if $l_{\mathbf{k}}$ on that segment is shorter (longer) than $l_{\mathbf{k}}$ on its neighbors. The area of the secondary loop adds algebraically to that of the primary loop. The sensitivity of A_l to the sense of circulation is conveniently expressed in the flux language [Eq. (4b)], since $A_l B$ is precisely the *net* magnetic flux threading the self-intersecting l curve. The geometric representation shows that σ_{xy}, rather than being a simple sum over independent segments, actually describes a global property of the FS curve (and the $\mathbf{k}$ dependence of $l_{\mathbf{k}}$). σ_{xy} cannot be computed by decomposing the 2D FS into a series of arcs and summing the contribution from each arc. This procedure introduces large errors at the points where arcs join (see Sec. VIII). Each segment's contribution is strongly influenced by the behavior of κ and $l_{\mathbf{k}}$ on its neighbors.

IV. ARBITRARY FS WITH ISOTROPIC l

Some general results are easily derived. First, I consider 2D closed FS curves, in which $l_{\mathbf{k}}$ is a constant l_0 everywhere (the "isotropic-l" assumption). The area of all secondary loops collapses to zero, and the primary loop is a circle of radius l_0. Thus, in the isotropic-l approximation, *all* closed FS shapes have the same value for σ_{xy}, aside from the scaling factor l_0^2. [In this case the Hall coefficient reduces to the isoperimetric ratio (Sec. VI).] This simple result contradicts model calculations which purport to show that (within the isotropic-l assumption) σ_{xy} may become electronlike even though the FS cross section has global holelike topology because large segments with electronlike curvature may dominate smaller segments with holelike curvature (see Sec. VIII).

As a second example, I consider an open FS in 2D (Fig. 3). The vector $\boldsymbol{l}(\mathbf{k})$ assumes the same orientation at two Bragg planes separated by a reciprocal lattice vector. As we move along the FS between the planes, $\boldsymbol{l}(\mathbf{k})$ traces out a closed loop which does not include the origin. If $l_{\mathbf{k}}$ is a constant, the area A_l vanishes. Hence, σ_{xy} of all open surfaces is zero, unless $l_{\mathbf{k}}$ is anisotropic, in which case the sign of σ_{xy} is determined by the segments with the largest $l_{\mathbf{k}}$.

V. ARBITRARY CLOSED FS WITH k-DEPENDENT l

In general, when $l_{\mathbf{k}}$ varies with $\mathbf{k}$, the l curve will be far from circular and, usually, nonsimple. As discussed above, secondary loops are generated when κ is negative over finite FS segments. For concreteness, I consider the two closed Fermi surfaces in Fig. 4. both have fourfold

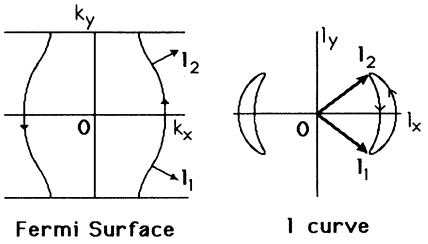

FIG. 3. The l curve (the two crescents) generated by an open FS with anisotropic, $l_{\mathbf{k}}$. If $l_{\mathbf{k}}$ is isotropic, σ_{xy} vanishes.

symmetry in the plane and have segments with very large curvature sandwiched between segments of small curvature (the "fingers" along [100] in Fig. 4(a), and the "fjords" along [110] in Fig. 4(b)). If I assume that $l_{\mathbf{k}}$ is larger on the low-κ segments than on the high-κ segments on both FS (as drawn), their respective l curves are shown in Figs. 4(c) and 4(d), respectively. In Fig. 4(c), the "flatter" segments along [110] generate the secondary loops. Since $l_{\mathbf{k}}$ is longer there, the loops are negative. In Fig. 4(d), the large, negative-κ segments on the fjords generate secondary loops that are positive because $l_{\mathbf{k}}$ is shorter there. [If the opposite assumption on $l_{\mathbf{k}}$ is adopted (i.e., $l_{\mathbf{k}}$ is longer on the high-curvature segments), the signs of the secondary loops will reverse in both FS.]

These examples illustrate the crucial role the anisotropy in $l_{\mathbf{k}}$ plays in determining the sign of the contribution from a particular segment. In Fig. 4(d), the secondary loops are positive despite the existence of negative-κ segments, so σ_{xy} is always positive, regardless of the magnitude of κ on the negative-κ segments. For a particular segment to generate a negative contribution to σ_{xy}, $l_{\mathbf{k}}$ must be longer on that segment than on its neighbors (in

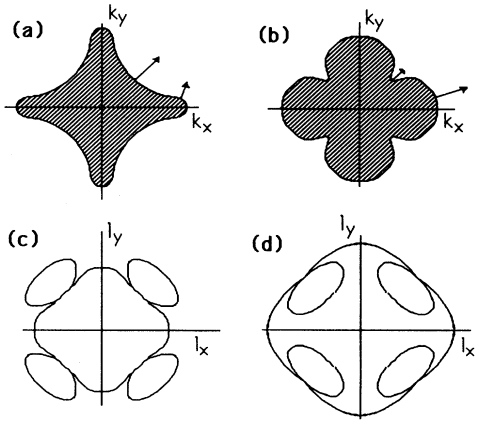

FIG. 4. Two nonconvex closed FS and their corresponding l curves. In panel *a* (b), the arc length of each negative-κ segment is long (short). If $l_{\mathbf{k}}$ varies as drawn (short on the large-curvature segments), the l curves qualitatively resemble the closed curves in panel *c* (d). However, if $l_{\mathbf{k}}$ has the opposite variation, the l curves in *c* and *d* are interchanged.

addition to κ being negative).

The global nature A_l is nicely illustrated by the clover leaf FS in Fig. 4(b). Let us consider the limit in which *both* the arc length and $l_{\mathbf{k}}$ of the high-curvature segments (the fjords) become very short. (Thus, the conductivity σ_{xx} is dominated by the positive-κ segments.) It might be inferred that the contribution of the fjords to σ_{xy} is also negligible. However, examination of Fig. 4(d) shows this to be incorrect. In this limit, the secondary loops increase in size until they are collectively comparable in area to the primary loop, i.e., their contribution to A_l becomes comparable to that of the primary (and of the same sign). In general, small segments of high curvature strongly influence σ_{xy}; their contribution is readily assessed by examining $\mathbf{A}_l$.

For a large class of FS curves that possess negative-κ segments, the secondary loops are negative. This is true if $l_{\mathbf{k}}$ decreases with increasing $k=|\mathbf{k}|$ everywhere [as in Fig. 4(c)]. (A segment that has negative κ is, on average, closer to the origin than its positive-κ neighbors. Thus, the condition $dl_{\mathbf{k}}/dk<0$ implies that $l_{\mathbf{k}}$ assumes a smaller value on this segment than on adjacent segments.) A common example occurs when the energy contours near the Fermi level are similar in shape except for a uniform scale factor. Under uniform dilation, segments with large, positive κ [the "fingers" in Fig. 4(a)] have contours that are more widely separated compared with the small-κ segments. Conversely, segments with large, *negative* κ have closely separated contours [the "fjords" in Fig. 4(b)]. Thus, the velocity $v_{\mathbf{k}}$ is small in segments with large, positive κ (and large in segments with large, negative κ.) I will call such behavior of $v_{\mathbf{k}}$ "conventional." If I assume that τ is isotropic, then $l_{\mathbf{k}}$ is proportional to $v_{\mathbf{k}}$, so that $dl_{\mathbf{k}}/dk<0$ is satisfied for both cases. Thus, FS shapes which are scale invariant always generate negative loops, provided τ is isotropic.

In general, however, the anisotropy of $l_{\mathbf{k}}$ is controlled by the physics of electron scattering, rather than by geometry.[12] At high $T(\geq\Theta_D$, the Debye temperature) scattering of the electrons by phonons tends to be isotropic so that the isotropic-τ assumption is valid. Dugdale and Firth[13] report that this assumption accurately describes their Hall measurements in Cu and Al lightly doped with impurities when $T\geq\Theta_D$. In this regime, $l_{\mathbf{k}}$ is proportional to $v_{\mathbf{k}}$ over the whole FS. At low temperatures, however ($T\ll\Theta_D$), scattering by phonons becomes more intense on FS segments with small calipers.[12] This tends to increase $l_{\mathbf{k}}$ on large-κ segments (the fingers and fjords in Fig. 4), relative to small-κ segments which have large diameters. In the case of Fig. 4(a), lowering T tends to favor the formation of negative loops [4(c)], whereas in the case of Fig. 4(b), it favors positive loops [4(d)].

In the impurity-scattering dominated regime at very low T, the scattering path length tends to become isotropic, since the average distance between impurities within the 2D plane is independent of direction. In this regime, A_l approaches πl_0^2, and ρ_{xy} approaches a number Γ which just measures the area-to-circumference ratio (Sec. VI). Such a situation may be realized experimentally by studying the low-T σ_{xy} in samples doped with impurities.

Thus, in the Bloch-Boltzmann theory with phonon scattering,[12] the variation of the anisotropy $\Delta l/l$ with T defines three characteristic regimes. In the first regime ($T\geq\Theta_D$) the isotropic-τ approximation is valid, and $\mathbf{k}$ dependence of $l_{\mathbf{k}}$ follows that of $v_{\mathbf{k}}$. Since $\Delta l/l$ does not change with further increase in T, R_H is T independent at high temperatures. In the second regime, in the vicinity of $0.2\Theta_D$, phonon scattering becomes highly anisotropic, and segments with high curvature and small caliper suffer increased scattering. R_H is strongly T dependent, often showing nonmonotonic behavior, as observed in transition metals. In the limit of zero T, however, the system approaches the isotropic-l approximation in the presence of impurities, and R_H is again a constant. (Inclusion of electron-electron interaction and weak localization effects complicate this regime.) The implications for ρ_{xy} will be described next.

VI. VARIATION OF ρ_{xy} WITH ANISOTROPY AND TEMPERATURE

Single band. Band-structure effects can greatly alter the value of the Hall coefficient R_H from the "free-electron" value $1/(ne)$. For a single band, it is customary to define the "Hall factor" $r\equiv(B/ne)\rho_{xy}$, with $\rho_{xy}=\sigma_{xy}/(\sigma_{xx}\sigma_{yy})$. I consider crystals that have N-fold symmetry in the x-y plane. If N is larger than 2, the in-plane conductivity is a scalar given by (Appendix A)

$$\sigma_{xx}/(e^2/h)=\sigma_{yy}/(e^2/h)=l_{\rm av}S/(2\pi)\ , \tag{7}$$

where S is the FS circumference and $l_{\rm av}$, the average of $l_{\mathbf{k}}$ over the FS, is given by

$$l_{\rm av}=\int ds\frac{l_{\mathbf{k}}}{S}\ . \tag{8}$$

Using Eqs. (4), (7), and (8), r is a scalar equal to

$$r=\Gamma A_l/(\pi l_{\rm av}^2)\ , \tag{9}$$

with

$$\Gamma=[(4\pi)A_{\rm FS}/S^2]\ . \tag{10}$$

Here, $A_{\rm FS}$ is the area enclosed by the FS curve [$n=A_{\rm FS}/(2\pi^2)$] and A_l is the Stokes area defined in Eq. (3). Equation (9) will be used to compare the present approach with calculations of ρ_{xy} on simplified FS models (Sec. VIII).

If we specialize to the isotropic-l case, r is simply equal to Γ (setting $l_{\rm av}=l_0$, and $A_l=\pi l_0^2$). By the isoperimetric inequality, $\Gamma\leq 1$, with equality holding for the circle. Thus, for a 2D metal with N-fold symmetry ($N>2$) and an *isotropic* $l_{\mathbf{k}}$, the Hall factor r is a direct measure of the FS area-to-circumference ratio, and is always ≤ 1. [The oft-repeated "rule"[4] that $r\leq 1$ in a wide range of FS shapes is merely a restatement of the isoperimetric inequality. In effect, the author(s) tacitly adopt the isotropic-l assumption.] For $N=2$, the same argument applies if the quantity $4\sigma_{xy}/(\sigma_{xx}+\sigma_{yy})^2$ is used in place of r (see Appendix A). If the FS is circular, but $l_{\mathbf{k}}$ is anisotropic, the inequality $\int(d\theta/2\pi)l_{\mathbf{k}}^2>l_{\rm av}^2$, together with

Eq. (9), implies that $r > 1$. Thus, any anisotropy in a circular FS always increases r, a result previously known.[14]

Returning to the general $l_{\mathbf{k}}$ case, Eq. (9) shows that r (or ρ_{xy}) is strongly influenced by three factors: the circumference S, the existence of negative curvature segments (which generate secondary loops in the l curve), and anisotropy in $l_{\mathbf{k}}$. While the effect of S is always to reduce r, as described for the case of isotropic l, the anisotropy in $l_{\mathbf{k}}$ may reduce or enhance r, depending on the local κ.

There are four cases to consider. For definiteness, I assume fourfold symmetry (the extension to N-fold is direct). These curves are either convex (cases I and II) or nonconvex (cases III and IV). In case I (II), $l_{\mathbf{k}}$ attains a maximum (minimum) on segments of large curvature. The nonconvex curves are distinguished by the sign of their secondary loops which are either negative (case III) or positive (IV). The results of Appendix B show that r deviates from Γ significantly as $\Delta l/l$ changes from zero. An examination of the four trends shows that, in the high-temperature (isotropic-τ) regime, cases II and III may be considered as conventional in the sense of Sec. V (i.e., $v_{\mathbf{k}}$, hence $l_{\mathbf{k}}$, scales inversely as $|\mathbf{k}|$), while the other two are anomalous. In the conventional cases, the bound $r \le \Gamma \le 1$ is valid (at high T). (The other cases I and IV, which require anomalous variation of $l_{\mathbf{k}}$ over the FS, have the opposite bound, $r > \Gamma$.)

Multiple bands. When the FS is comprised of several pockets, the above considerations for R_H are altered in an interesting way. If the arguments above are generalized to multibands in the standard way, I get

$$eR_H = (2\pi)^3 \left[\sum_i A_{li}\right] \Big/ \left[\pi \sum_i l_{\mathrm{av}i} S_i\right]^2 , \tag{11}$$

where subscript i denotes the ith band. (Each FS pocket generates its own A_l.) Again the two important regimes are the isotropic-τ ($T > \Theta_D$) and isotropic-l ($T \to 0$) extremes. It is convenient to introduce the reduced quantity $r' = (\sum_i n_i) eR_H$, where n_i is the carrier density enclosed in pocket i.

In the isotropic-τ case $l_{\mathrm{av}i}$ scales as the Fermi velocity in the ith pocket v_{Fi}. First, consider the denominator in Eq. (13). By the isoperimetric inequality, $S_i \le \sqrt{[(2\pi)^3 n_i]}$. Thus, r' satisfies the inequality

$$r' \le \left[\sum_i n_i\right]\left[\sum_i A_{li}\right] \Big/ \left[\pi \sum_i l_{\mathrm{av}i}\sqrt{n_i}\right]^2 . \tag{12}$$

In the important case when all pockets obey a quadratic dispersion (low carrier density metals), $l_{\mathrm{av}i}$ is proportional to $\sqrt{\varepsilon_F} \sim \sqrt{n_i}$, viz.,

$$l_{\mathrm{av}i} = \beta\sqrt{n_i} \quad (\varepsilon \sim k^2 \text{ and } \tau \text{ isotropic}) \tag{13}$$

(where β is a constant). For the numerator, the discussion in Sec. V implies that for "conventional" variation of $l_{\mathbf{k}}$ in the isotropic-τ limit

$$|A_{li}| \le \pi l_{\mathrm{av}i}^2 . \tag{14}$$

Using inequality (14) and Eq. (13) in Eq. (12), I get the general bound

$$r' \le \left[\sum_i A_{li}\right] \Big/ \left[\pi\beta^2 \sum_i n_i\right] \le \left[\sum_i (l_{\mathrm{av}i}/\beta)^2\right] \Big/ \left[\sum_i n_i\right] = 1 . \tag{15}$$

For this bound to hold, it is sufficient that τ be isotropic, and Eqs. (13) and (14) be valid in all pocketes. [We have assumed that all A_{li}'s > 0. If some of the A_{li}'s are negative, the second inequality in Eq. (15) is stronger.] Actually, in physical systems, the interband anisotropy overwhelms the intraband anisotropy, i.e., $|\Delta l/l|_{\mathrm{inter}} \gg |\Delta l/l|_{\mathrm{intra}}$, so that the bound Eq. (14) is less crucial than Eq. (13) for the derivation [i.e., the former may be slightly violated without invalidating Eq. (15)]. *Thus, in the high-T limit, the Hall coefficient of a 2D multiband system, normalized to the total carrier density (counting all FS pockets), is less than or equal to one, if quadratic dispersion is valid.* For the special case when all FS pockets are circular r' is precisely 1 (even if the $l_{\mathrm{av}i}$'s are all distinct). We encounter an example in the next section.

The zero-T case will be considered briefly. If $l_{\mathbf{k}}$ is isotropic with the same value in all FS pockets, Eq. (12) reduces to

$$r' \le \left[\sum_i n_i\right] M \Big/ \left[\sum_i \sqrt{n_i}\right]^2 \equiv \Omega , \tag{16}$$

where M is the total number of pockets and the number Ω is as defined. It is readily seen that $\Omega \ge 1$ for all $\{n_i\}$, so that a useful bound cannot be imposed on r' (except when all the pockets are circular, in which case $r' = \Omega \ge 1$). This situation contrasts with the single-band case where r equals $\Gamma \le 1$. The physical origin of this difference is that in the isotropic-l limit, all pockets contribute equally to σ_{xy} whereas σ_{xx} is dominated by the large caliper pockets. (We have not considered scattering mechanisms other than impurity scattering in this limit. The inclusion of electron-electron scattering may alter this picture significantly.)

VII. APPLICATIONS

2H-NbSe$_2$. The dichalcogenide 2*H*-NbSe$_2$ is a quasi-2D layered compound that undergoes a charge-density-wave (CDW)[15] transition at 38 K. The band-structure results of Mattheiss[16] show that the FS in the first zone is comprised of a cylindrical hole surface (FS1) with its axis along KH (Fig. 5).[17] A second hole surface (FS2) is either closed (an oblate spheroid centered at A) or an open cylinder like FS1 with axis along ΓA (the calculation is not accurate enough to decide).

The measured[18] Hall coefficient R_H^{meas} equals $+4.60\times10^{-10}$ m^3/C, and changes by $<2\%$ between 350 and 60 K. Below $\sim$40 K, R_H^{meas} decreases steeply, becoming negative at low T. From Mattheiss' calculations,[16,17] the FS cross sections are very close to circular. I compute the FS areas A_{FS1}, A_{FS2} to be 0.129 and 0.245, respectively, in units of $(2\pi/a)^2$, which implies that $n_1 = 2.19\times10^{14}$ cm^{-2}, $n_2 = 4.14\times10^{14}$ cm^{-2}. (The lattice parameter $a = 3.440$ Å. Note that $2n_1 = n_2$, to a few

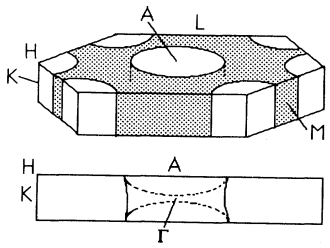

FIG. 5. The first zone of $2H$-$NbSe_2$ showing the holelike Fermi surfaces calculated by Mattheiss (Ref. 16). The lower drawing shows a vertical section the the zone. The surface FS2 is either an oblate spheroid centered at A (broken lines) or an open cylinder (solid lines).

% accuracy.) Thus, the total density predicted from the band structure is $n_{2D}=2n_1+n_2=8.51\times10^{14}$ cm^{-2}. (There are two FS1 pockets.) This number (0.872/Nb) is $\sim$13% lower than the "chemical" estimate of 1 hole/Nb. However, if I calculate the reduced quantity $r'=n_{2D}eR_H{}^{\rm meas}$, instead, I find that it equals 1.00, i.e.,

$$eR_H{}^{\rm meas}=1/(2n_1+n_2)\quad (T\sim350\ \text{K})\ . \tag{17}$$

Since the system is quasi-2D above 38 K (the anisotropy ρ_c/ρ_a is between 20–30, and T independent[19]), I may apply the reasoning (in reverse) of Eqs. (11)–(15) to this three-pocket system. From Eq. (11),

$$eR_H=(2\pi)^3(2A_{l1}+A_{l2})/\pi(2l_{\rm av1}S_1+l_{\rm av2}S_2)^2\ . \tag{18}$$

As the FS pockets are close to circular, $A_{li}=\pi l_{{\rm av}i}{}^2$, and $S_i=\sqrt{[(2\pi)^3n_i]}$. Also, $2n_1=n_2$. Substituting these in Eq. (18), and comparing with the measured value in Eq. (17), I obtain the equation $(2l_{\rm av1}{}^2+l_{\rm av2}{}^2)/(2l_{\rm av1}+l_{\rm av2}\sqrt{2})^2=\frac{1}{4}$, which has the solution $l_{\rm av1}/l_{\rm av2}=1/\sqrt{2}$, i.e., the path length on FS1 is shorter than that on FS2 by the factor $\sqrt{2}$. Since the Fermi velocity ratio $v_{F1}/v_{F2}=1/\sqrt{2}$ (both FS1 and FS2 have quadratic dispersion[16]), I conclude that at 350 K the scattering time is isotropic, and that the high-T R_H in $2H$-$NbSe_2$ confirms Eq. (15).

The two assumptions that the dispersion in $NbSe_2$ is 2D above 60 K, and that the FS cross sections are as given by Mattheiss, provide a quantitatively accurate and self-consistent analysis of the high-T Hall coefficient. The scattering time is found to be isotropic, and the scattering path lengths scale as the Fermi velocities in the two inequivalent pockets, as expected for $T\geq\Theta_D$. The self-consistency achieved favors an open cylinder for FS2 over an oblate spheroid.

This system should also provide a test of the $T\rightarrow0$ case [Eq. (16)], but the situation below the CDW transition is quite complicated. Magnetothermal oscillation experiments[17] below 3.5 K reveal the existence of a small FS pocket that is not apparent in the band structure. The large swing of R_H to negative values implies a drastic change in FS topology.

Layered cuprate perovskites. An important class of quasi-2D conductors are the cuprate oxides that become superconducting at high T. Conductivity anisotropies in these families range from 10^2 (in $YBa_2Cu_3O_7$ or "1:2:3") to 10^4 in the Bi-based materials ($Bi_2Sr_2CaCu_2O_8$ "Bi 2:2:1:2" and $Bi_2Sr_2CuO_6$ "Bi 2:2:0:1").[20] The Hall coefficient has been extensively studied in single crystals of these compounds as well as in $La_{2-x}Sr_xCuO_4$ ("214").[21] There appears to be a pervasive pattern of behavior in the R_H vs T curves.[22] In the compounds 1:2:3, Bi 2:2:1:2, and 214 ($0.14<x<0.25$), R_H is strongly T dependence, and approximates the behavior $R_H=(c+dT)^{-1}$. In Bi 2:2:0:1, the T dependence is less pronounced, but still significant.

The interesting question is Can the T dependence of R_H (particularly in 1:2:3) be explained within the conventional Bloch-Boltzmann theory, assuming phonon scattering alone? As discussed in Sec. V, in the conventional theory R_H is T dependent only if the anisotropy $\Delta l/l$ changes with T. In principle, even with one band, an increase of $\Delta l/l$ with decreasing T can cause R_H to grow significantly (as in case III).

However, if phonon scattering alone is operative, the variation of $\Delta l/l$ must display the three distinct regimes determined by Θ_D (as discussed in Sec. V). At $T>\Theta_D$($\sim$440 K in 1:2:3), R_H should saturate to a T-independent value. In 1:2:3, Fiory and Grader have measured[23] R_H up to 600 K, and find no evidence of saturation. More high-temperature measurements of R_H are clearly desirable. At very low T, R_H again saturates to a different constant. Unfortunately, the low-T regime is inaccessible in 1:2:3. Interestingly, in $Nd_{2-x}Ce_xCuO_4$, in which the superconductivity can be suppressed with an 8-T field, Wang *et al.*[24] have found that R_H remains *strongly* T dependent down to 2 K, in contrast with σ_{xx} which is T independent below 20 K. The persistence of the monotonic variation of R_H with T in these cuprates over such a wide range in T is anomalous. The absence of any saturation in R_H in 1:2:3 at T above 440 K (if further corroborated) shows that the T dependence is not caused by phonon scattering. This anomaly shows, even more clearly than the linear-T behavior of the resistivity, that the scattering mechanism is *not* phononic alone but dominated by an anomalous mechanism, probably electronic in origin.

VIII. COMPARISON WITH PREVIOUS MODEL CALCULATIONS

Over the years there have been systematic efforts[4–8] to understand the factors that influence the weak-field σ_{xy}. Most of these studies have relied on direct calculations on families of FS shapes approximated by simple geometric figures (cubes, spheroids, dodecahedra, etc.). Some empirical rules have been deduced, based on these studies. However, due to their empirical nature, these rules have limited ranges of validity. A significant advantage of the present geometric picture is that many such calculations are greatly simplified. Moreover, the insight afforded clarifies why certain approaches fail in the weak-field limit.

First, we consider a cubic FS in 3D (or a square in 2D). Early attempts[4] found that, in the isotropic-l approximation, r equals $\frac{1}{2}$ in both 3D and 2D. ("Edge" and "face" of a 3D polygon will be taken to mean "vertex" and 'side" when referring to a 2D polygon.) Subsequently, it was realized that the large Hall angle at the sharp edges violated the weak-field assumption. Instead, if the edges are replaced by a rounded surface of curvature $1/\Delta k$, then r is found to be $\pi/4$ (in the limit $\Delta k \to 0$).[6] This may be seen to be a direct consequence of the isoperimetric inequality. By Eq. (9), r simply equals Γ (when l is isotropic) for any value of Δk, and $\Gamma = \pi/4$ for a square. It is instructive to see why the direct computation yields an r $(=\frac{1}{2})$ that differs from the limiting approach $(\pi/4)$. In the former scheme, the velocity [hence $\boldsymbol{l}(\mathbf{k})$] is a constant, say $v\hat{\mathbf{x}}$, along one face of the cube (or one side of the square in 2D), and another constant $v\hat{\mathbf{y}}$ along an adjacent face. I assume $\mathbf{E}\|\hat{\mathbf{x}}$ and $\mathbf{H}\|\hat{\mathbf{z}}$. Under the Lorentz force, all electrons on the (100) face and within δk of the edge "spill" over the edge onto the (010) face $(\delta k = e\tau vH/\hbar)$.[5] The total Hall current i_y equals $(4/h^3)\delta k(2k_F eE\tau/\hbar)ev$ [I will call this the "edge-current" argument. For instance, this argument is central to the "planar-faced energy surface" (PFES) scheme,[5] which approximates the FS by a finite number of plane faces.]

Now, since $\mathbf{v}$ is a constant vector over each planar face, the whole face maps into a *single* point $\boldsymbol{l}(\mathbf{k}) = \mathbf{v}\tau$, when we go into the space of $\boldsymbol{l}(\mathbf{k})$. Therefore, the edge-current model for an N-sided FS defines *only* N points in $\boldsymbol{l}$ space, instead of a closed curve. Without specifying in detail the path between two two points, the computation of A_l (and hence, r) is left ambiguous. In general, then, σ_{xy} cannot be computed by approximating a FS by a finite number of planar surfaces (or polygon in 2D). The contribution of the planes to σ_{xy} is identically zero (because their image in $\boldsymbol{l}$ space has zero measure). On the other hand, the jump in the angle of $\mathbf{v}$ at the edge actually accounts for *all* of the area, but its path is left undefined in such models. The path corresponding to the edge-current argument disagrees with the isotropic-l path.

Six years after the introduction of the PFES scheme, Cowley and Allgaier (CA)[7] in effect showed its failure in the weak-field limit by rounding the sharp cusps in the four-fold "star" FS (Fig. 6). (CA relegate the PFES results to the "intermediate-field" regime, but this has not been justified, to my knowledge.) It is interesting to derive their results from the present picture. To study anisotropy effects, CA also let $l_{\mathbf{k}}$ vary linearly on each face from l_0 on the "α edge" to l_1 on the "β edge." They obtain

$$r = (\pi/8)(1+\sin\alpha-\cos\alpha)(1+6q+q^2-8q\alpha/\pi) ,$$

where $q = (l_1 - l_0)/(l_1 + l_0)$. CA implicitly assume that $l_{\mathbf{k}}$ is frozen at the value $l_0(l_1)$ when $\mathbf{k}$ moves around the smoothed α edge (β edge). Thus, the l curve is a complicated self-intersecting loop that resembles a fan with four blades (Fig. 6). Its Stokes area is that of the hub minus that of the blades, i.e.,

$$\begin{aligned} A_l &= \pi l_0^2 - \pi(l_1^2 - l_0^2)(1-2\alpha/\pi) \\ &= \pi l_0^2(1+6q+q^2-8q\alpha/\pi)/(1+q^2) . \end{aligned}$$

Using the result that $l_{\mathrm{av}} = (l_0 + l_1)/2$, we see that r is just the product of $\Gamma = (\pi/8)[1+\sin\alpha-\cos\alpha]$ and the factor $A/\pi l_{\mathrm{av}}{}^2$, in agreement with Eq. (9). [When q is <0 (>0) this example belongs to case III (IV), as described in Sec. V.]

Banik and Overhauser[8] also use the edge-current argument to explain why aluminum has an electron-type weak-field σ_{xy} even though the FS has global holelike topology. They approximate the FS by N arcs joined at a sharp cusp. Using the isotropic-l approximation, they purport to show that r is determined by the arcs and is therefore electronlike. However, in the weak-field regime, the results of Sec. VI show that r will be strictly equal to Γ, and no change in sign can be expected. (The contributions of the cusps, if appropriately rounded and the limit $\Delta k \to 0$ taken, must dominate the negative curvature arcs to give precisely πl_0^2 for A_l, as shown in Sec. IV.) Banik and Overhauser[8] add the caveat that their model does not apply to the weak-field limit (the cusps are assumed to be infinitely sharp).

The present picture also shows a way to calculate accurately the weak-field σ_{xy} of an arbitrary FS. The prescription is that the actual summation must be done in $\boldsymbol{l}$ space. If a 2D FS is approximated by an N-sided polygon, the map into $\boldsymbol{l}$ space consists of N points. σ_{xy} is best estimated by computing the Stokes area of the l curve approximated by joining the N points with smooth arcs, rather than by computing the edge current between two adjacent faces. [From the preceding, it is also necessary to know (or specify) how $l_{\mathbf{k}}$ changes from face to face, because by assuming an isotropic $l_{\mathbf{k}}$, one merely recovers $r = \Gamma$.]

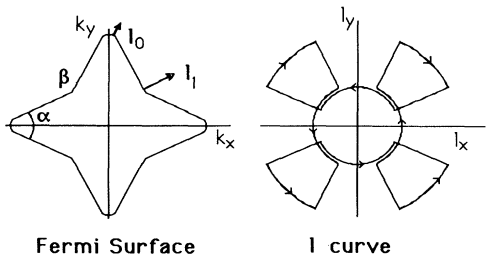

FIG. 6. The star-shaped FS studied in Ref. 6. On each plane surface, $l_{\mathbf{k}}$ varies linearly from l_1 to l_0 as shown. The l curve of this FS resembles a fan with four blades. (Overlapping lines are drawn displaced for clarity.)

IX. DISCUSSION

The simple geometric representation for σ_{xy} in 2D suggests that there may exist a connection between the flux threading A_l and the Hall current, that is more fundamental than indicated by the Boltzmann derivation. Although I have not uncovered this connection, it seems

that there are two lines of speculation worth pursuing. Is there a relationship between Eqs. (4) and the quantized Hall effect[24] (QHE), i.e., is the former a semiclassical "echo" of the latter? A positive answer may seem unlikely since the two phenomena are at opposite extremes of field. [In the QHE case, $\sigma_{xy}/(e^2/h)=n\phi_0/B$, i.e., σ_{xy} varies inversely with B, whereas in the weak-field limit, σ_{xy} is linear in B.] Nonetheless, the fact that in both limits the number of flux quanta (in a physically well-defined area) precisely determines σ_{xy} seems to me highly suggestive. [The extra factor of 2 in Eq. (4b) is due to spin degeneracy.]

The close similarity between A_l and magnetic flux threading through a twisted loop also suggests the following picture. Equation (4b) implies that the weak-field σ_{xy} normalized to e^2/h is the total change in the Aharonov-Bohm phase $\Delta\phi/2\pi$ of a charged particle taken around the closed l curve. Within the time τ, each electron maintains phase coherence (I assume τ is the dephasing time). It may be possible to relate A_l to the phase change of electrons on the FS in the time τ, and from the phase change to σ_{xy}, in suggestive analogy with Thouless *et al.*[25] One may also interpret the area A_l as a measure of the "divergence" of the flow of the electron gas in the time τ. (If the FS is flat, both the divergence and A_l are zero, whereas for a convex FS both quantities are a maximum.) It is plausible that the notion of spreading implied by the divergence is related to the phase change of the total wave function of the system, through the flux penetrating the spreading area in the time τ, although this has not been demonstrated.

These speculations aside, the geometric representation provides a powerful way to assess the Hall current in quasi-2D systems regardless of FS shape. A major difficulty in comparing the experimental σ_{xy} with predictions of band-structure models is that the integral in Eq. (1) is difficult to estimate.[1,2] However, most of this difficulty stems from working in the wrong space, as I have argued here. In $\mathbf{k}$ space, σ_{xy} is dominated by FS regions with very large curvature, and schemes such as the edge-current argument exacerbate errors by introducing artificially sharp edges. By going into the space of l, these difficulties are largely avoided. Part of the appeal of the geometric representation is its simplicity. Regardless of the FS shape, σ_{xy} may be effectively calculated, often simply by approximating the area A_l, as shown in Appendix B.

The new perspective enables one to disentangle the influences of three major factors that determine σ_{xy} the ratio of $4\pi A_{\rm FS}/S^2$, the anisotropy λ, and the local curvature κ. The separation of physical effects (scattering) from purely geometric effects (such as the isoperimetric inequality) clarifies the variation of R_H with T. It seems that Hall measurements performed near the Debye temperature are considerably easier to compare with band-structure calculations, than those obtained at low T. In both single and multiband systems, the bound $r\leq 1$ is valid under a wide range of conditions for $T\geq\Theta_D$, and this is often sufficient to discriminate competing models applied to novel metals. Where the conventional Bloch-Boltzmann theory with phonon scattering is valid, the analysis is quite effective, as in the case of $2H$-$NbSe_2$. Unfortunately, there seems to be a paucity of high-T Hall results in anisotropic (2D) metals. In systems, such as the cuprate perovskites, where the conventional theory fails, the present analysis shows why the R_H-T profile is anomalous, and a nonphononic scattering mechanism is required.

ACKNOWLEDGMENTS

I thank Paul Wiegmann for helpful discussions. This research is supported by the Office of Naval Research (Contract No. N00014-90-J-1013) and a Grant from the Seaver Foundation.

APPENDIX A

In 2D, the diagonal components of the conductivity are given by[12] $\sigma_{xx}=(e^2/h\pi)\int ds\, l_{\mathbf{k}}\cos^2\theta_{\mathbf{k}}$, where the integral is over the FS, and $\theta_{\mathbf{k}}$ is the angle between $\boldsymbol{l}(\mathbf{k})$ and $\hat{\mathbf{x}}$. For a 2D crystal with N-fold symmetry ($N>2$), I divide the FS into N identical wedges, which may be mapped into each other by one of the crystal group rotations, except that $\theta_{\mathbf{k}}$ is changed to $(\theta_{\mathbf{k}}+m\alpha)$ on the mth wedge. Hence, σ_{xx} simplifies to

$$\sigma_{xx}/(e^2/h)=\pi^{-1}\int_{\Delta S} ds\, l_{\mathbf{k}} \sum_{m=1}^{N}\cos^2(\theta_{\mathbf{k}}+m\alpha)\ , \tag{A1}$$

where $\alpha=(2\pi)/N$. The integration is confined to one wedge of length $\Delta S=S/N$. The sum equals $N/2$. Using the relationship $l_{\rm av}S=N\int ds\, l_{\mathbf{k}}$, I get Eqs. (7) and (8). The case $N=2$ needs a slight modification. Adding σ_{xx} and σ_{yy} together, I get $(\sigma_{xx}+\sigma_{yy})/2=(e^2/h)(l_{\rm av}/2\pi)S$. The discussion in Sec. VI on ρ_{xy} can be extended to include the $N=2$ case by considering the quantity $4\sigma_{xy}/(\sigma_{xx}+\sigma_{yy})^2$, instead of ρ_{xy} itself.

APPENDIX B

The variation of r with anisotropy with a FS with fourfold symmetry is estimated here for the four cases considered in Sec. V.

Cases I and II are closed convex curves with four short segments that have large κ. The extreme example is a square of side $2k_F$ with rounded corners that have curvature $1/\Delta k$. I assume that the anisotropy in $l_{\mathbf{k}}$ is given by a Gaussian curve centered at $s=s_0$, viz.,

$$l(s)=l_0+\Delta l\exp[-(s-s_0)^2/\Delta s_1^2]\ . \tag{B1}$$

Δl is positive (negative) in case I (II). The width Δs_1 is chosen to satisfy $\Delta s_1/\Delta k=\pi/4$, so that $l(s)$ is essentially constant over the flat segments. With this variation in $l(s)$, r becomes [Eq. (9)]

$$r=\Gamma[1+(\sqrt{\pi})\lambda+(\sqrt{\pi}/8)\lambda^2]/[1+(\pi^{3/2}/8)\eta\lambda]^2\ , \tag{B2}$$

where $\lambda=\Delta l/l_0$ and $\eta\equiv\Delta k/k_F\ll 1$. For positive λ

(case I), r increases linearly with λ, and then saturates to the value c/η^2 (c a constant) for very large λ. In this limit, the corners dominate both σ_{xx} and σ_{xy}, so that r is mainly determined by a small FS "pocket" of effective area $\pi\Delta k^2$. For negative Δl (case II), r decreases monotonically to zero (as λ approaches -1, its largest possible negative value). This limit corresponds to a large FS with flat surfaces. The isotropic-τ regime ($T \geq \Theta_D$) corresponds to case II, whereas the isotropic-l regime corresponds to the point $\lambda=0$. (Case I is anomalous at high T since the velocity at the corner is usually lower than on the flat face.)

In cases III and IV, the negative-κ segments generate secondary loops which are negative and positive, respectively (see Fig. 4). First, consider a negative loop. Its area I estimate as follows. At the two points on the FS where κ vanishes, let the vector $\boldsymbol{l}(\mathbf{k})$ be $\boldsymbol{l}_1(\mathbf{k}_1)$ and $\boldsymbol{l}_2(\mathbf{k}_2)$, and their angles (relative to $\mathbf{x}$) be θ_1 and θ_2, respectively (see Fig. 2). Since $\mathbf{k}_1$ and $\mathbf{k}_2$ are "turning" points of $\boldsymbol{l}(\mathbf{k})$, and are symmetrically located around a symmetry axis, the vectors $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ define the cone that the secondary loop subtends at the origin ($|\boldsymbol{l}_1|=|\boldsymbol{l}_2|=l_c$). The cone angle is the difference angle $\Delta\theta=\theta_2-\theta_1$ given by $\Delta\theta = \int_1^2 ds\,\kappa$. The secondary loop may be approximated by an ellipse of semiradii $\Delta l/2$ and $l_c\sin(\Delta\theta/2)$. [$\Delta l=(l_{\max}-l_0)$, where $l_{\max}$ is the maximum value of $l_{\mathbf{k}}$ on PQ and l_0 the minimum value of $l_{\mathbf{k}}$ over the whole FS.] The area A_s for a single loop is approximated by

$$A_s \sim (\pi/2)\Delta l l_c \sin(\Delta\theta/2) \ . \tag{B3}$$

I obtain the criterion for a negative σ_{xy} by making $4A_s$ larger than A_p. If I approximate l_c by $l_0+\Delta l/2$) and A_p by πl_0^2, I find that the critical anisotropy (at which σ_{xy} vanishes) is given by $(\Delta l/l_0)_c = \{1+[\sec(\Delta\theta/2)]/4\}^{1/2}-1$.

For case III curves, the area A_l is the difference between A_p and $N_s A_s$, with A_s given by Eq. (B3). If $l(s)$ is given by Eq. (B1) ($\Delta l>0$), and A_p is approximated by πl_0^2, the Hall factor equals

$$r=\Gamma[1-\lambda(1+\lambda/2)\sin(\Delta\theta/2)]/[1+4\lambda(\Delta s/S)\sqrt{\pi}]^2, (\lambda>0) \ . \tag{B4}$$

The width Δs is of the order $\Delta\theta/2|\kappa_n|$, where κ_n is the average curvature on the negative-κ segments. As λ increases from 0, r decreases linearly from Γ, reaching zero at $(\Delta l/l_0)_c$. For very large λ, r saturates to the negative value $-\Gamma(S/\Delta s)^2/16\pi$. The variation of case IV may be estimated from Eq. (B4) with $\lambda<0$. As $|\lambda|$ increases from 0, r increases linearly, eventually saturating to a positive value as $\lambda\rightarrow -1$.

The high-T, isotropic-τ regime corresponds to case III (negative secondary loops), but r may be of either sign depending on λ. As discussed in the main text, case IV is anomalous since it requires negative-κ segments to have shorter $l_{\mathbf{k}}$. For example, in Fig. 4(b) one physically expects $l_{\mathbf{k}}$ to be larger at the fjords since the velocity is higher there, so that negative loops are generated at high T.

[1] C. M. Hurd, *The Hall Effect in Metal and Alloys* (Plenum, New York, 1972).

[2] *The Hall Effect and Its Applications,* edited by C. L. Chien and C. R. Westgate (Plenum, New York, 1979).

[3] M. Tsuji, J. Phys. Soc. Jpn. **13**, 979 (1958).

[4] R. S. Allgaier, in *The Hall Effect and its Applications* (Ref. 2), p. 375.

[5] R. S. Allgaier, Phys. Rev. **165**, 775 (1968).

[6] R. S. Allgaier and Robert Perl, Phys. Rev. B **2**, 877 (1970).

[7] P. H. Cowley and R. S. Allgaier, Philos. Mag. **29**, 111 (1974).

[8] N. C. Banik and A. W. Overhauser, Phys. Rev. B **18**, 1521 (1978).

[9] J. J. Rhyne, Phys. Rev. **172**, 523 (1969); A. Fert, Physica B **86**, 491 (1976).

[10] L. M. Falicov, A. B. Pippard, and P. R. Sievert, Phys. Rev. **151**, 498 (1966); L. M. Falicov and P. R. Sievert, Phys. Rev. **138**, A88 (1965).

[11] H. Jones and C. Zener, Proc. R. Soc. A **145** (1934); see Colin M. Hurd, in *The Hall Effect and its Applications* (Ref. 2), p. 1, for an excellent survey.

[12] J. M. Ziman, Phys. Rev. **21**, 1320 (1961); J. M. Ziman, *Electrons and Phonons* (Clarendon, Oxford, 1960).

[13] J. S. Dugdale and L. D. Firth, J. Phys. C **2**, 1272 (1969).

[14] K. Boning, Phys. Konden. Mater. **11**, 177 (1970).

[15] J. A. Wilson, F. J. DiSalvo, and S. Mahajan, Adv. Phys. **24**, 117 (1975).

[16] L. F. Mattheiss, Phys. Rev. B **8**, 3719 (1973).

[17] John E. Graebner and M. Robbins, Phys. Rev. Lett. **36**, 422 (1976).

[18] H. N. S. Lee, H. McKinzie, D. S. Tannhauser, and A Wold, J. Appl. Phys. **40**, 602 (1969).

[19] T. W. Jing and N. P. Ong (unpublished).

[20] S. W. Tozer, A. W. Kleinsasser, T. Penney, D. Kaiser, and F. Holtzberg, Phys. Rev. Lett. **59**, 1768 (1987); S. J. Hagen, T. W. Jing, Z. Z. Wang, J. Horvath, and N. P. Ong, Phys. Rev. B **37**, 7928 (1988); D. A. Brawner, Z. Z. Wang, and N. P. Ong, *ibid.* **40**, 9329 (1989).

[21] For a review, see N. P. Ong, in *Physical Properties of High-Temperature Superconductors II,* edited by D. M. Ginsberg (World Scientific, Singapore, 1990), p. 459.

[22] J. Clayhold, N. P. Ong, Z. Z. Wang, J. M. Tarascon, and P. Barboux, Phys. Rev. B **39**, 7324 (1989).

[23] A. T. Fiory and G. S. Grader, Phys. Rev. B **38**, 9198 (1988).

[24] Z. Z. Wang, T. R. Chien, N. P. Ong, J. M. Tarascon, and E. Wang (unpublished); also see *Physical Properties of High-Temperature Superconductors II* (Ref. 21).

[25] For a survey, see *The Quantum Hall Effect,* edited by Richard E. Prange and Steven M. Girvin (Springer, Berlin, 1987).

[26] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982); D. J. Thouless (Ref. 25).