Hall conductivity in the presence of spin-orbit interaction and disorder

Louis-François Arsenault*
Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854-8019, USA

B. Movaghar†
Département de Génie Physique and Regroupement Québécois sur les Matériaux de Pointe (RQMP), École Polytechnique de Montréal, CP 6079, Succursale "Centre-Ville," Montréal, Québec, Canada H3C 3A7

(Received 8 January 2008; revised manuscript received 14 August 2008; published 5 December 2008)

Starting from the Kubo formula, we expand the Hall conductivity using a cumulant approach which converges quickly at high temperatures ($k_BT >$ energy differences of initial and final scattering states) and can be extended to low temperatures. The theory can deal with the sign, the ordinary and the anomalous contributions to the Hall effect. When applied to include the spin-orbit interaction to first order, we recover what is essentially the Karplus-Luttinger result for the anomalous Hall effect. Contact is made to the Chazalviel and Nozières-Lewiner formulae. A side-jump-like formula is obtained by using an exact application of linear response. We show that there exists an exact rigid Hall current which is not a Fermi level property. We introduce a relationship between mass and diffusivity which allows us to generalize the theory to strong disorder and even introduce a mobility edge. The formalism provides a systematic and practical way of analyzing both ordinary and anomalous contributions to the Hall conduction including the changes of sign, and in the presence of serious disorder. As a by-product of the method, we show that the anomalous Hall coefficient can vary with resistance to the power n, with $1 \leq n \leq 2$ depending on the degree of coherence.

DOI: 10.1103/PhysRevB.78.214408 PACS number(s): 75.47.−m, 72.10.Bg, 72.15.Gd, 72.80.Ng

I. INTRODUCTION

The Hall conductivity of materials exhibits a wide and rich variety of behavior. The interpretation is, in general, still very difficult, even though, in principle, the information is contained in the Kubo formula. This is true at least in linear response to the applied electric field. The Kubo formula is therefore the starting point of our analysis of Hall conductivity. We include the spin-orbit interaction and discuss the so-called anomalous Hall effect (AHE). Previously, most work was focused on understanding the sign change of the Hall coefficient in ordered and disordered systems,[1] localization, and the quantum Hall effect (QHE).[2] More recently, the problem has been to understand the effect of magnetism and of many-body corrections on the Hall effect. Considerable progress has been made recently by integrating the spin-orbit coupling into the Bloch wave function formalism of Kane in crystals, and applying these wave functions and other first-principles numerical methods[3] to study the AHE effect in magnetic materials.[4] In these papers, the emphasis is on order, and disorder is only represented by a uniform lifetime. Most magnetically doped semiconductors and magnetic alloys can, however, not really be considered to be in the weak scattering regime, and therefore, in the present approach, we have inverted this priority. We emphasize the absence of Bloch symmetry rather than its presence.[5–10] The aim is to develop practical formulae which can deal with disorder, the sign, the side-jump Hall effect problem, and the skew scattering/intrinsic Hall effect problem.

The AHE is now a well established observation in magnets, and a number of explanations have been proposed,[7] and will be discussed here. All explanations are based, in one way or another, on the spin-orbit mechanism. Thus, it is accepted that spin-orbit coupling causes the anomalous Hall conductivity contribution.

Much progress has been made recently on understanding the origin of the AHE and this has generated exciting physics. Traditionally it was thought that only the spin-orbit *skew scattering mechanism*[1,11,12] gives rise to a magnetism (M_z) dependent Hall coefficient. This process depends on the conductivity relaxation time or relaxation time squared, depending on whether the skew scattering is itself rate determining for conductivity or not. Traditionally, skew scattering is derived as an extrinsic effect; i.e., it is due to impurity scattering, and not to the host spin-orbit interaction. The intrinsic spin-orbit interaction produced by the host crystal potential had been invoked as a source of AHE by Karplus and Luttinger[10] but later rejected by Smit[13] who claimed that the intrinsic effect is negligible. Earlier, it had been shown by Mott and Massey[14] that electron scattering from Coulomb potentials is asymmetric with respect to spin direction, with spin up going more to one side, and those with spin down more to the other. When the electron gas is magnetized, there is a net transverse Hall current. When the asymmetric scattering is at impurity sites, this process is called skew scattering as mentioned above, and has been discussed by several authors.[11,12] The important point about skew scattering is that the net spin-orbit coupling at the impurity site can be strongly enhanced by the host crystal. This is really what makes this process so important. For a detailed account of the history and progress in understanding the AHE, see Ref. 8.

II. KUBO FORMULA FOR THE LONGITUDINAL AND TRANSVERSE CONDUCTIVITY

The frequency (ω) dependent conductivity in linear response to an electric field is usually written as[1,15,16]

$$\sigma_{\mu\nu} = \frac{i\hbar e^2}{\Omega} \lim_{\delta \to 0} \sum_{\alpha,\beta} \frac{\langle \alpha | v_\mu | \beta \rangle \langle \beta | v_\nu | \alpha \rangle}{\varepsilon_\alpha - \varepsilon_\beta + \hbar\omega + i\delta} \frac{f(\varepsilon_\alpha) - f(\varepsilon_\beta)}{\varepsilon_\beta - \varepsilon_\alpha}. \quad (1)$$

This form is general for any exact set of eigenstates $|\alpha\rangle$ and energies ε_α; spin summation is implied. The $f(\varepsilon)$ are the Fermi functions and Ω is the volume. The velocity operators v_μ are given by Heisenberg's equation of motion.

For the general case of a material which need not be periodic, the spin-orbit term in the Hamiltonian is

$$H_{\rm so} = \frac{\hbar}{4m^2c^2} [\nabla V(\mathbf{r}) \times \mathbf{p}] \cdot \boldsymbol{\sigma}, \quad (2)$$

where

$$V(\mathbf{r}) = \sum_n \frac{eZ_n}{4\pi\varepsilon\varepsilon_0 |\mathbf{r} - \mathbf{R}_n|} - eFx. \quad (3)$$

In Eqs. (2) and (3), m is the bare electron mass, c is the speed of light, $\mathbf{p}$ is the momentum operator, and $\boldsymbol{\sigma}$ is the Pauli spin operator which is a vector containing the Pauli's matrices, i.e., $[\sigma_x, \sigma_y, \sigma_z]$. In Eq. (3), F is the external applied electric field, e is the electric charge, Z_n is the effective local charge, $\varepsilon\varepsilon_0$ is the permittivity, $\mathbf{r}$ and x are positions operators for the charges, and $\mathbf{R}_n$ is the position of the fixed ions that make the lattice. The velocity operators can then be written as

$$v_x = v_x^0 + \frac{\hbar}{4m^2c^2} [\nabla_z V(\mathbf{r}) \sigma_y - \nabla_y V(\mathbf{r}) \sigma_z], \quad (4)$$

$$v_y = v_y^0 - \frac{eB_z x}{m} + \frac{\hbar}{4m^2c^2} [\nabla_x V(\mathbf{r}) \sigma_z - \nabla_z V(\mathbf{r}) \sigma_x], \quad (5)$$

where we have used the Landau gauge for the vector potential, $\mathbf{A} = (0, B_z x, 0)$. The v_x^0 and v_y^0 are $-\frac{i\hbar}{m}\frac{\partial}{\partial x}$ and $-\frac{i\hbar}{m}\frac{\partial}{\partial y}$, respectively. The choice of the minus sign in $-eB_z x/m$ implies that $e = -|e|$ since the kinetic part of the Hamiltonian for a charge q in a field is $T = 1/2m(\mathbf{p} - q\mathbf{A})^2$.

The spin-dependent terms in Eqs. (4) and (5) can be important in magnets. We shall consider them explicitly in Sec. V.

The Hall effect is given by the antisymmetric part of the transverse conductivity.[11,17] To write the antisymmetric Hall conductivity, we note that for nonzero frequency it is the anti-Hermitian part[11] $\sigma_{xy}^a = \frac{1}{2}(\sigma_{xy} - \sigma_{yx}^*)$. In the dc limit we have

$$\sigma_{xy}^a = \frac{\hbar e^2}{\Omega} \lim_{\delta \to 0} \sum_{\alpha,\beta} -i \langle \alpha | v_x | \beta \rangle \langle \beta | v_y | \alpha \rangle \frac{f(\varepsilon_\alpha) - f(\varepsilon_\beta)}{(\varepsilon_\beta - \varepsilon_\alpha)^2 + \delta^2}. \quad (6)$$

We shall now derive the Hall conductivity in a general way which will include both the skew scattering and the intrinsic contributions.

III. HALL EFFECT USING A CUMULANT EXPANSION OF THE KUBO FORMULA

One way to derive the contributions to the Hall effect, with disorder present, is to rewrite the Kubo formula using the Heisenberg equation of motion,

$$\langle \alpha | x | \beta \rangle = -i\hbar \frac{\langle \alpha | v_x | \beta \rangle}{\varepsilon_\alpha - \varepsilon_\beta}, \quad (7)$$

which is true in any finite box (length L) without dissipation. Then Eq. (6) becomes

$$\sigma_{xy}^a = \frac{e^2}{\Omega} \lim_{\delta \to 0} \sum_{\alpha,\beta} \frac{[f(\varepsilon_\alpha) - f(\varepsilon_\beta)](\varepsilon_\alpha - \varepsilon_\beta)}{(\varepsilon_\beta - \varepsilon_\alpha)^2 + \delta^2} \langle \alpha | x | \beta \rangle \langle \beta | v_y | \alpha \rangle. \quad (8)$$

We drop the a superscript for the Hall conductivity, but it will be implicit that we refer to the antisymmetric part unless otherwise mentioned. In order to demonstrate the cumulant technique, we first consider the limits of high temperatures when $(\varepsilon_\beta - \varepsilon_\alpha) < k_B T$ and, of small magnetic fields $k_B T \gg \hbar\omega_c$. This is the limit when the matrix elements are dominated by intraband scattering with weak B field. We expand the function

$$\begin{aligned} f(\varepsilon_\alpha) - f(\varepsilon_\beta) &\approx e^{-(\varepsilon_\alpha - \varepsilon_f)/k_B T} - e^{-(\varepsilon_\beta - \varepsilon_f)/k_B T} \\ &= \left(\frac{\varepsilon_\beta - \varepsilon_\alpha}{k_B T} \right) e^{-(\varepsilon_\alpha - \varepsilon_f)/k_B T} \\ &\quad - \frac{1}{2!} \left(\frac{\varepsilon_\beta - \varepsilon_\alpha}{k_B T} \right)^2 e^{-(\varepsilon_\alpha - \varepsilon_f)/k_B T} + \ldots \end{aligned} \quad (9)$$

and substitute Eq. (9) and Eq. (7) into Eq. (6). This gives us a cumulant expansion of the Hall conductivity in powers of $\frac{1}{k_B T}$ where one can use the operator identity

$$(\varepsilon_\beta - \varepsilon_\alpha) A_{\alpha\beta} = \langle \alpha | [A, H] | \beta \rangle, \quad (10)$$

where $[A, H]$ is the commutator of A with the Hamiltonian H and $A_{\alpha\beta} = \langle \alpha | A | \beta \rangle$, to reduce and evaluate the terms generated by the expansion.

The high T expansion converges rapidly as soon as $k_B T$ is larger than the typical energy difference in the matrix elements. In particular, this is obviously true as soon as $k_B T$ exceeds the bandwidth. But, we can also have a more general expansion that may also be valid at low temperature though the justification in this case is more complicated. To obtain the more general cumulant expansion at any temperature, we observe that in Eq. (6) the dominant region is around $\varepsilon_\beta \approx \varepsilon_\alpha$. Thus, we carry out a complete Taylor expansion of $f(\varepsilon_\beta)$ around ε_α and obtain

$$f(\varepsilon_\alpha) - f(\varepsilon_\beta) = \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) (\varepsilon_\beta - \varepsilon_\alpha) + \ldots \quad (11)$$

When we substitute Eq. (11) [strictly equivalent to Eq. (9) at high T] into the Kubo formula [Eq. (8)], we obtain the first-order term

$$\begin{aligned} \sigma_{xy}^{\{1\}} &= -\frac{e^2}{\Omega} \lim_{\delta \to 0} \sum_{\alpha,\beta} \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \frac{(\varepsilon_\beta - \varepsilon_\alpha)^2 + \delta^2 - \delta^2}{(\varepsilon_\beta - \varepsilon_\alpha)^2 + \delta^2} \\ &\quad \times \langle \alpha | x | \beta \rangle \langle \beta | v_y | \alpha \rangle \\ &= -\frac{e^2}{\Omega} \left[\sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \langle \alpha | x v_y | \alpha \rangle - \lim_{\delta \to 0} \sum_{\alpha,\beta} \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \right. \end{aligned}$$

$$\times \frac{\delta^2}{(\varepsilon_\beta - \varepsilon_\alpha)^2 + \delta^2} \langle \alpha | x | \beta \rangle \langle \beta | v_y | \alpha \rangle \Bigg]$$

$$= -\frac{e^2}{\Omega} \sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \langle \alpha | x v_y | \alpha \rangle, \tag{12}$$

where we have used the identity $\lim_{\delta \to 0} \frac{\delta^2}{(\varepsilon_\beta - \varepsilon_\alpha)^2 + \delta^2} = \pi \delta(\varepsilon_\alpha - \varepsilon_\beta) \lim_{\delta \to 0} \delta = 0$.

Apart from this first-order term [Eq. (12)], we obtain an infinite series of higher order cumulant which can be generated via Eq. (10). These higher order cumulants represent correction to the first-order result. The first-order term will turn out to have very simple physical interpretation. The corrections generated by the higher order cumulants resulting from Eq. (11), and when evaluated at low temperatures, must be studied in detail as it is not immediately self-evident that they represent lower order corrections. This is done later in the paper, but at this stage, it is already possible to note that the corrections represent, via Heisenberg's equation of motion, and from Eq. (10), higher and higher time derivatives of the velocity operator. When using the effective-mass Hamiltonian, one can show that beyond the second-order cumulant, the corrections which are linear in the cyclotron frequency or magnetization all scale with the disorder potential and spin-orbit coupling, and therefore are lower order corrections. From the mathematical structure and to keep the terms linear in B_z, it is thus essential therefore only to keep the second-order cumulant at low temperatures. This will now be shown step by step as we proceed with the analysis of the various contributions to the Hall current. Let us also remember that in Eq. (12), $|\alpha\rangle$ is a magnetic-field-dependent exact eigenstate.

In order to rewrite the matrix element in a simple way, we show in Appendix A that we can write the derivative of the eigenvalue with respect to the magnetic field as

$$\frac{\partial \varepsilon_\alpha(B_z)}{\partial B_z} = -e \langle \alpha | x v_y | \alpha \rangle - \frac{g \mu_B}{2} \langle \alpha | \sum_i \sigma_z^i | \alpha \rangle, \tag{13}$$

where v_y does not contain the spin-orbit term [Eq. (5)] contrary to v_y in Eq. (12). The spin-orbit term of v_y will be treated later in Sec. V C.

The Hall conductivity only comes from the first term on the right-hand side of Eq. (13), the orbital term. The way to handle this is to introduce initially two different magnetic fields, one acting on the orbital part B_{orb} and one giving the Zeeman energies B_z. We can thus rewrite Eq. (12) as

$$\sigma_{xy}^{\{1\}} = \frac{e^2}{\Omega} \frac{1}{e} \sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \frac{\partial \varepsilon_\alpha}{\partial B_{\text{orb}}}. \tag{14}$$

Only the derivative of the energy eigenvalues with respect to the orbital field B_{orb} gives the Hall conductivity. An advantage of this representation is that the spin-orbit energy can now be treated in first-order perturbation theory, as we shall see later.

The way disorder should be treated for numerical calculations is as follows. The Hall conductivity as given by the general conductivity formula [Eq. (1)] or the formulae obtained using the cumulant approach [Eq. (14) and equations thereafter] is evaluated for a specific configuration of disorder. The wave functions $|\alpha\rangle$ and energies ε_α are therefore the exact corresponding eigenstates and eigenvalues. Now the procedure is repeated for all the possible configurations and averaged with the appropriate weighting factors. Before examining the higher order terms (Appendixes B and C), let us understand the significance of this result and compare with other well-known approaches.

IV. COMPARISON TO OTHER THEORIES

A. Streda result

We can rewrite Eq. (14) in the form

$$\sigma_{xy}^{\{1\}} = -e \left[\frac{\partial}{\partial B_z} \int_{-\infty}^{E} \rho(E', B_z) dE' \right]_{E=\varepsilon_f}, \tag{15}$$

where $\rho(E', B_z)$ is the density of states (DOS) with magnetic field.

This expression is the quantum term of the Streda formula.[18] This author reduced the Kubo formula to two terms called σ_{xy}^{I} and σ_{xy}^{II} [Eq. (15) above]. The antisymmetric part of the other term, σ_{xy}^{Ia}, should be contained in the remainder of the cumulant expansion. Our term [Eq. (15)] differs from Streda's[18] by a minus sign. The sign problem can be traced in Ref. 18 to one transformation [Eq. 11 of Ref. 18] where there should be a minus sign on the right-hand side.

B. Classical limit

If, instead of Eq. (7), we write as in Ref. 19

$$\langle \alpha | x | \beta \rangle = -i\hbar \frac{\langle \alpha | v_x | \beta \rangle}{\varepsilon_\alpha - \varepsilon_\beta + \frac{i\hbar}{\tau}}, \tag{16}$$

where τ is a lifetime, we can reduce the first term of the cumulant expansion to the classical result. By adding a lifetime, we effectively assume that the electrons are subject to resistive scattering processes. Using the effective-mass Hamiltonian for the periodic part of the Hamiltonian also allows us to replace m with the effective m^*. The finite lifetime then represents all the scattering processes that break the translational invariance of a Bloch electron with effective mass m^*. This includes disorder and electron-phonon scattering treated in the Born approximation. When working in this approximation, there is no configuration average to be done anymore.

We consider only the term linear in B_z in Eq. (8), the one involving the $eB_z x/m$ part of v_y [see Eq. (5)], which gives the diagonal mass tensor term. Then, Eq. (8) together with Eq. (16) yields the cumulant,

$$\sigma_{xy}^{\{1\}} = \frac{e^2}{\Omega m} (eB_z) \sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \langle \alpha | x^2 | \alpha \rangle$$

$$\approx \frac{e^2}{\Omega m^*} (eB_z) \sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \sum_\beta \tau^2 |\langle \alpha | v_x | \beta \rangle|^2. \tag{17}$$

This leads to the very well known classical result

$$\sigma_{xy}^{\{1\}} = \frac{Ne^2\tau}{m^*}\frac{eB_z\tau}{m^*} = \sigma_{xx}\frac{eB_z\tau}{m^*}, \tag{18}$$

where $N=\rho(\varepsilon_f)m^*v_f^2$.

V. INFLUENCE OF THE SPIN-ORBIT COUPLING

Spin-orbit coupling introduces a number of contributions to the Hall current. New terms arise, due to the spin-dependent velocities from Eqs. (4) and (5) and from the effect of the spin-orbit interaction on the energy levels. We examine this last effect first. Also, as pointed out in Sec. IV B, for the class of problems where we have a periodic system+impurities, wherever the bare mass appears in the following text (except when it comes from the spin-orbit Hamiltonian), we can replace bare mass with the effective mass (m^*) and drop the periodic part of the Hamiltonian in the remaining analysis.

A. Effect of the spin-orbit interaction on the Hall current from the changes in the energy levels and wave functions

Consider the first-order cumulant result with the zero-order velocity operators (including the $-\frac{eB_zx}{m}$ term). The advantage of the cumulant expansion is that it allows us to analyze a very complex phenomenon, the effect of the spin-orbit coupling on the Hall conductivity, via the eigenstates. We write, to first order in perturbation theory,

$$\varepsilon_\alpha(B_z) = \varepsilon_\alpha^0(B_z) + \langle\alpha|H_{\rm so}|\alpha\rangle. \tag{19}$$

The spin-orbit Hamiltonian being dependent upon the Pauli's matrices, we should remember that the state $|\alpha\rangle$ must now be a spinor (two components vector). The action of taking the bracket will leave a scalar and, as we use the same state at this order, only the z component survives,

$$\frac{\partial\varepsilon_\alpha(B_z)}{\partial B_{\rm orb}} = -e\langle\alpha|xv_y|\alpha\rangle = \frac{\partial\varepsilon_\alpha^0(B_z)}{\partial B_{\rm orb}} + \sigma_z^\alpha\frac{\partial}{\partial B_{\rm orb}}\langle\alpha|\sum_i \lambda_i l_{i,z}|\alpha\rangle, \tag{20}$$

where

$$\lambda_n = \frac{\hbar eZ_n}{4m^2c^2(4\pi\varepsilon\varepsilon_0)|\mathbf{r}-\mathbf{R}_n|^3}, \tag{21}$$

$\sigma_z^\alpha \equiv \langle\alpha|\sigma_z|\alpha\rangle$, and $l_{i,z}$ is the z component of the orbital angular momentum at site i.

The first term in Eq. (20) is the one we examined above in the classical limit and is intuitively very attractive. The magnitude of the Hall current per eigenstate is related to the sensitivity of the energy level to an external magnetic field. Its sign depends on whether the magnetic field increases or decreases the energy of the eigenstate. In particular, it also follows that the contribution of a localized state is negligible. In reality, localized states should actually give exactly zero. It trivially follows that the zero is recovered only after summing the remaining contributions in the cumulant series. Keeping only the first cumulant does not give the exact result when the level in question is a localized level, with discrete energy levels. But the first-order result is close to zero, and therefore can be said to represent a good approximation. See Appendixes B and C for the analysis of the higher order cumulants.

For delocalized states, there is more information in the first term of Eq. (20). Normally, for weak scattering, when ε_f is near the top of the band, we have the hole sign, because the magnetic field can only lower the energy near the top of the band. Near the bottom of a band, we have the electron sign because the magnetic field confines the carrier and raises the energy of the electrons. This rule is also true for disordered eigenstates. The effect of the magnetic field on the energy levels can be evaluated in second order perturbation theory in the presence of disorder.

The anomalous term can now be studied by going to first order perturbation theory in B-field with exact eigenstates. The term $-\mu_B\mathbf{L}\cdot\mathbf{B}_{\rm orb}$ generates

$$|\alpha\rangle = |\alpha_0\rangle + \sum_{\beta_0}\langle\beta_0|-\mu_B\mathbf{L}\cdot\mathbf{B}_{\rm orb}|\alpha_0\rangle\frac{|\beta_0\rangle}{\varepsilon_{\alpha_0}-\varepsilon_{\beta_0}}. \tag{22}$$

Substituting Eq. (22) in the second term of Eq. (20), keeping only the z components, we obtain terms in the energy which involve a factor of the type

$$\Delta g_\alpha^{zz} = \sum_\beta \frac{\sum_i \lambda_i l_{\alpha\beta}^{i,z}\sum_j l_{\beta\alpha}^{j,z}}{\varepsilon_\alpha-\varepsilon_\beta}. \tag{23}$$

The sign of the anomalous process depends upon the sign of a quantity which is closely related to the electron g shift and which itself can be electronlike or holelike. Thus the first-order cumulant can be written as

$$\sigma_{xy}^{\{1\}} = \sigma_{xy}^{\{1\}(n)} + \frac{e^2}{\Omega}\frac{1}{|e|}\sum_\alpha\left(-\frac{\partial f(\varepsilon_\alpha)}{\partial\varepsilon_\alpha}\right)\mu_B\sigma_z^\alpha\Delta g_\alpha^{zz}. \tag{24}$$

Let us compare the relative magnitude of the two terms of Eq. (24). At very low T $[-\frac{\partial f}{\partial\varepsilon}=\delta(\varepsilon-\varepsilon_f)]$ we can rewrite the Hall conductivity [Eq. (24)] and the current is

$$J_y = \sigma_{xy}^{\{1\}(n)}F_x + e^2\rho(\varepsilon_f)\frac{\hbar}{2m}\langle\sigma_z\rangle_{\varepsilon_f}\Delta g^{zz}(\varepsilon_f)F_x. \tag{25}$$

With $\rho(\varepsilon_f)\sim 10^{45}/\mathrm{m}^3$ J and $F_x=10^4$ V/m we have for the anomalous contribution $J_y^{\rm an}=10^3\Delta g\langle\sigma_z\rangle F_x$ A/m^2, which is the same order of magnitude as the normal process with $N=10^{26}/\mathrm{m}^3$, giving the normal Hall current $J_y^n=10^6$ A/m^2 or $10^5(\hbar\omega_c\tau)F_x$ A/m^2 where $\hbar\omega_c\tau\sim 10^{-3}$. In principle, the normal and anomalous terms can have opposite signs. The results of Eqs. (24) and (25) are very elegant, and by writing $\Delta g=2-g^*$ we have essentially recovered the Fermi level version of the Chazalviel[20] and Nozières and Lewiner[21] result. Chazalviel[20] computes the single carrier wave packet motion in an electric field without using the Kubo linear-response formalism. We also note that in this form, the anomalous term apparently has no dependence on the relaxation time. Finally, and most importantly, it can also be related to the Karplus and Luttinger[10] result in the limit of high temperature when the Karplus and Luttinger energy gap between the Kane-Luttinger subbands is taken as $\Delta\sim k_BT$ and $\Delta g\sim\frac{\lambda_{\rm so}}{\Delta}$.

In the Karplus-Lutttinger[10] Bloch wave-function formalism, the anomalous Hall effect, even though it is intrinsic, is a Fermi level property at low temperatures. In Ref. 10, the spin-orbit interaction is treated in first-order perturbation theory using Bloch functions. We have also used first-order perturbation, but the Fermi level property, here, is a result of keeping the first-order cumulant. An interesting point is that Karplus and Luttinger did not use the Kubo formula. They arrived at a similar expression, except that the matrix elements are always interband matrix elements. The reason is that their starting point is the Bloch function, so that in the absence of an explicit treatment of disorder scattering, only interband matrix elements are left when the matrix elements of position and momentum operators are considered. Note that one way of calculating the g shift, when we have weak disorder, is to use the Kohn-Luttinger wave functions. One can compute the g shift $\Delta g_{\mathbf{k},\alpha}$ in the exact band states.[22]

In summary, in this section, we have derived a result which can be related to the Karplus-Luttinger intrinsic AHE (Ref. 10) and we have made contact with the Chazalviel,[20] Nozières and Lewiner,[21] and Sinova[8,9] results using a simple and unified formalism.

B. Side-jump Hall current

The velocity terms generated by the spin-orbit coupling in Eqs. (4) and (5) have been shown by Wölfle and Muttalib[23] to give rise to a term of a form called the "side-jump Hall effect" in the linear-response Kubo formula, which is

$$J_y = -Ne^2\langle\sigma_z\rangle\frac{\hbar}{4m^2c^2}F_x. \tag{26}$$

However, this result is obtained only after a complex diagrammatic sum of potential scattering events. A similar result was derived previously and more simply from the same velocity term by Lyo and Holstein[24] using scattering theory and originally by Berger.[25] We shall now rederive a contribution to the Hall current which has the same structure but is more closely related to the Rashba effect.[26] It is derived using linear response with the external-field-induced spin-orbit term in the Hamiltonian below. The linear-response analysis will show us how the presence of the lattice goes to modify the Hall conductivity even when we have disorder. Treated in linear response, the Rashba coupling term will give a side-jump-like Hall current. We do not use the cumulant approach or effective-mass approximation because it is easy to treat this term exactly.

Consider again that part of the spin-orbit coupling which is itself dependent on the applied external field F_x [see Eq. (3)]. This term is a contribution to the total energy which directly depends on the external field. As part of the Hamiltonian, this term creates a departure from equilibrium, and must therefore be treated on the same footing as the usual electric potential eF_xx. We therefore start from first principles, with the density matrix. When we take all such terms in the Hamiltonian as the perturbation H_{pert}, the change in the density matrix is given by

$$\Delta D_{\alpha\beta} = \langle\alpha|H_{\text{pert}}|\beta\rangle\frac{f(\varepsilon_\alpha)-f(\varepsilon_\beta)}{\varepsilon_\beta-\varepsilon_\alpha}, \tag{27}$$

where

$$H_{\text{pert}} = -eF_xx + \frac{\hbar}{4m^2c^2}[\nabla V_{\text{ext}}(\mathbf{r})\times\mathbf{p}]\cdot\boldsymbol{\sigma}, \tag{28}$$

with $V_{\text{ext}}(\mathbf{r})=-eF_xx$.

The second term of Eq. (28) involves the external applied electric field, and one can use linear response and ask: what Hall current does it produce in the presence of disorder? We can evaluate the thermally averaged velocities in the usual way. We consider the external-field-independent eigenstates, including disorder and the Zeeman splitting. These states can therefore be picked to have either spin-up or spin-down eigenstates in a chosen direction. Thus the y current produced by the second term from Eqs. (27) and (28) is given by (spin projection in z direction and keeping only the y-momentum term)

$$J_y = -\frac{e^2}{\Omega}F_x\sum_{\alpha,\beta}\langle\beta|v_y|\alpha\rangle\langle\alpha|\frac{\hbar}{4m^2c^2}p_y|\beta\rangle\frac{f(\varepsilon_\alpha)-f(\varepsilon_\beta)}{\varepsilon_\beta-\varepsilon_\alpha}\sigma_z^\alpha. \tag{29}$$

If we now use $mv_y=p_y$ and the sum rule (to be discussed in detail in Sec. V D)

$$\frac{1}{2m_\alpha} = \sum_\beta\frac{|\langle\alpha|v_\mu|\beta\rangle|^2}{\varepsilon_\beta-\varepsilon_\alpha}, \tag{30}$$

where $\mu=x$ or y, applied to the y operator, we have the very simple and elegant result

$$J_y = -\frac{e^2}{\Omega}\left(\frac{\hbar}{4m^2c^2}\right)F_x\sum_\alpha f(\varepsilon_\alpha)\frac{m}{m_\alpha}\sigma_z^\alpha, \tag{31}$$

which apparently only depends upon the effective mass. This is true as long as Eq. (30) can be used to define effective mass, i.e., if all incoherence is neglected. The interpretation of this sum rule in the Kubo formula context is not trivial. If we include the entire infinite spectrum in the sum of Eq. (30), then we have the trivial result $m_\alpha=m$, and we obtain a contribution which looks exactly like the Lyo and Holstein[24] and Berger[25] side-jump Hall current but is clearly based on a different logic (Rashba term). Note that in our theory, the spin polarization is now summed over the entire band and is not just the Fermi level spin.

In the framework of a finite band model, one can interpret Eq. (30) as the *effective mass*. The spin-orbit terms associated with the lattice can be included in Eq. (30). In the weak disorder limit, one could compute Eq. (30) using the Kane method. One can see that, in the Kane model, Eq. (30) is indeed the effective mass. The effective-mass correction in Eq. (31) can, then in principle, increase the current up to 2 orders of magnitude (InSb for example). But the side jump

[Eq. (31)] is, even with effective mass, for extended states, much smaller than the Karplus-Luttinger contribution. We have $J_y = 10^{-29}\langle\sigma_z\rangle NF_x$ A/m in two dimensions. In three dimensions, with the same numbers we have $10^{-3}F_x\langle\sigma_z\rangle(m/m^*)$ A/m^2 compared to $10^3F_x\langle\sigma_z\rangle\Delta g$ A/m^2 from Eq. (25). Note that the huge Bloch function enhancement evaluated in various forms by Chazalviel,[20] Berger,[25] and Fivaz[27] and which make the side-jump term important does not appear in Eq. (31). Chazalviel, for example, used the Heisenberg commutator for v_y and Kane wave functions to derive the Hall velocity and then derives a similar expression for the Hall current using Drude theory. We have derived an expression which is similar in structure to what is called the side-jump Hall current in the literature. The derivation we have used is however not the same as that of Wölfle and Mutallib,[23] Lyo and Holstein,[24] and Berger.[25] In our formula, the spin-orbit coupling can be enhanced by the lattice, but the effect is relatively small, and directly related to the effective-mass lowering [see Eq. (31)]. The so-called side-jump theories[24,25] in which the mechanism is due to the spin-orbit scattering induced sideways jump at impurity potentials, and the associated large enhancements caused by Bloch functions, have not been recovered using the present Kubo formula method. To complete the analysis of Eq. (31) we need a discussion of the sum rule of Eq. (30) and we will defer this to Sec. V D because a similar term is encountered in Sec. V C.

C. Effect of the spin-dependent velocity on the Hall current: The terms which are due to the internal potentials

Let us consider the contributions to the Hall current which results from including the contribution of the remaining spin-orbit velocity terms [Eqs. (4) and (5)] in the Kubo formula. These now involve the lattice potentials as sources of *velocity*. For Coulomb potentials, we have

$$\begin{aligned}\langle\alpha|v_x|\beta\rangle\langle\beta|v_y|\alpha\rangle &\to \langle\alpha|\sum_n \frac{eZ_n\hbar}{4m^2c^2(4\pi\varepsilon\varepsilon_0)}\frac{y-Y_n}{|\mathbf{r}-\mathbf{R}_n|^3}\sigma_z|\beta\rangle\langle\beta|v_y|\alpha\rangle \\ &\approx \sum_\eta \langle\alpha|\sum_n \varsigma\frac{1}{|\mathbf{r}-\mathbf{R}_n|^3}|\eta\rangle\langle\eta|y\sigma_z|\beta\rangle\langle\beta|v_y|\alpha\rangle \\ &\approx \langle\alpha|\sum_n \varsigma\frac{1}{|\mathbf{r}-\mathbf{R}_n|^3}|\alpha\rangle\langle\alpha|y\sigma_z|\beta\rangle\langle\beta|v_y|\alpha\rangle, \end{aligned} \tag{32}$$

where $\varsigma \equiv \frac{eZ_n\hbar}{4m^2c^2(4\pi\varepsilon\varepsilon_0)}$.

Two approximations were made here. Y_n takes alternatively positive and negative values and the term involving it would be zero if $\mathbf{r}$ were not present. But, even if we include $\mathbf{r}$, it will always give a smaller contribution compared to the first and we therefore neglect it. Second, $\frac{1}{|\mathbf{r}-\mathbf{R}_n|^3}$ is local and therefore cannot couple different sites but, in tight-binding for example, could couple different orbitals at the same sites. We considered that the main contribution comes from the matrix element taken between the same eigenstate and this is why we considered only the $|\eta\rangle = |\alpha\rangle$ term in the previous equation. One has to note that the original integrals in Eq. (32) are convergent but when one breaks them up, then the integral $\langle\alpha|\frac{1}{|\mathbf{r}-\mathbf{R}_n|^3}|\alpha\rangle$ is strictly speaking not convergent because one has taken one position term $y-Y_n$ out of it. The underlying assumption which allows us to make this decoupling is that the orbit radius is never allowed to be smaller than the effective atomic orbit of the valence state so that the cubic singularity does not occur.

This term gives rise to a contribution in the first cumulant. To obtain it, we start from Eq. (6) instead of Eq. (8). We substitute Eq. (32) in Eq. (6) and use Eq. (7) to transform $\langle\alpha|y|\beta\rangle$ to $\langle\alpha|v_y|\beta\rangle$. The result is

$$\sigma_{xy} = \frac{\hbar^2e^2}{\Omega}\sum_\alpha\left(-\frac{\partial f(\varepsilon_\alpha)}{\partial\varepsilon_\alpha}\right)\Gamma_\alpha\frac{\partial}{\partial\varepsilon_\alpha}\left(\frac{1}{2m_\alpha}\right)\sigma_z^\alpha, \tag{33}$$

where

$$\Gamma_\alpha = \langle\alpha|\sum_n \frac{eZ_n\hbar}{4m^2c^2(4\pi\varepsilon\varepsilon_0)}\frac{1}{|\mathbf{r}-\mathbf{R}_n|^3}|\alpha\rangle. \tag{34}$$

Again we have used Eq. (30). Following Datta,[19] we will assume that, in a crystal, the sum rule of Eq. (30) is indeed the effective mass. The sum rule [Eq. (30)] will be discussed in detail in Sec. V D. Note that if we use Eq. (16) with broadening, then an exact result can be obtained by taking the derivative with respect to the broadening ($\frac{i\hbar}{\tau}$) instead of the energy in Eq. (33). This relation will be used to derive Eq. (36) for the strong scattering limit.

D. Problem of the sum rule of Eq. (30)

There is a certain arbitrariness in the use of the sum rule of Eq. (30) which we should clarify. We note that if we use the standard transformation of Eq. (7), then it follows, in principle, that when we sum over the entire real spectrum of

the Hamiltonian, we obtain the free electron mass on the left-hand side. This is simply a consequence of the trivial identity $p_x x - x p_x = -i\hbar$ with $p_x = m v_x$.

However, if we follow Datta,[19] and use the linear-response density matrix to compute the *acceleration* of a particle in an electric field, then we obtain the effective mass, in the sense of Newton's law, as given by Eq. (30). Consequently, this gives the absurd result that the accelerating particle is always free, irrespective of what its initial state is. In effect, the trivial result signifies that if we wait long enough, then even a strongly bound electron will eventually be free in an electric field. That is, this result (mass is free mass) would represent the very long time behavior, when the history of the particle is irrelevant, and its acceleration in a constant field is truly dominated by what happens when it has reached its final free state. Datta[19] concluded that he should use a finite band in the evaluation of Eq. (30), and then the left-hand side is indeed the effective mass in the sense of the tight-binding band structure, for example.

The solution of this problem, in general, seems to be that, in a transport situation, where electrons are injected at one end and absorbed at the other, the sum can only run over that part of the spectrum which is accessible to the carrier in its lifetime, i.e., for which $\hbar \Sigma_\beta |\langle \alpha | v_x | \beta \rangle|^2 \delta(\varepsilon_\beta - \varepsilon_\alpha - \hbar\omega)$ is finite. In weak scattering, the particle *lives* in energy levels near the Fermi level which have an effective mass, because it relaxes and emits energy to the lattice. In a strong scattering situation, the kinetic energy of the carrier can be of the same order as the scattering energy uncertainty $\frac{\hbar}{\tau}$. So here, we can relate the sum on the right-hand side of Eq. (30) directly to the quantum diffusivity D_α^0 (see Appendix D). We propose therefore, in the strong scattering limit, where Bloch's theorem does not apply, to define the effective mass, in the sum rule of Eq. (30) by the relation

$$\frac{1}{m_\alpha} = 2 c_\alpha \frac{D_\alpha^0}{\hbar}, \tag{35}$$

where c_α is a constant $\sim \langle \frac{(\varepsilon_\alpha - \varepsilon_\beta)\tau_\beta}{\hbar} \rangle_\beta$ which carries a sign and is averaged over the band (see Appendix D for a formal representation). It is of order 1 when the energy ε_α is near the bottom of the band or in a rapidly changing region of the DOS. We shall henceforth absorb this constant in the definition of an effective diffusivity, D_α. Equation (35) is exact (Appendix D).

In the random-phase limit, c_α is a relatively weak function of energy and can be treated as a constant. The unit of time is the scattering time. Acceleration with strong disorder is therefore drift velocity divided by scattering time. With the same definition, in the semiclassical limit we therefore have

$$\frac{\partial}{\partial \varepsilon_\alpha}\left(\frac{1}{2m_\alpha}\right) \sim -\frac{D_\alpha \tau}{\hbar^2}. \tag{36}$$

It follows that a localized initial state $|\alpha\rangle$ has no acceleration without phonons; i.e., at zero temperature, its effective mass in the sense of Eq. (30) is infinite. If one evaluates what diffusivity one needs to reproduce the electron mass using Eq. (35), one has $D \sim 1$ cm^2 s^{-1} which is not a small value in a disordered system. We now have a way of interpreting terms involving the effective mass m_α. For example, for strong disorder, expressions of the form $\Sigma_\alpha \frac{1}{m_\alpha} \frac{\partial^2 f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha^2}$ as encountered in Appendix B for the second-order cumulant can be written as

$$-\int \frac{D(\varepsilon)\rho(\varepsilon)}{\hbar} \frac{\partial}{\partial \varepsilon}[\delta(\varepsilon - \varepsilon_f)] = \frac{1}{\hbar} \left. \frac{\partial \sigma(\varepsilon)}{\partial \varepsilon} \right|_{\varepsilon_f}, \tag{37}$$

with $\sigma(\varepsilon) = \rho(\varepsilon) D(\varepsilon)$, where $\sigma(\varepsilon)$ is the energy-dependent conductivity with universal scaling properties near mobility edges.[2]

E. Resistance dependence of the anomalous term

Experimentally one is normally interested in the resistance dependence of the AHE and this is the criterion used to discriminate between possible models. For example, the observed resistance independence of the experimental anomalous Hall conductivity in dilute magnetic semiconductors is taken as proof that the AHE is intrinsic. However, it follows from the present analysis that the apparent scattering-time independence of the anomalous Hall conductivity, as derived here in the form of Eq. (24), and by Karplus and Luttinger,[10] and by Jungwirth *et al.*,[7] constitutes a very special limit. The scattering-time independence of the g shift [Eq. (23)] follows when the matrix elements are dominated by interband processes where the energy differences are $\gg \frac{\hbar}{\tau}$, (see Ref. 10 for example). But this implies that the basic band structure is Bloch-type, with well-defined semiconductor bands. This is obviously not always the case, and there can be many situations where disorder and band crossings wash out the Kane subband gaps, and give rise to arbitrarily small energy denominators in which the lifetime directly enters the AHE as well. Thus, the g shift [Eq. (23)] can very well involve the conductivity scattering time.

The assumption of Bloch functions is made by Jungwirth *et al.*,[7] and the resistance independence of their result also rests on the existence of well-defined subbands, and the sums run over all occupied levels. Thus we conclude that for strong spin-orbit scattering and weak disorder scattering, the anomalous Hall conductivity will not depend strongly on the resistance, and this then accounts for some of the experimental observations on the AHE.[28] In contrast, for weak spin-orbit coupling and strong disorder, the anomalous Hall conductivity can vary as the scattering time τ^n, $0 < n < 2$. The Bloch matrix elements[20] are not appropriate in the strong disorder limit, and the present approach, though perturbational, is more appropriate.

VI. DISCUSSION

The Kubo formula was expanded in a cumulant expansion which converges quickly at high temperatures. One can also use the series expansion at lower temperatures, provided one examines the higher order contributions for convergence. Working with the effective-mass Hamiltonian, and keeping only the contributions to first order in magnetic field, allows one to stop the expansion after the second-order cumulant.

The higher order cumulant contributions are examined in Appendixes B and C.

The second-order cumulant derived in Appendix B also has a very interesting structure. It gives terms which scale as the derivative of the density of states at the Fermi level, both for the normal and the anomalous contributions. Previously, Bush and Guentherodt[29] had suggested that the sign of the Hall effect in disordered materials scales as the sign of the derivative of the density of states at the Fermi level (electronlike for increasing and vice versa). In the present theory, we have actually successfully identified the quantities that determine the sign of the Hall effect. Indeed, we have made the very interesting observation that the sign of the normal term, as expressed in the first term, is related to the derivative of the energy with respect to B field. In the second cumulant contribution, the sign is related to the derivative of the density of states at the Fermi level (at low T). Interestingly the two terms have opposite trends because an increasing density of states actually gives the hole sign, and not as one would intuitively expect and as Bush and Guentherodt[29] suggest, the electron sign. This is a truly remarkable result when applied to the strong scattering limit, because normally in weak scattering, the first term is dominant and that gives the intuitive result from Ref. 29. But in very strong scattering, for example in the amorphous limit, and in the region near the mobility edge, it may happen that the second term dominates. If this happens, then we have a sign anomaly because the increasing density of states at ε_f for n doping gives the hole sign and vice versa. This is exactly what is observed in the band edge of doped amorphous silicon.[30] This observation merits a more detailed investigation which goes beyond the scope of this paper.

In disordered systems, one can use the coherent potential approximation (CPA) (Ref. 31) to describe the disordered band structure for example[32,33] and get explicit results for the sign of the Hall effect. It turns out that there is no simple rule for the sign of the Hall coefficient in CPA either, but at the band edges, we do indeed have the same behavior as predicted here. This has been discussed in detail in Refs. 1 and 33. In Ref. 1, the case of an impurity band is also considered.

The disadvantage of the cumulant method is that the low-temperature limit has to be examined with care for convergence. In situations with Bloch symmetry where the dominant matrix elements are on the same energy shell, this is no problem. But in general, with disorder, there is, in the present formulation, in lowest order in magnetic field B_z, unfortunately, an infinite number of terms. This seems a big problem at first, but then it resolves itself. The resolution of the difficulty is most obvious when we apply the first-order term in the limit that the states are localized at ε_f. We obtain a Hall conductivity which is small, but nonzero. This small contribution must be canceled by the remaining linear terms in the series. Nevertheless the approximation is still good because it gives a negligible contribution to the Hall current, knowing that the exact result should be strictly zero. Equations (B3) and (C1) have terms which scale as B_z. They are generally smaller than the first-order contributions we derived in Sec. V. However, they involve higher derivatives of the Fermi function and can be dangerous to handle at low T. One may infer that, if the system has a density of states and scattering times which are only weak functions of energy at the Fermi level, these higher order cumulant terms are negligible. If the density of states is a strong function of energy, the expansion will not converge so easily. Indeed, near the band edges, there will be mobility edges and localized levels which must give rise to a null result without phonons, but the null result must be arrived at by cancellation of many, albeit, small contributions.

At high temperatures, the higher order terms cause no problem and can be neglected. At any temperature we may conclude that a very good approximation is obtained by keeping only the first and second cumulants, the spin-orbit velocity contribution, and the external-field-induced spin-orbit term. The final approximate formula for the Hall conductivity thus becomes

$$\begin{aligned}\sigma_{xy} = {} & \frac{e^2}{\Omega}\sum_{\alpha}\left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha}\right)\left[\frac{-1}{|e|}\frac{\partial \varepsilon_\alpha}{\partial B_{\mathrm{orb}}} + \frac{\hbar}{2m|e|}\sigma_z^\alpha \Delta g_\alpha^{zz}\right] \\ & + \frac{e^2}{2!\,\Omega}\sum_{\alpha}\frac{\hbar}{2m_\alpha}\frac{\partial^2 f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha^2}\left[\frac{\hbar|e|B_z}{m} + \hbar\Gamma_\alpha\sigma_z^\alpha\right] \\ & - \frac{e^2}{\Omega}\frac{\hbar}{4m^2c^2}\sum_{\alpha} f(\varepsilon_\alpha)\frac{m}{m_\alpha}\sigma_z^\alpha \\ & + \frac{e^2\hbar^2}{\Omega}\sum_{\alpha}\left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha}\right)\Gamma_\alpha\frac{\partial}{\partial \varepsilon_\alpha}\left(\frac{1}{2m_\alpha}\right)\sigma_z. \end{aligned} \tag{38}$$

We may call Eq. (38) the weak-to-intermediate scattering Hall conductivity. When we evaluate the cumulants, one can use the effective-mass Hamiltonian so that the periodic potential is not part of the Hamiltonian with which the higher order commutators [Eq. (10)] are to be evaluated. This means that the only terms which contribute above the second-order cumulant, and which scale linearly with B_z, will be those which depend on the disorder and spin-orbit part of the Hamiltonian and this makes the approximation of only keeping up to second order very accurate. Each of the four additive terms of Eq. (38) will now be discussed and a simple interpretation given.

The first expression in the bracket of the first term has been discussed and is easy to interpret, but it is not completely trivial to see that it simply reduces to the classical result [Eq. (18)] if we used Eq. (16) or Bloch functions in the Kubo formula with a constant lifetime. In the pure quantum limit it gives the Streda result but it is actually the *normal* Hall effect. The way this term should be handled depends on the problem in question. In weak scattering it again gives Eq. (18), with a free electron mass. In the tight-binding representation, one can evaluate it using second-order perturbation theory in the magnetic-field-dependent term in the Hamiltonian.

The second part of the first bracket involves the g shift of the delocalized levels above the mobility edge. In comparison, the localized g shift is negligibly small. In the weak scattering limit, the g shift can be evaluated using the Kane-Luttinger wave functions.[10] This is given by Roth *et al.*[22] In an incoherent situation, a disordered system with no Bloch

symmetry and strong scattering, we should use $m_\alpha = \frac{\hbar}{D_\alpha}$ as the effective mass.

For the second term (see Appendix B), the one involving a second derivative of the Fermi function, one can, using the $m_\alpha = \frac{\hbar}{D_\alpha}$ approximation, obtain integral products of the type $\sigma(\varepsilon) = \rho(\varepsilon) D(\varepsilon)$ as shown in Eq. (37) which are energy-dependent conductivities, and which obey well-known universal scaling relations near the mobility edges.

The third term was shown to be due to the effect of the external-field-induced spin-orbit energy (Sec. V B). With $m_\alpha = m^*$, this term reduces to the same form as the so-called side-jump contribution. Our theory shows that it can be enhanced via a small effective mass and even extended to apply to strong scattering via Eq. (35).

Considering the second and third terms, we propose that, to a good approximation, in most situations where Bloch functions cannot be used, we may replace $m_\alpha = \frac{\hbar}{D_\alpha}$, once the sign has been determined via Eq. (30). This accounts for localized states if any are present, because D_α is zero.

The last term in Eq. (38) is due to the spin-orbit contribution to the velocity operators via the internal potentials from Eqs. (4) and (5). This also vanishes for localized states. Above the mobility edge, this contribution is comparable to the second expression in the bracket of the second term. This can be seen using the strong scattering correspondence [Eq. (36)]. The last term in Eq. (38) then reduces to a term which resembles the usual skew scattering term[12] provided we interpret the skew scattering rate for Coulomb spherical potentials as

$$\frac{1}{\tau_s^\alpha} = \frac{\hbar e}{4m^2c^2(4\pi\varepsilon\varepsilon_0)} \langle \alpha | \sum_n \frac{Z_n}{|\mathbf{r} - \mathbf{R}_n|^3} | \alpha \rangle. \tag{39}$$

Thus, for disordered systems, the last term of Eq. (38) can be written as

$$\sigma_{xy}^{\text{skew}} = -\frac{\hbar e^2}{\Omega m} \sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \frac{\tau}{\tau_s^\alpha} \frac{m D_\alpha}{\hbar} \sigma_z^\alpha. \tag{40}$$

Note that the sum runs over all the potentials in the lattice. So the skew scattering and g-shift terms include both extrinsic and intrinsic contributions. Note that very recently, Chudnovsky,[34] using a different approach, and for spin Hall effect, also obtained a term where all the potentials are included. In the form of Eq. (40), appropriate for disordered systems with no Bloch symmetry, the Bloch enhancement does not appear. If we neglect the host spin-orbit coupling and only include the impurities, then both g shift and skew term are, by definition, extrinsic contributions, and the mass in Eq. (39) is necessarily the effective mass. It is the effective-mass particle which generates the impurity spin-orbit magnetic field.

Let us now reexamine the question of the theoretical resistance dependence of the AHE. From the above analysis we note that this all has to do with the way we treat the matrix elements and at what stage we introduce incoherence and lifetime. This can already be seen in the first term, which can be treated as a quantum effect, as in Streda[18] or in the semiclassical limit. The same is true for the spin-orbit terms. Thus, if we keep to the notion of effective mass, we have the quantum result. If we use the transformation of Eq. (35), which involves the diffusivity, then we have the connection with conductivity. As an example, we can take the first two dominant terms in Eq. (38) and write them, using the definition of the g shift given by Roth *et al.*[22] (we assume Δ and E_g to still be defined), in the incoherent limit as

$$\sigma_{xy}^{(1)} = \sigma_{xx} \left(\frac{eB_z \tau}{m^*} \right) + \frac{e^2}{\Omega} \sum_\alpha \left(-\frac{\partial f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha} \right) \times \frac{\hbar}{m} \sigma_z^\alpha \left[-\left(m \frac{D_\alpha}{\hbar} - 1 \right) \frac{\Delta}{3E_g + 2\Delta} \right], \tag{41}$$

where we have used Eq. (35) in the g-shift term.

Now, we see that what was a pure quantum term with effective mass, has, in this limit, become a term which depends on the diffusivity. The AHE Hall coefficient can, it seems, change from a linear scaling with resistance in Eq. (41) to a resistance squared (relaxation time squared) behavior if $\frac{D_\alpha}{\hbar}$ becomes m_α^{-1}, as is observed in diluted magnetic semiconductors.[28] The experimentally observed $R_A \propto \rho_{xx}^2$ dependence implies, therefore, that the effective-mass concept remains valid in these systems.

We have focused our attention on the anomalous contributions, and how they undergo a transformation, in going from the weak scattering to the strong scattering limit. A similar change must occur for the first term of Eq. (41). Here too, we must replace the Drude σ_{xx} by

$$\sigma_{xx} = e^2 \int d\varepsilon \left(-\frac{\partial f(\varepsilon)}{\partial \varepsilon} \right) \rho(\varepsilon) D(\varepsilon) \tag{42}$$

and the band mobility term $\frac{e\tau B_z}{m^*}$ by

$$\frac{e\tau B_z}{m^*} = eB_z \frac{D}{\langle \varepsilon \rangle}, \tag{43}$$

at low temperatures. The notation in terms of diffusivity D is valid in the weak scattering limit too, but now one can see what happens as we approach the mobility edge using standard localization theory and $\sigma(\varepsilon) = \rho(\varepsilon) D(\varepsilon)$.

We have shown that the AHE Hall coefficient R_A can vary with resistance ρ_{xx}^n with $1 \le n \le 2$ depending on the degree of coherence. The essential point seems to be the way one treats the sum rule given by Eq. (30). This was already a problem for Datta[19] in his analysis of the Hall effect in his 1980 paper. Equation (30) may be treated as the well-known f-sum rule, but it is clearly not realistic to sum over an infinite excited state spectrum, without taking into account the finite lifetime of the states. The question then becomes: when is the sum on the right-hand side of Eq. (30) $\frac{1}{2m^*}$ and when is it closer to $\frac{D}{\hbar}$ which is the sum evaluated in the semiclassical limit, and also derivable from $\langle \alpha | x^2 | \alpha \rangle \sim D_\alpha \tau$? The interesting and important side of this last result is that it gives the correct null result for transport in localized states, and therefore must be the correct approximation near the mobility edges.

We have derived a side-jump-like contribution to the Hall current using linear response to the (Rashba) spin-orbit interaction[26] which is produced by motion in the external

potential. Our result is closely related to the Rashba current[26] and is in principle, exact, up to the evaluation of the sum rule [Eq. (30)]. The sum rule has to be treated with care in both the coherent and the strong scattering limit. In the incoherent limit, and especially near mobility edges, we have the relation $m_\alpha = \frac{\hbar}{D_\alpha}$ which then gives a very appealing result, namely

$$\sigma_{xy}^{\mathrm{sj}} = -\frac{e^2}{\Omega}\frac{\hbar}{4m^2c^2}\sum_\alpha f(\varepsilon_\alpha)\frac{mD_\alpha}{\hbar}\sigma_z^\alpha. \tag{44}$$

This gives a vanishing side-jump-like Hall effect (Rashba) contribution from those states in the localized part of the band.

Adding Eqs. (41) and (44) we have the strong scattering Hall conductivity limit. As in the QHE, every conducting state contributes. However, the full enhancement of the spin-orbit coupling (see Refs. 12 and 35 for example) does not appear in our expression. The most interesting aspect of the side-jump-like (Rashba) term [Eq. (44)] is that if the Fermi level is in a region of localized states, say near the top of the band, then this term dominates the Hall current since all other terms are Fermi level terms, and vanish at low temperatures. It would give us (using the free mass) $J_y \sim 10^{-29} N_{\mathrm{del}} \langle \sigma_z \rangle F_x$ A/m^2 (N_{del}=delocalized electron density) which is, of course, a small but *rigid* current, analogous to the quantum Hall current.

VII. CONCLUSION

The principal result of the cumulant method developed in this paper is given by Eq. (38). It is an expression for the Hall current which allows for disorder and spin-orbit coupling. For a typical doped semiconductor, we might expect the following contributions to the Hall effect:

(1) The normal process [first term of Eq. (41)].

(2) The intrinsic AHE side-jump-like process [Rashba term, Eq. (44)].

(3) The intrinsic AHE Karplus-Luttinger process [second term of Eq. (41)].

(4) The skew scattering process due to impurities [Eq. (40)]. Here, the basic spin-orbit coupling has the effective mass and not the free mass.

The skew term [Eq. (40)] is only relevant if both the enhancement and the basic spin-orbit coupling are large and the concentration of impurities high enough (see Ref. 12 for estimates). The side-jump term is only important if the effective mass is small or the Fermi level is in a region of localized states. In most situations with weak disorder, we expect the Karplus-Luttinger intrinsic term to dominate the AHE. In the limit of strong scattering, when the use of Bloch functions and uniform lifetimes is no longer valid, we may replace the effective mass in Eq. (38) using the concept of quantum diffusivity; i.e., put $m_\alpha^{-1} \to \frac{D_\alpha}{\hbar}$, and use derivatives, as explained in the text.

For magneto-optical Faraday angle measurements, for example, we need the frequency-dependent Hall conductivity. One can show that in weak scattering, where we use the effective mass for Eq. (30), we add a factor $\frac{1}{1+(\omega\tau)^2}$ to the four main Hall conductivity results listed above. In strong scattering, where we use the diffusivity to describe Eq. (30), $D(\varepsilon)$ is replaced by $D(\varepsilon, \omega)$.

ACKNOWLEDGMENTS

The authors thank Arthur Yelon for reading and commenting on the manuscript. L.-F.A. acknowledges financial support from A.-M. S. Tremblay during the writing of the paper.

APPENDIX A: DERIVATIVE OF THE HAMILTONIAN WITH RESPECT TO THE MAGNETIC FIELD

With $\mathbf{A}_i = (0, B_z x_i)$ and $p_y^0 = m v_y^0$, the kinetic part of the Hamiltonian is

$$T = \sum_i \left[\frac{m(v_i^0)^2}{2} - eB_z x_i v_y^{0,i} + \frac{e^2 B_z^2}{2m} x_i^2 \right]. \tag{A1}$$

Thus, with $H = T + V + H_{\mathrm{so}} - \frac{g\mu_B}{2} B_z \Sigma_i \sigma_z^i$, we can write, remembering that without the spin-orbit term in the velocity operator $v_y^i = v_y^{0,i} - \frac{eB_z}{m} x_i$,

$$\frac{\partial H}{\partial B_z} = \sum_i \left[-e x_i v_y^i - \frac{g\mu_B}{2} \sigma_z^i \right]. \tag{A2}$$

Thus,

$$\langle \alpha | \frac{\partial H}{\partial B_z} | \alpha \rangle = -e \langle \alpha | x v_y | \alpha \rangle - \frac{g\mu_B}{2} \langle \alpha | \sum_i \sigma_z^i | \alpha \rangle. \tag{A3}$$

We now want to transform the left-hand side of Eq. (A3) to obtain Eq. (13). To do so, we just apply the well-known Hellman-Feynman theorem and we obtain $\langle \alpha | \frac{\partial H}{\partial B_z} | \alpha \rangle = \frac{\partial \varepsilon_\alpha}{\partial B_z}$ and thus Eq. (13). Remember that v_y in the equation does not contain the spin-orbit term of Eq. (5).

APPENDIX B: THE SECOND-ORDER TERMS IN THE CUMULANT EXPANSION

Consider now the second-order contribution to the Hall conductivity. This can be written as

$$\sigma_{xy}^{\{2\}} = \frac{e^2}{2!\,\Omega} \sum_{\alpha,\beta} \langle \alpha | x | \beta \rangle \langle \beta | H v_y - v_y H | \alpha \rangle \frac{\partial^2 f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha^2}, \tag{B1}$$

where, as previously, $v_y = \frac{-i\hbar}{m}\frac{\partial}{\partial y} - \frac{eB_z x}{m}$, dropping the spin-orbit current operator. After evaluating the commutator we are left with three terms,

$$i\hbar \frac{eB_z}{m} v_x^0, \quad \frac{i\hbar}{m}\frac{\partial V}{\partial y}, \quad -\frac{i\hbar}{m} p_x \sum_n \lambda_n \sigma_z. \tag{B2}$$

Only the first and third term give significant contributions in this order. The first term is

$$\sigma_{xy,1}^{\{2\}} = \frac{e^2}{2!\,\Omega} \sum_{\alpha,\beta} \langle \alpha | x | \beta \rangle \langle \beta | i\hbar \frac{eB_z}{m} v_x^0 | \alpha \rangle \frac{\partial^2 f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha^2}$$

$$= \frac{e^2}{2!\,\Omega} \frac{\hbar^2 e B_z}{m} \sum_\alpha \left(-\frac{1}{2m_\alpha} \right) \frac{\partial^2 f(\varepsilon_\alpha)}{\partial \varepsilon_\alpha^2}, \tag{B3}$$

where we have used the sum rule of Eq. (30) restricted to the active energy band. For Bloch states restricted to the eight Kane bands with spin-orbit coupling, the right-hand side of Eq. (30) reproduces the eight-band $\mathbf{k}\cdot\mathbf{p}$ calculated effective masses. To conclude, we note that the cumulant expansion is a powerful method when the bandwidth is very narrow and $<k_BT$. At low temperatures, when we use $m_\alpha=m^*$, we obtain for Eq. (B3)

$$\sigma_{xy,1}^{\{2\}}=-\frac{e^2}{2!}\frac{\hbar}{m}\left(\frac{\hbar eB_z}{2m^*}\right)\left.\frac{\partial\rho(\varepsilon)}{\partial\varepsilon}\right|_{\varepsilon=\varepsilon_f}. \tag{B4}$$

Note that the cumulant expansion is not purely an expansion in powers of B_z. There is a term linear in B_z in almost every order. Using the effective-mass method, the linear term in B_z becomes an expansion in powers of the $V_{\text{disorder}}/\varepsilon_f$. The convergence is problematic near band edges where we have localization. We neglect the term given by the second term of Eq. (B2). The spin-orbit term, given by the third term of Eq. (B2), becomes

$$\sigma_{xy,3}^{\{2\}}=\frac{e^2\hbar^2}{2\,!\,\Omega}\sum_\alpha\frac{1}{2m_\alpha}\Gamma_\alpha\frac{\partial^2 f(\varepsilon_\alpha)}{\partial\varepsilon_\alpha^2}\sigma_z^\alpha. \tag{B5}$$

Again, at low T (neglecting other energy dependence under the integral), we can use

$$\sum_\alpha\frac{\partial^2 f(\varepsilon_\alpha)}{\partial\varepsilon_\alpha^2}=\left.\frac{\partial\rho}{\partial\varepsilon}\right|_{\varepsilon=\varepsilon_f}. \tag{B6}$$

If we assume that $\frac{\partial\rho}{\partial\varepsilon}|_{\varepsilon=\varepsilon_f}\sim\frac{\rho}{\varepsilon_f}$, again, we have the same structure as before in Eq. (25), this time with $\Delta g=\frac{\hbar\langle\Gamma_\alpha\rangle}{\varepsilon_f}$. With the numerator of order 10^{-4} eV, this term corresponds to an effective g shift of 10^{-4} which is therefore smaller than the first-order cumulant of its type. But, in general we remind the reader that the cumulant expansion, as is also true for the configurationally decoupled Kubo formula, is not accurate near the band edges for reasons of Anderson localization. When used in this region it can only apply above the mobility edge.

APPENDIX C: THIRD-ORDER CUMULANT

In the third-order cumulant, if we neglect terms of order B_z^2 and λ^2, we have only the terms

$$\sigma_{xy}^{\{3\}}=\frac{e^2}{3\,!\,\Omega}\sum_\alpha\frac{1}{m_\alpha}\frac{\partial^3 f(\varepsilon_\alpha)}{\partial\varepsilon_\alpha^3}\langle\alpha|x^2\nabla_y^2V(\mathbf{r})\times\left[\frac{\hbar eB_z}{m}+\sum_n\hbar\lambda_n(\mathbf{r})\langle\sigma_z\rangle\right]|\alpha\rangle\frac{\hbar}{m}. \tag{C1}$$

The terms linear in B_z and spin-orbit coupling always go in pairs, the spin-orbit term acting like an effective magnetic field. The higher order terms form an infinite series with products involving derivatives of the lattice potential. These terms are small at high temperature and renormalize the first-order linear terms in B_z and λ_n. In the effective-mass approach, the lattice potential no longer appears in the Hamiltonian so $V(\mathbf{r})$ appearing in Eq. (C1) is due to impurities or disorder.

APPENDIX D: EFFECTIVE MASS AND DIFFUSIVITY IN THE STRONG SCATTERING LIMIT

Since the excited states can decay and must have a finite lifetime, we can write the sum of Eq. (30) as

$$\frac{1}{m(\varepsilon_\alpha)}=2\int_{-W}^{W}d\varepsilon_\beta\rho(\varepsilon_\beta)\frac{\langle\alpha|v_x|\beta\rangle^2(\varepsilon_\beta-\varepsilon_\alpha)}{(\varepsilon_\beta-\varepsilon_\alpha)^2+\left(\frac{\hbar}{\tau_\beta}\right)^2}. \tag{D1}$$

The quantum diffusivity is defined as

$$D_\alpha=\hbar\sum_\beta|\langle\alpha|v_x|\beta\rangle|^2\delta(\varepsilon_\beta-\varepsilon_\alpha). \tag{D2}$$

Then it follows that we can write

$$\frac{1}{m(\varepsilon_\alpha)}=2\int_{-W}^{W}d\varepsilon_\beta\left\{\rho(\varepsilon_\beta)\frac{|\langle\alpha|v_x|\beta\rangle|^2\frac{\hbar}{\tau_\beta}}{(\varepsilon_\beta-\varepsilon_\alpha)^2+\left(\frac{\hbar}{\tau_\beta}\right)^2}\right\}\left\{\frac{(\varepsilon_\beta-\varepsilon_\alpha)}{\frac{\hbar}{\tau_\beta}}\right\}=\frac{2c_\alpha D(\varepsilon_\alpha)}{\hbar}, \tag{D3}$$

where c_α is a constant, defined by Eq. (D3). Its value and sign depends on the density of states variation. One can also see this by doing an integration by parts using the product of the two functions in brackets. The constant carries a sign (as does the effective mass) which depends on the energy. We can rewrite it as $c_\alpha=\langle(\varepsilon_\beta-\varepsilon_\alpha)\frac{\tau_\beta}{\hbar}\rangle_\beta$, where the average is to be taken with the diffusivity weighting function under the integral as defined by Eq. (D3). This equation is still *exact*. In the random-phase approximation, kinetic energies are of the same order as the energy uncertainty and c_α is ~-1 or $+1$ depending on whether we are near the top or the bottom of the band.

*Present address: Département de Physique and RQMP, Université de Sherbrooke, Sherbrooke, QC, Canada.
†Present address: Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois, USA.

[1] B. Movaghar and R. W. Cochrane, Phys. Status Solidi B **166**, 311 (1991).
[2] T. Ando, A. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[3] Y. Yao, L. Kleinman, A. H. MacDonald, J. Sinova, T. Jungwirth, D.-S. Wang, E. Wang, and Q. Niu, Phys. Rev. Lett. **92**, 037204

(2004).
[4] M. Abolfath, T. Jungwirth, J. Brum, and A. H. MacDonald, Phys. Rev. B **63**, 054418 (2001).
[5] T. Jungwirth, Q. Niu, and A. H. MacDonald, Phys. Rev. Lett. **88**, 207208 (2002).
[6] T. Jungwirth, J. Sinova, K. Y. Wang, K. W. Edmonds, R. P. Campion, B. L. Gallagher, C. T. Foxon, Qian Niu, and A. H. MacDonald, Appl. Phys. Lett. **83**, 320 (2003).
[7] T. Jungwirth, J. Sinova, J. Masek, J. Kucera, and A. H. MacDonald, Rev. Mod. Phys. **78**, 809 (2006).
[8] J. Sinova, T. Jungwirth, and J. Cerne, Int. J. Mod. Phys. B **18**, 1083 (2004).
[9] J. Sinova, D. Culcer, Q. Niu, N. A. Sinitsyn, T. Jungwirth, and A. H. MacDonald, Phys. Rev. Lett. **92**, 126603 (2004).
[10] R. Karplus and J. M. Luttinger, Phys. Rev. **95**, 1154 (1954).
[11] L. E. Ballentine, in *Liquid Metals 1976*, edited by R. Evans and D. A. Greenwood, IOP Conference Proceedings No. 30 (Institute of Physics, London, 1977), Pt. 2, Chap. 1, pp. 188–199.
[12] H. A. Engel, B. I. Halperin, and E. I. Rashba, Phys. Rev. Lett. **95**, 166605 (2005).
[13] J. Smit, Physica (Amsterdam) **24**, 39 (1958).
[14] N. F. Mott and H. S. W. Massey, *Theory of Atomic Collisions* (Clarendon, Oxford, 1965).
[15] L. M. Roth, in *Liquid Metals 1976*, edited by R. Evans and D. A. Greenwood, IOP Conference Proceedings No. 30 (Institute of Physics, London, 1977), Pt. 2, Chap. 1, pp. 200–206.
[16] B. Movaghar and R. W. Cochrane, Z. Phys. B: Condens. Matter **85**, 217 (1991).
[17] T. Matsubara and T. Kaneyoshi, Prog. Theor. Phys. **40**, 1257 (1968).
[18] P. Streda, J. Phys. C **15**, L717 (1982).
[19] S. Datta, Phys. Rev. Lett. **44**, 828 (1980).
[20] J.-N. Chazalviel, Phys. Rev. B **11**, 3918 (1975).
[21] P. Nozieres and L. Lewiner, J. Phys. (France) **34**, 901 (1973).
[22] L. M. Roth, B. Lax, and S. Zwerdling, Phys. Rev. **114**, 90 (1959).
[23] P. Wölfle and K. A. Muttalib, Ann. Phys. **15**, 508 (2006).
[24] S. K. Lyo and T. Holstein, Phys. Rev. Lett. **29**, 423 (1972).
[25] L. Berger, Phys. Rev. B **2**, 4559 (1970).
[26] E. I. Rashba, Phys. Rev. B **68**, 241315(R) (2003).
[27] R. C. Fivaz, Phys. Rev. **183**, 586 (1969).
[28] D. Ruzmetov, J. Scherschligt, D. V. Baxter, T. Wojtowicz, X. Liu, Y. Sasaki, J. K. Furdyna, K. M. Yu, and W. Walukiewicz, Phys. Rev. B **69**, 155207 (2004).
[29] G. Bush and J. J. Guentherodt, Solid State Phys. **29**, 335 (1974).
[30] H. Overhof and P. Thomas, *Electronic Transport in Hydrogenated Amorphous Semiconductors* (Springer-Verlag, Berlin, 1989).
[31] B. Velicky, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. **175**, 747 (1968).
[32] L.-F. Arsenault, B. Movaghar, P. Desjardins, and A. Yelon, Phys. Rev. B **77**, 115211 (2008).
[33] L.-F. Arsenault, M.S. thesis, École Polytechnique de Montréal, 2006.
[34] E. M. Chudnovsky, Phys. Rev. Lett. **99**, 206601 (2007).
[35] E. A. de Andrada e Silva, G. C. La Rocca, and F. Bassani, Phys. Rev. B **50**, 8523 (1994).