The Dissertation Committee for Dimitrie Max Culcer
certifies that this is the approved version of the following dissertation:

Novel intrinsic effects in charge and spin transport in semiconductors

Committee:

Qian Niu, Supervisor

Allan H. MacDonald

Michael Marder

Leonard Kleinman

Lorenzo Sadun

Novel intrinsic effects in charge and spin transport in semiconductors

by

Dimitrie Max Culcer, M. Phys.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2005

Acknowledgments

I wish to acknowledge the patient, tireless and unquestioning support offered to me by my supervisor, Prof. Qian Niu, without whose guidance this work would have been at best a misguided investigation on the periphery of solid state physics. I am grateful also for the many fruitful discussions and collaborations with Prof. Allan H. MacDonald over the years. Austinites who have helped me at one time or another include Junren Shi, Yugui Yao, Di Xiao, Jie Liu, Chuanwei Zhang, Chih Piao Chuu, Ping Zhang, Roberto Diener and Artem Dudarev. I wish to thank them all for their kind assistance. I wish to thank, in addition, Profs. Qikun Xue, Enge Wang, Rui Rui Du, Wuming Liu and Biao Wu of the Institute of Physics of the Chinese Academy of Sciences for their support and inspiration during my repeated stays in their most intriguing and inspiring land.

Dimitrie Max Culcer

The University of Texas at Austin

December 2005

Novel intrinsic effects in charge and spin transport in semiconductors

Publication No. _____

Dimitrie Max Culcer, Ph.D.
The University of Texas at Austin, 2005

Supervisor: Qian Niu

This work investigates intrinsic (scattering-independent) contributions to charge and spin transport in semiconductors which have been at the forefront of research in recent years. The first chapter explores an intrinsic contribution to the anomalous Hall effect, which has been related to the Berry phase, in two dimensional paramagnetic systems. The following two chapters develop a theory of spin transport in inversion-asymmetric semiconductors. Following that, a theory describing charge and spin transport in degenerate bands is formulated. Possibilities of experimental observation are emphasized in each case. Throughout these chapters, the point of view is that of wave packet dynamics. Finally, the last chapter surveys all the above phenomena from a density matrix point of view.

Contents

List of Tables

List of Figures

Chapter 1

Anomalous Hall effect in two dimensional paramagnetic systems

1.1 Introduction

When a nonferromagnetic metallic sample is exposed to a perpendicular external magnetic field, the Lorentz force acting on the current carriers gives rise to a transverse voltage in the plane of the sample. The transverse component of the resistivity ρ_{xy} depends on the magnetic field through:

$$\rho_{xy} = R_0 B \tag{1.1}$$

where $\mathrm{R}_0 = \frac{1}{\mathrm{ne}}$ is known as the Hall coefficient. This phenomenon is known as the ordinary Hall effect.

In many ferromagnets, however, the transverse resistivity acquires an additional term which is often seen to be proportional to the magnetization of the sample,

and becomes constant once the sample has reached its saturation magnetization M_s. Empirically one writes:

$$\rho_{xy} = R_0 B + R_s M \tag{1.2}$$

The effect is referred to as the anomalous Hall effect while the constant R_s is called the anomalous Hall coefficient. It can be seen from the second term above that ferromagnets display a spontaneous Hall conductivity in the absence of an external field. The effect was subsequently noted in a large number of bulk alloys, as well as, in recent experiments, in materials which exhibit colossal magnetoresistance [1, 2] and ferromagnetic semiconductors. Recent studies of ferromagnetic semiconductors such as (Ga,Mn)N films have in fact reported ferromagnetic behavior at room temperature [3, 4].

Although it has been known for close to half a century, the AHE has had a controversial history and it remains a somewhat poorly understood phenomenon. Karplus and Luttinger [5] pioneered the theory of the AHE, finding that the spin splitting of bands can give rise to a Hall conductivity in the presence of spin orbit coupling. Smit [6] countered that in a perfectly periodic lattice the AHE could not occur without scattering from impurities, and introduced the skew scattering mechanism to explain it. This mechanism, in which an electron is scattered at an angle to its original direction, gives a contribution proportional to ρ, the diagonal resistivity. In a more complete treatment, Luttinger [7] found a term corresponding to skew scattering but maintained that the scattering free contribution to the AHE still remains. There has been much debate on the possibility of this scattering free contribution in principle and on its relative importance, if it exists, in real materials.

Later, a new mechanism, called side jump was introduced by Berger [8] to explain the observed ρ^2 dependence, although the scattering free contribution gives the same dependence. In side jump, the electron incident into the area of influence of the potential emerges parallel to its original direction but displaced perpendicular

to it. This latter term is supposed to dominate in alloys, where ρ is high. It is nevertheless not clear how to relate the side jump mechanism to the systematic theory proposed by Luttinger.

In recent years, the scattering free contribution of Luttinger and Karplus was rederived in a semiclassical analysis of wave packet motion in Bloch bands by Chang and Niu [9] and Sundaram and Niu [10] and was attributed to a Berry phase effect in k-space. A more rigorous derivation [11] based on the Kubo formula gives the same result. This contribution was also evaluated for the mean-field bands of semiconductor ferromagnets, yielding good agreement with experiments without any parameter fitting. This theory [12] of the AHE is based on the Stoner description of ferromagnetism, considering the charge carriers to be quasiparticles in spontaneously split Bloch bands. It is to be distinguished from the mechanism of Ye *et al.* [13] based on the Berry phase in real space. The motivation behind the current effort is to provide a conceptual framework for the theoretical study of the AHE in magnetic quantum wells and heterostructures, which have been realized in recent years. These structures constitute the simplest systems in which the Berry phase can be evaluated analytically from the Hamiltonian including the Rashba spin orbit coupling and provide a suitable ground for testing a theory based on fundamental physics. We shall concentrate our attention on the conduction band and the topmost valence band of an inversion asymmetric semiconductor heterostructure in an exchange field supplied through doping with Mn and calculate the anomalous Hall conductivities for the two bands. Although ferromagnetic behavior has not been observed in II-VI heterostructures we shall concentrate on II-VI semiconductors, as they can be doped with Mn more heavily than III-V semiconductors.

The chapter is organized as follows. Section 1.2 is devoted to the symmetry aspects of the anomalous Hall effect. In section 1.3, within the framework of the effective mass approximation applied to a doubly degenerate band, we calculate the

Berry phase of the wave function which yields the off-diagonal conductivity σ_{xy}. We consider an ideal situation with T=0, where we follow the method used by Chang and Niu [9] and Sundaram and Niu [10] for the semiclassical treatment of carrier motion in 2D. Under certain circumstances one can make the approximation that the anomalous Hall conductivity is quantized, taking the values:

$$|\sigma_{xy}| = \frac{e^2}{2h} \tag{1.3}$$

In sections 1.4 and 1.5 we apply the theory to wurtzite and zincblende structures respectively. We consider finite temperature corrections and discuss the conditions under which equation (1.3) holds. Moreover, we investigate the variation of the conductivity with temperature, exchange coupling and spin orbit constant. In the last section we examine the effect of placing the system in an external magnetic field and determine the optimal parameters needed for the observation of the AHE in heterostructures in the laboratory.

1.2 Symmetry considerations

In a perfect crystal, according to Bloch's theorem, the wave function for a band with index n is decomposed into two parts:

$$|\Psi_n(\mathbf{k},\mathbf{r})\rangle = e^{i\mathbf{k}\cdot\mathbf{r}}|u_n(\mathbf{k},\mathbf{r})\rangle \tag{1.4}$$

where $|\mathrm{u_n}(\mathbf{k},\mathbf{r})\rangle$ is a function with the periodicity of the lattice. The semiclassical motion of a charge carrier through the crystal is described by constructing a wave packet out of Bloch wave functions. The dynamics of such a wave packet are given

by the following equations of motion:

$$\dot{\mathbf{r}} = \frac{1}{\hbar}\frac{\partial \varepsilon_n}{\partial \mathbf{k}} - \dot{\mathbf{k}} \times \mathbf{\Omega}_n \tag{1.5}$$

$$\dot{\mathbf{k}} = -\frac{e}{\hbar}(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}) \tag{1.6}$$

which determine the position vector and wave vector of the center of the wave packet in the presence of external electromagnetic fields, with $\mathbf{\Omega}$ the Berry curvature. The Berry curvature of a band is defined by the following expression:

$$\mathbf{\Omega}_n = -\Im\langle \frac{\partial u_n}{\partial \mathbf{k}} | \times | \frac{\partial u_n}{\partial \mathbf{k}} \rangle, \tag{1.7}$$

where $\Im$ denotes the imaginary part. The term containing the Berry curvature is usually neglected, due to the fact that it frequently vanishes by symmetry, as in crystals which are invariant with respect to both time reversal and spatial inversion (e.g. non-magnetic Bravais crystals [10]).

In the AHE, the additional contribution to the current is perpendicular to the direction of the electric field and independent of the magnetic field. We now show it to be related to the Berry curvature, $\mathbf{\Omega}_{\mathrm{n}}$ which appears in the equations of motion as an additional term in the velocity. This is the same as the velocity correction derived previously by Luttinger. From the two equations it is apparent that this correction term is perpendicular to $\dot{\mathbf{k}}$ and therefore perpendicular to the direction of the Lorentz force. In the absence of an external magnetic field $\mathbf{B}$, this term is seen to be perpendicular to the electric field $\mathbf{E}$, giving a transverse component of the velocity. This velocity adds a transverse term in the current producing a contribution to the off diagonal conductivity. Therefore, as long as $\mathbf{\Omega}_{\mathrm{n}}$ is nonzero it is possible to have an off diagonal conductivity term which is independent of $\mathbf{B}$.

The Berry curvature is related to the Berry phase [14] (denoted by γ_{n}), which is the phase acquired by the wave function upon being transported around a loop

in k-space. According to Stokes' theorem:

$$\int_A d\mathbf{S}\cdot\mathbf{\Omega}_n = \int_{\partial A} d\mathbf{k}\cdot\langle u_n|\frac{\partial}{\partial\mathbf{k}}|u_n\rangle = \gamma_n \tag{1.8}$$

In the above the loop around which the wave function is transported is denoted by ∂A and the area enclosed by the loop by A. The Berry curvature can therefore be regarded as the Berry phase per unit area of k-space.

In the following we give an analysis of the main symmetry aspects of the problem. From the requirement that the semiclassical equations be invariant under time reversal it is apparent that $\mathbf{\Omega}_\mathrm{n}$ must be odd under this transformation, namely $\mathbf{\Omega}_\mathrm{n}(\text{-}\mathbf{k})=\text{-}\mathbf{\Omega}_\mathrm{n}(\mathbf{k})$. A geometric argument can also be made by noting that the Berry phase is a path dependent quantity. Under time reversal, both the path along which the wave function is transported and the orientation of the wave vector $\mathbf{k}$ are reversed. A clockwise path spanning a set of wave vectors $\{\mathbf{k}\}$ becomes an anticlockwise path spanning the set of wave vectors $\{-\mathbf{k}\}$. This implies that the Berry phase changes sign uner time reversal and the Berry curvature satisfies the above constraint.

One can also obtain this result by carrying out the explicit transformation of Eq.(7) under time reversal. If $|\mathrm{u_n}\rangle$ is written in terms of the real and imaginary parts of its components:

$$|u_n\rangle = \begin{pmatrix} \Re|v_n\rangle + i\Im|v_n\rangle \\ \Re|w_n\rangle + i\Im|w_n\rangle \end{pmatrix}, \tag{1.9}$$

in which $\Re$ denotes the real part, then application of the time reversal operator will result in:

$$\mathcal{T}|u_n\rangle = \begin{pmatrix} -i\Re|w_n\rangle - \Im|w_n\rangle \\ i\Re|v_n\rangle + \Im|v_n\rangle \end{pmatrix} \tag{1.10}$$

producing a change of sign in the Berry curvature.

If time reversal symmetry is present, Kramers degeneracy must also be present, imposing $\varepsilon_{\mathrm{n}}(\mathbf{k}) = \varepsilon_{\mathrm{n}}(\text{-}\mathbf{k})$. Therefore, if the state at wave vector $\mathbf{k}$ is occupied so is the state at wave vector -$\mathbf{k}$. This, together with the condition $\mathbf{\Omega}_{\mathrm{n}}(\text{-}\mathbf{k})=\text{-}\mathbf{\Omega}_{\mathrm{n}}(\mathbf{k})$ implies that the integral of $\mathbf{\Omega}_{\mathrm{n}}$ over all filled states vanishes. Therefore, in general it is always necessary for the system to lack time reversal symmetry in order for the AHE to occur.

To obtain a nonzero anomalous Hall conductivity, the spin orbit interaction must also be present in order to couple the spin up and spin down bands. This coupling transfers the time reversal violation from the spin degree of freedom to the orbital motion, which is responsible for the Berry curvature. An example is provided by ferromagnetic (Ga,Mn)As [15, 16] crystals, in which, without spin orbit, the valence band wave functions at k=0 are eigenstates of $\hat{\mathbf{L}}$, the orbital angular momentum operator, with eigenvalue l=1 and thus sixfold degenerate. When spin orbit is included the k=0 band wave functions are eigenstates of the total angular momentum operator $\hat{\mathbf{J}}$, splitting into a fourfold degenerate j=3/2 level (containing the heavy holes and the light holes) and a twofold degenerate j=1/2 level (the split off band). Away from k=0, there is a correction proportional to $(\hat{\mathbf{J}} \cdot \mathbf{k})^2$, which partially lifts the degeneracy of the bands. This term provides a k-dependent quantization direction for the angular momentum, so that as the wave vector is displaced the angular momentum is rotated and it is possible to obtain a non-zero Berry curvature.

In 2D, the quantum confinement lifts the degeneracy of the heavy hole and light hole bands, so that at k=0 it is possible to separate the Hamiltonian into independent 2×2 blocks. For finite k, each block remains degenerate in the presence of both time reversal and spatial inversion symmetries, based on Kramers' theorem. With time reversal symmetry $|\mathbf{k},\uparrow\rangle$ is equivalent to $|-\mathbf{k},\downarrow\rangle$ while with space inversion symmetry $|\mathbf{k},\uparrow\rangle$ is equivalent to $|-\mathbf{k},\uparrow\rangle$. Therefore, with both symmetries $|\mathbf{k},\uparrow\rangle$ is

equivalent to $|\mathbf{k},\downarrow\rangle$. In the absence of space inversion symmetry it is possible to break the degeneracy at each finite k. The space inversion asymmetry gives rise to the Rashba spin orbit interaction [17]:

$$V_{so} = \alpha_{mat} f(k)(\sigma \times \mathbf{k}) \cdot \hat{\mathbf{z}} \tag{1.11}$$

where α_{mat} is a constant, σ is the vector of Pauli spin matrices, $\mathbf{k}$ is the two dimensional wave vector in the xy-plane and f(k) depends only on the magnitude of the wave vector. The asymmetry can originate from either the crystal structure (bulk inversion asymmetry) or the confinement potential (structure inversion asymmetry). The Rashba interaction has been found to be the main mechanism responsible for the zero field spin splitting in 2DEGs [18, 19, 20, 21].

It is apparent from the above that the spin orbit coupling provides a k-dependent quantization direction for the charge carriers' spins. The spins prefer to lie in the xy-plane and be perpendicular to the wave vector. As a result, when the wave vector sweeps a circle around the origin, the spins are rotated by a solid angle of 2π, and acquire a Berry phase of π. Since this phase is independent of the area enclosed, it follows that the Berry curvature is singular at the origin and is zero everywhere else. When an exchange field is applied, the spins are tilted out of the xy-plane. The amount of tilting depends on the competition between the Rashba term and the exchange field. From the wave vector dependence of the Rashba term it can be seen that the solid angle swept by the spins is different from 2π and depends on the size of k, tending to zero as the radius of the circle tends to zero. This implies that the Berry curvature is now spread out and is finite at the origin. As will be shown in more detail in the following sections, such a Berry curvature will lead to a finite contribution to the AHE.

1.3 General treatment of Berry curvature and Hall conductivity

The 2D anomalous Hall conductivity is calculated at T=0 and shown to be quantized. We consider a 2×2 Hamiltonian describing the spin split conduction/valence band of a semimagnetic semiconductor in the presence of an exchange field and spin orbit coupling. Effects of band mixing are neglected, which is a suitable approximation for the bandstructures of the materials we shall consider-the top two valence bands and the conduction band in wurtzite structures and the conduction band of zincblende materials.

The exchange field due to the magnetic impurities is taken to be uniform and directed along the z-axis, normal to the heterostructure. Based on a mean field model, we consider the interaction to be described by a vector $\mathbf{h}_0$, which for simplicity has units of energy. The magnitude of the interaction is tuned by controlling the concentration of Mn but its effect will be masked by thermal fluctuations once $\mathrm{h_0} \leq \mathrm{k_B T}$.

In a narrow quantum well in which the subbands are widely separated the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian, with m^* the band electron effective mass, $\gamma = \frac{\hbar^2}{2\mathrm{m}^*}$ and $\mathrm{k_\pm}=\mathrm{k_x} \pm \mathrm{ik_y}$ is:

$$H = \gamma k^2 I_{2\times 2} + \begin{pmatrix} h_0 & i\alpha_{mat} f(k) k_- \\ -i\alpha_{mat} f(k) k_+ & -h_0 \end{pmatrix} \tag{1.12}$$

It is readily seen to have the eigenvalues:

$$E_\pm = \gamma k^2 \pm \sqrt{h_0^2 + \alpha_{mat}^2 k^2 f(k)^2} \tag{1.13}$$

yielding two subbands, separated by $2\mathrm{h_0}$.

Assuming T=0 for the time being we take the bottom subband to be occupied and the top one to be empty, while the Fermi level corresponds to $\mathrm{k_F}=(4\pi \mathrm{n})^{1/2}$.

The form of the Berry curvature for a general f(k) is:

$$\Omega_z^{\uparrow/\downarrow} = \mp\frac{1}{2}\frac{\alpha_{mat}^2 h_0 f(k)\frac{d}{dk}[kf(k)]}{[h_0^2+\alpha_{mat}^2 k^2 f(k)^2]^{3/2}}. \tag{1.14}$$

The geometrical phase factor is the integral of the curvature over all wave vectors [11]. As the upper band is empty the integral over it is zero and one only needs to consider the curvature of the lower band, $\Omega^{\downarrow}$:

$$\Gamma^{\downarrow} = \iint_{k<k_F} \Omega^{\downarrow}{}_z d^2\mathbf{k} = \pi[1 - \frac{h_0}{[h_0^2+\alpha_{mat}^2 k_F^2 f(k_F)^2]^{1/2}}]. \tag{1.15}$$

To maximize the conductivity, the interval between k = 0 and k = k_F should cover the region over which the Berry curvature is significant, so that k_F must be equal to several times k_c, the wave vector at which the curvature falls to half its maximum value. As k_F is fixed by the number density the way to accomplish this is to have $h_0 \ll \alpha_{mat}\ k_F$. When this relation holds the phase $\Gamma^{\downarrow}$ is very nearly π.

At zero temperature the conductivity σ for a full band is equal to the integral over the Brillouin zone of the component of the Berry curvature parallel to z and is thus proportional to the Berry phase. The upper limit of the integral can be taken to infinity:

$$\sigma_{xy}^{\downarrow} = -\frac{e^2}{h}\iint_{k_F\to\infty} \Omega_z^{\downarrow}\frac{d^2\mathbf{k}}{2\pi} \tag{1.16}$$

which results in:

$$\sigma_{xy} = -\frac{e^2}{2h}\int_0^{\infty} dk\frac{\alpha_{mat}^2 h_0 k f(k)\frac{d}{dk}[kf(k)]}{[h_0^2+\alpha_{mat}^2 k^2 f(k)^2]^{3/2}} = -\frac{e^2}{2h}. \tag{1.17}$$

From the above we see that the conductivity is approximately quantized, regardless of the form of f(k). It is worth noting that σ does not depend on the size of the spin orbit splitting constant α_{mat} nor on the magnitude of the external magnetic field and that exact quantization occurs when the Berry phase is π, i.e. the spin lies in

the xy-plane.

1.4 Wurtzite structures at finite temperatures

The conduction band and the bottom valence band of wurtzite transform according to the Γ_7 representation of the rotation group at k=0, while the top valence band transforms according to Γ_9. The latter however is known empirically not to exhibit a linear spin splitting.

The coefficient $\alpha_{\rm mat}$ introduced above is replaced by $\alpha_{\rm w}$. Then the interaction for the Γ_7 band is given by:

$$V_{so} = \alpha_w(\sigma \times \mathbf{k}) \cdot \hat{\mathbf{z}}. \tag{1.18}$$

The energy bands, corresponding to the dispersion relation:

$$E_{\pm} = \gamma k^2 \pm \sqrt{h_0^2 + \alpha_w^2 k^2} \tag{1.19}$$

are plotted in Fig. 1.1 as a function of k.

The Berry curvature is pointing along the z-axis:

$$\Omega_z^{\uparrow/\downarrow} = \mp \frac{1}{2} \frac{\alpha_w^2 h_0}{(\alpha_w^2 k^2 + h_0^2)^{3/2}}. \tag{1.20}$$

The absolute value of the Berry curvature, $\Omega_{\rm z}$ is plotted in Fig. 1.2. It falls to half its maximum value when the wave vector is equal to:

$$k_c = \pm \frac{0.77 h_0}{\alpha_w} \tag{1.21}$$

and its effect becomes negligible once the magnitude of k exceeds several times that

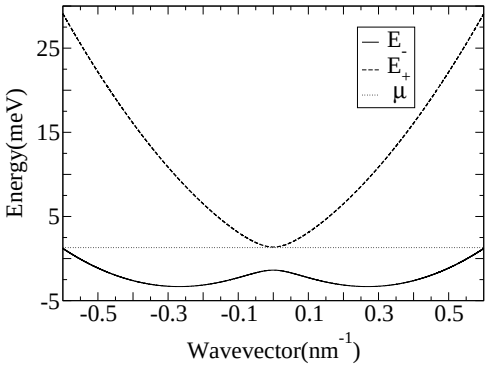

Figure 1.1: Band dispersion relation for 2D holes in the Γ_7 valence band of wurtzite structures. The parameters are n=2.9×10^{12}cm^{-2}, α_w=23meVnm, h_0=1.38meV and $m^*=0.9m_0$.

of k_c.

At finite temperatures one must take into account the fact that the Fermi-Dirac distribution deviates from the step function at T=0, which is done by incorporating the distribution function into the expression for σ_{xy}. It is also important to maintain a carrier number density in the range in which AHE is not overshadowed by disorder effects. High densities cause interface effects to become important whereas low densities will cause pockets of electrons to be isolated in localized states. In addition to the above, one must consider the contribution from both the lower and the upper band as there exists a finite fraction of carriers excited into the E_+ band. The two conductivities are:

$$\sigma_{\uparrow/\downarrow} = \pm\frac{e^2}{2h}\int_0^\infty dk \frac{k\alpha_w^2 h_0}{(h_0^2+\alpha_w^2k^2)^{3/2}}\frac{1}{e^{(E_\pm(k)-\mu)/k_BT}+1} \tag{1.22}$$

with E(k) given by (1.19). The total conductivity σ_{xy} is the sum of the two:

$$\sigma_{xy} = \sigma_{xy}^{\uparrow} + \sigma_{xy}^{\downarrow} \tag{1.23}$$

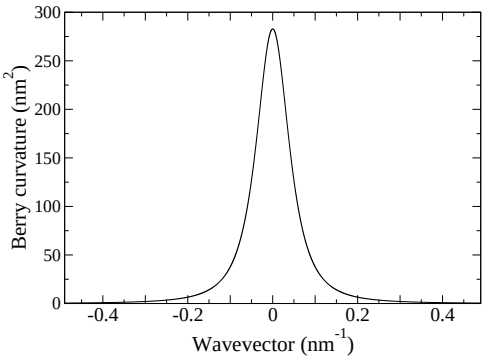

Figure 1.2: Absolute value of Berry curvature Ω_z as a function of wavevector for the Γ_7 valence band of wurtzite structures.

We consider the conduction band and concentrate on CdSe, where α_w has been measured to be 10meVnm and the effective mass m^* is $0.13m_0$. μ is determined by the number density and exchange field, which are fixed at 1×10^{11} cm^{-2} and 0.8meV. Our numerical calculations show that under these conditions the maximum conductivity is:

$$|\sigma_{xy}| = 0.125\frac{e^2}{h}. \tag{1.24}$$

It is not quantized, but the effect is still observable.

In the case of the valence band, theory gives an estimate for α_w of 23meVnm while experiment sets an upper limit of 90meVnm (see below), and we employ the theoretical value as a worst case scenario. The effective mass is $0.9m_0$, the number density is set to 2.9×10^{12} cm^{-2}, h_0 is fixed at 1.38meV and the temperature at 0.1K. Repeating the calculation yields:

$$|\sigma_{xy}| = 0.45\frac{e^2}{h} \tag{1.25}$$

showing that the conductivity approaches the quantized value.

We now investigate the dependence of the integral in equation (1.16) upon the

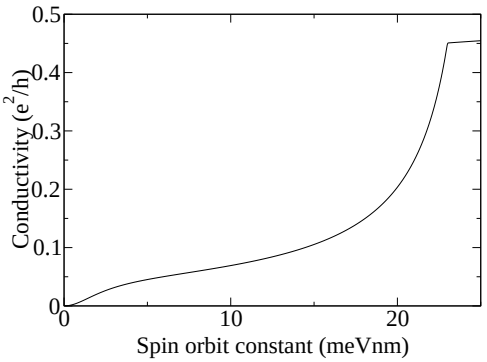

Figure 1.3: Variation of the conductivity with α_w in the case of the Γ_7 valence band of wurtzite structures.

spin orbit coupling constant (Fig. 1.3), maintaining the other parameters at their values for the valence band. We find that at T=0.1K it increases with increasing α_w, saturating to $0.45\frac{e^2}{h}$. The shape of the graph can be explained by noting that the effect of increasing α_w is to bring down the chemical potential and flatten the lower band in such a way that its Fermi wave vector is unchanged. The point where the conductivity reaches its maximum corresponds to the point where the chemical potential crosses from the top band into the bottom one so that at very low temperatures only the latter is occupied. Since there are only carriers in the lower band, as α_w increases they acquire approximately the same Berry phase until the chemical potential touches the band maximum at k=0, beyond which our theory does not apply. The shape of the curve as far as the plateau follows from the fact that as the chemical potential is lowered fewer states are available in the upper band. The plateau itself is understood by noting that increasing α_w makes the curvature narrower but after a point almost all the area over which Ω is appreciable has been covered, so further increasing α_w will not make a considerable difference.

The dependence upon the exchange coupling is studied next (Fig. 1.4). It

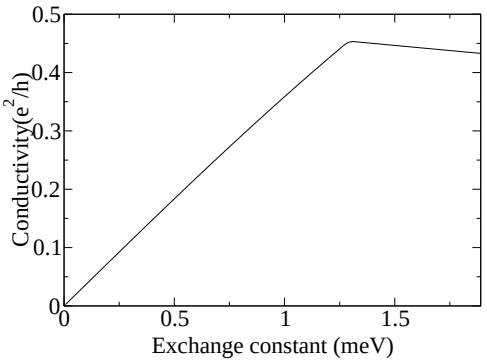

Figure 1.4: Variation of the conductivity with h_0 for the Γ_7 valence band of wurtzite structures.

can be seen that σ_{xy} reaches a maximum when h_0 is approximately 1.38 meV, after which it drops. At first, when there is no magnetic interaction, the spin lies in the xy-plane. As h_0 increases the spin is tilted out of the plane by larger amounts, increasing the phase acquired by the wave function, until it reaches a maximum. As $h_0 \to \infty$ the spin becomes parallel to h_0 and the phase gradually falls to zero. Increasing h_0 makes Ω wider so less of the curvature is covered in the range $k = 0$ to k_F. The sudden fall in the conductivity beyond the maximum is therefore a combined effect - the magnitude of the curvature is smaller and less of the curvature is covered in the integral.

These two plots illustrate the fundamental physics of the system, namely the interplay between the Rashba and Zeeman effects giving rise to the anomalous Hall conductivity through the Berry phase acquired by the wave function. The dynamics can be viewed as a competition between the Zeeman term, which by itself would align the spin with the z-axis and the Rashba term, which draws it towards the xy-plane. Without the spin orbit interaction, the Berry phase is zero yielding zero conductivity whereas without the exchange field the energy gap vanishes and the bands overlap. What is more, as h_0 tends to infinity the spins align themselves

along z in such a way that the wave function does not acquire a Berry phase. At this stage the wave vector precesses on the Fermi surface at an infinite rate, which is equivalent to no precession at all. Lastly, as $\alpha_{\rm w}$ tends to zero the spins once more align with the z-axis.

Finally, we have observed the temperature dependence of the integral in equation (1.23), with $\alpha_{\rm w}$ chosen as before. As Fig. 1.5 shows, the conductivity declines over the range T=10mK to T=1K, which is attributed to the fact that raising the temperature causes more carriers to be excited across the gap, increasing the size of the negative contribution.

These two situations are similar to the limit $h_0 \to 0$: the bandgap here does not disappear, but it is bridged by facilitating the movement of carriers across it.

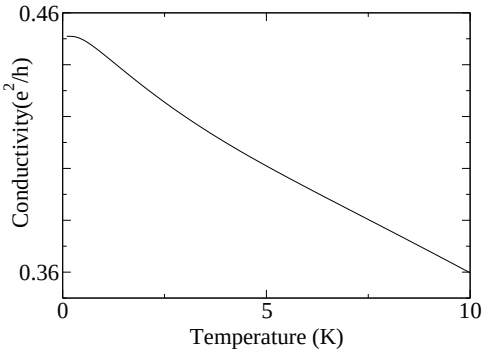

Figure 1.5: Variation of the conductivity with temperature for the Γ_7 valence band of wurtzite structures.

1.5 Zincblende structures at finite temperatures

Having investigated the underlying physics of the problem for a wurtzite QW, we turn our attention to a case which promises immediate experimental realization. We shall restrict our discussion of zincblende materials to the conduction band of

$Hg_{1-x}Mn_xTe$, in which the linear term in k is not allowed by symmetry. Instead, the first term in the expansion is cubic in k and the spin orbit term takes the form:

$$V_{so} = \alpha_{zb} k^2 (\sigma \times \mathbf{k}) \cdot \hat{\mathbf{z}} \tag{1.26}$$

where α_{zb} replaces α_{mat}. This expression is valid near k=0, but is not accurate as k approaches k_F. In order to improve the accuracy we have chosen the polynomial coefficients b_1 and b_2 so as to match the dispersion relation with that shown in Fig. 7 of Ref. [22], namely:

$$V_{so} = \frac{\alpha_{zb} k^2 (\sigma \times \mathbf{k}) \cdot \hat{\mathbf{z}}}{1 + b_1 k^2 + b_2 k^4} \tag{1.27}$$

yielding the energy bands (Fig. 1.6):

$$E_{\pm} = \gamma k^2 \pm \sqrt{h_0^2 + \frac{\alpha_{zb}^2 k^6}{(1 + b_1 k^2 + b_2 k^4)^2}} \tag{1.28}$$

and the absolute value of the z-component of the Berry curvature (Fig. 1.7):

$$\Omega_z^{\uparrow/\downarrow} = \mp \frac{\alpha_{zb}^2 h_0 k^2}{2(h_0^2 + \alpha_{zb}^2 k^6)^{3/2}} \frac{(3k^2 + b_1 k^4 - b_2 k^6)}{(1 + b_1 k^2 + b_2 k^4)^3} \tag{1.29}$$

We consider the optimum achievable conditions for the observation of the anomalous Hall conductivity. The doping density is n=2.8×10^{11}cm^{-2}, the spin orbit constant from Ref. [22] is approximately α_{zb}=10000meVnm^3, we set the exchange field to be equal to 3.38meV and m*=0.034m$_0$. Under these conditions the conductivity is:

$$|\sigma_{xy}| = 0.34 \frac{e^2}{h} \tag{1.30}$$

In Fig. 1.8-1.10 we have plotted the conductivity as a function of the spin orbit constant, exchange field and temperature. The graphs will be seen to have very similar features to the corresponding ones for wurtzite. These common features

have identical explanations in terms of the modification of the shape of the bands and the movement of the chemical potential relative to them, as discussed above.

It will be noticed that the zincblende graphs are smoother and the plateau in the spin orbit constant graph is missing. The qualitative differences in the behavior of the conductivity come about due to the difference in the shape of the bandstructure and Berry curvature in the two structures. In wurtzite Ω peaks at the origin and is appreciable within a disc centered at k=0. In zincblende on the other hand the curvature is zero at the origin and is concentrated within a ring on either side of the values of $\mathbf{k}$ at which it peaks. If the magnitude of the wave vector at which Ω has its maximum is denoted by k_Ω, it emerges that in order to maximize the anomalous conductivity the parameters must be adjusted such that k_F is large enough to contain the ring on the outer side of k_Ω but small enough for the number of states available in the upper band not to cause the contribution from it to cancel out the curvature from the lower band. In the spin orbit constant graph, after reaching a maximum, the conductivity quickly declines, since increasing α_{zb} has the effect of lowering the chemical potential, so that fewer states in the bottom band are integrated over. Due to the shape of the curvature, lowering the chemical potential causes those wave vectors at which the Berry curvature is significant to be omitted, resulting in a sharp decrease in the conductivity. Moreover, the fact that in zincblende the lower band does not have a maximum at k=0 means that our theory can be applied regardless of where in the band the chemical potential lies.

1.6 Other materials

In general, the expression for α_{mat} is $a_{mat}\langle E\rangle$. Here $\langle E\rangle$ is the expectation value of the total electric field felt by the carriers and a_{mat} a material specific parameter which is straightforwardly calculated using third order perturbation theory. It is

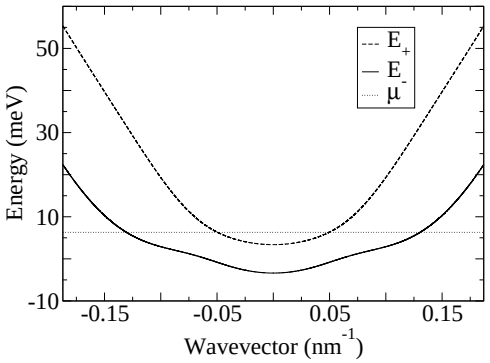

Figure 1.6: Band dispersion relation for 2D electrons in a zincblende lattice. The parameters are n = 2.8×10^{11}cm^{-2}, α_{zb} = 10000meV nm^3, h_0 = 3.38meV and $m^*=0.034m_0$.

customary to assume that the gradient of the confining potential has only a z-component, with the result that α_{mat} is given by $a_{mat}\langle E_z\rangle$. In the literature the size of the spin orbit coupling is parametrized either by direct measurements of α_{mat} (Table 1.1), calculated values of a_{mat} (Table 1.2), or the magnitude of the energy splitting at k=0 or k=k_F. This disguises the fact that the character of E_z is poorly understood and little literature is available on the topic. In a recent experiment, the electric field in the valence band of a GaAs/$Ga_{1-x}Al_xAs$ heterostructure was determined by Jusserand *et al.* [23] to be 17 mVnm^{-1}. E_z is assumed to scale with the band offset and can be increased by up to a factor of approximately 3.5 by applying a gate potential [24]. In addition it was pointed out by Lassnig [25] that the conduction band spin splitting is due to the electric field in the valence band, the two fields differing through the contributions of the interfaces [25, 26].

We present in Table 1.1 the maximum observed/calculated values of α_{mat} in the bulk for semiconductors with a strong spin-orbit interaction due to bulk inversion asymmetry (BIA). The small bulk GaAs spin-orbit coupling constant renders the effect of V_{so} negligible in GaAs, but in the other materials the size of α_{mat} is several

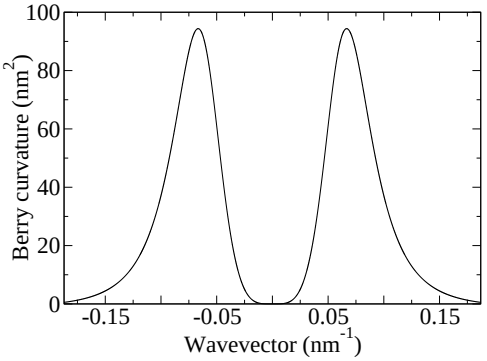

Figure 1.7: Absolute value of the Berry curvature for the conduction band of zincblende structures.

orders of magnitude larger. In Table 1.2 we list calculated values for the coefficient

Table 1.1: Spin orbit constant(meVnm)

GaAs [20]	0.69
HgMnTe [27]	100
InAs [28, 29]	30-45
HgTe gated QW [30]	40
CdTe/HgTe/CdTe [31]	40
$In_{0.75}Ga_{0.25}As/In_{0.75}Al_{0.25}As$ [32]	29.2
$In_{0.75}Ga_{0.25}As/InP$ [33, 34]	4.71-15
CdSe(holes) [35]	< 10 (expt)
CdSe(holes) [35]	6 (theory)
CdSe(electrons) [35]	< 90 (expt)
CdSe(electrons) [35]	23 (theory)

a_{46} for different materials. By comparing with the corresponding values of $\alpha_{\rm mat}$ one can obtain a rough estimate of the electric field in the valence band in the absence of a gate. In the case of InAs this field would lie in the range 25-40meVnm^{-1}.

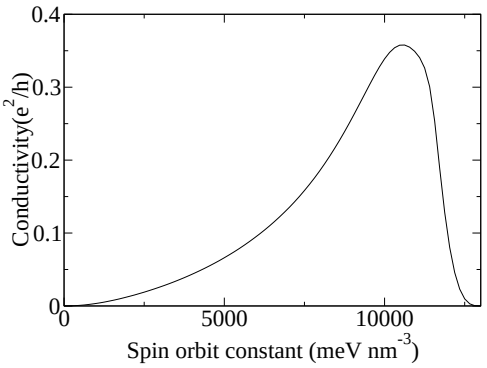

Figure 1.8: Variation of the conductivity with the spin orbit constant for zincblende structures.

1.7 Effect of magnetic field and disorder

AHE was first observed in ferromagnetic materials in the absence of an external magnetic field (in this case one would only need to apply a field to magnetize the material, lowering it to zero afterwards). As ferromagnetic heterostructures are yet to be realized and as ferromagnetism has not been observed in II-VI semiconductors, it is more sensible to consider the case of a paramagnetic system, in which the exchange field can be maintained only by applying an external magnetic field. In order to determine the regime in which AHE can be observed in a weak magnetic field one needs to consider the fact that a magnetic field will cause the system to be quantized into Landau levels - where the semiclassical approximation is not valid, as well as give rise to the ordinary Hall effect.

Table 1.2: The coefficient $a_{mat}(enm^2)$

GaAs [20]	0.055
$Hg_{0.8}Cd_{0.2}Te$ [27]	19.3
InAs [26]	1.17
InSb [26]	4-5.23
ZnSe [26]	0.01

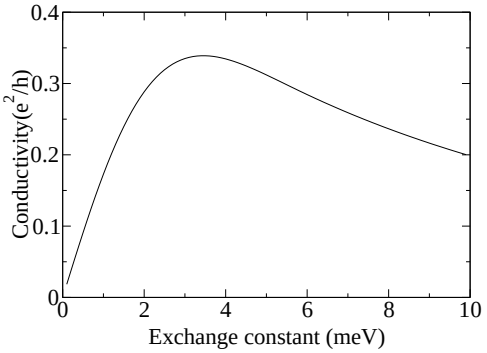

Figure 1.9: Variation of the conductivity with the exchange field for zincblende structures.

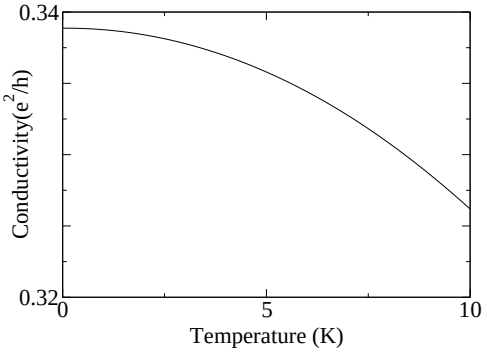

Figure 1.10: Variation of the conductivity with temperature for zincblende structures.

The first obstacle is circumvented by the presence of disorder in the sample, as the impurity scattering causes the Landau levels to broaden so that for a small enough magnetic field they overlap. The effect of disorder is parametrized by an impurity scattering time τ, which in II-VI heterostructures is of the order of 0.1ps [36]. To get round the second problem, the parameters must be matched so that the ordinary Hall conductivity does not overwhelm the anomalous one, making observation of the latter contribution clear. If the magnetic field and the scattering time are small enough to make the Landau levels overlap, $\omega_c \tau < 1$ must hold, where ω_c is

the cyclotron frequency. The condition that $\omega_c\tau < 1$ ensures that the semiclassical approximation is valid, but does not guarantee that the ordinary Hall conductivity will not greatly exceed the anomalous one. For small $\omega_c\tau$ the ordinary Hall contribution, which, in the absence of quantum oscillations is given by the Drude formula:

$$\sigma_{xy}^{OHE} = \frac{ne^2\tau}{m^*}\frac{\omega_c\tau}{1+\omega_c^2\tau^2} \tag{1.31}$$

tends to zero. To ensure the AHE is the dominant effect we set:

$$\sigma_{xy}^{OHE} < \sigma_{xy}^{AHE} \tag{1.32}$$

These two equations yield $(\frac{\mathrm{nh}}{\mathrm{m}^*})\omega_c\tau^2 < 1$.

It is also imperative to ensure the AHE itself is not completely overshadowed by disorder. To satisfy this requirement, the exchange splitting h_0 must exceed the energy fluctuation due to disorder, $\frac{\hbar}{\tau}$. It follows that the condition for the observation of AHE is:

$$\frac{2\pi n\hbar^2}{m^*}\omega_c\tau < \frac{\hbar}{\tau} < h_0 \tag{1.33}$$

For τ=0.1ps, the fluctuation $\frac{\hbar}{\tau}$ represents an energy of 6.5meV.

As it is desired to work with a narrow well, so as to keep the subbands as far from each other as possible, we shall set the well width at 10nm, close to the smallest that can be manufactured. Furthermore, the laboratory temperature will be fixed at 0.1K. We use the exchange constants $N_0\alpha$ and $N_0\beta$ in Table V of Ref. [37] to determine the optimal Mn concentration and external magnetic field for the observation of the AHE in $Cd_{1-x}Mn_xSe$ and $Hg_{1-x}Mn_xTe$ quantum wells.

For wurtzite ($Cd_{1-x}Mn_xSe$), with the value of α_w fixed we have chosen the carrier density n and exchange field h_0 in such a way as to have an observable conductivity in the valence band: n=2.9×10^{12} cm^{-2} and h_0=7meV. The Mn doping density will have to be 2.2%. At 0.1K, in order for the Brillouin function to saturate

the magnetic field must be approximately 1T. At this field, the ordinary Hall conductivity is less than 0.05 of the conductivity quantum, while the anomalous one is approximately 0.27.

In the case of zincblende ($Hg_{1-x}Mn_xTe$), the act of balancing the ordinary and anomalous conductivities is more difficult. The magnetic field cannot be as high as 1T, for that will produce a large ordinary contribution, but that is compensated by the fact that the exchange constant $N_0\beta$ is larger. In order to maintain the exchange splitting above the disorder broadening, i.e. h_0=7meV, it is sufficient to apply B=130mT and keep the Mn density unchanged at 2.2% (corresponding to 3.3×10^{26} m^{-3}, which is well within the experimentally achievable range [37]). At a carrier density of 1×10^{11} cm^{-2}, the ordinary and anomalous conductivities will be equal to just over 0.14 of the conductivity quantum.

Chapter 2

Semiclassical spin transport in spin-orbit coupled bands

2.1 Introduction

Electrical control of spins in systems with spin-orbit interactions is of basic interest and has great potential in semiconductor spintronics [18, 38, 39]. In recent years, steady progress has been made towards realization of convenient semiconducting ferromagnets and spin injection into semiconductors from ferromagnetic metals [40, 41, 42, 43, 44, 45, 46, 47, 48]. The spin transport theory presented in this chapter is motivated generally by current interest in novel spin-related transport effects in semiconductors, and particularly by interest in various schemes that generate spin-polarized currents [49, 50, 51, 52, 53, 54, 55, 56, 57, 58]. Using a semiclassical wave packet approach, we find that the spin current can be expressed as the sum of several physically transparent terms which are grouped together in a Kubo formula description. As an example, we use our theory to derive an expression for the intrinsic spin-Hall conductivity [49, 50] of a hole-doped semiconductor.

Semiclassical formulations of transport theory exploit the smooth variation

of transport fields on atomic length scales. Previous semiclassical theories of *spin* transport [59, 60, 61, 62, 63] have not accounted explicitly for intrinsic spin-orbit interaction in the crystal apart from, occasionally, its role in the relaxation of non-equilibrium spin-polarizations. In this chapter, we apply the wave packet approach introduced by Sundaram and Niu [10], which captures the consequences of the wave vector dependence of the Bloch spinors, to treat spin transport in spin-orbit coupled bands. This wave packet approach has already been successful in describing the anomalous Hall effect in ferromagnetic semiconductors [15, 16] and transition metals [64], interpreting it as a consequence of the Berry phase correction to the group velocity induced by the intrinsic spin-orbit interaction. We show here that the Hall spin current in response to an electric field is non-zero even in paramagnetic systems and that, in addition to the Berry phase term evaluated in a recent paper [49], other contributions must be taken into account. First of all, there is a contribution from the electric field correction to the average spin orientation of a wave packet. In addition, there are also contributions from the spin dipole and torque dipole of a carrier, which arise from the fact that spin and torque distribution within a wave packet generally differ from that of the charge. Including all these contributions, we obtain a total semiclassical spin current which is in agreement with the Kubo formula expression for the same quantity. We show that nonequilibrium spin polarization near the sample edge is driven not by the spin current alone but by the sum of the spin current and torque dipole density.

This chapter is organized as follows. In section 2.2 we outline the theory of semiclassical spin dynamics. In section 2.3 we determine the form and magnitude of the spin current in the spherical four-band Luttinger model.

2.2 Semiclassical spin dynamics

The semiclassical dynamics of each spin-charge carrier in a non-degenerate band is described by a wave packet, whose charge centroid has coordinates $(\mathbf{r}_c, \mathbf{k}_c)$. Wave packet construction is thoroughly explained in Refs. [10, 65]. When expanded in the basis of Bloch eigenstates, the wave packet has the form:

$$|w\rangle = \int d^3k \; a(\mathbf{k}, t) e^{i\mathbf{k}\cdot\hat{\mathbf{r}}} |u(\mathbf{r}_c, \mathbf{k}, t)\rangle. \tag{2.1}$$

In the above, the wave functions $|\mathrm{u}\rangle$ contain corrections linear in the electric field. We will be concerned with a system in which only a constant uniform electric field is present. Making a convenient choice of gauge, this electric field can be included in the Hamiltonian through the electromagnetic vector potential $\mathbf{A}(\mathbf{r}, \mathrm{t})$ only. This results in a nonadiabatic mixing of the bands, so that the Bloch wave functions $|\mathrm{u_n}\rangle$ have the following form:

$$|u_n\rangle = |\phi_n\rangle - \sum_{m \neq n} \frac{\langle \phi_m | i\hbar \frac{d}{dt} | \phi_n \rangle}{\epsilon_n - \epsilon_m} |\phi_m\rangle, \tag{2.2}$$

where the $\{|\phi_\mathrm{n}\rangle\}$ are the unperturbed Bloch eigenstates. The only time dependence comes from the fact that $\mathbf{k}$ drifts under the action of the electric field, as in (3.1). The $|\mathrm{u_n}\rangle$ form a complete set and retain the Bloch periodicity. The function $a(\mathbf{k}, t)$ is a narrow distribution sharply peaked at $\mathbf{k}_\mathrm{c}$, and its phase specifies the center of charge position $\mathbf{r}_\mathrm{c}$. The size of the wave packet in k-space must be considerably smaller than that of the Brillouin zone. In real space, this implies that the wave packet must stretch over many unit cells.

In the presence of a uniform electric field, the semiclassical equations of

motion for a non-degenerate band read [10]:

$$\hbar\dot{\mathbf{r}}_c = \frac{\partial\varepsilon}{\partial\mathbf{k}_c} - e\mathbf{E}\times\mathbf{\Omega} \text{ and } \hbar\dot{\mathbf{k}}_c = e\mathbf{E}, \tag{2.3}$$

where e is the carrier charge, ε is the band dispersion, and $\mathbf{\Omega}$ is the Berry curvature of the Bloch state [10]. Henceforth, $\mathbf{k}_c$ will be abbreviated to $\mathbf{k}$. The effect of the electric field is thus twofold: it drives the center of the wave packet in k-space, and it gives rise to a non-adiabatic correction to the wave functions, which mixes the bands at each $\mathbf{k}$.

Following the strategy of Boltzmann transport theory, we consider a collection of particles described by a phase space distribution f($\mathbf{r}_c$, $\mathbf{k}$, t). This distribution can drift according to the semiclassical equations of motion (2.3), and can also change due to collisions:

$$\frac{\partial f}{\partial t} + \dot{\mathbf{r}}_c\cdot\frac{\partial f}{\partial\mathbf{r}_c} + \dot{\mathbf{k}}\cdot\frac{\partial f}{\partial\mathbf{k}} = \frac{df}{dt}|_{coll}. \tag{2.4}$$

The collision term on the right hand side may be modeled by relaxation times or more accurately by collision integrals as usual.

The spin density distribution is defined as

$$S(\mathbf{r},t) \equiv \int\int d^3k d^3r_c f(\mathbf{r}_c,\mathbf{k},t)\langle\delta(\mathbf{r}-\hat{\mathbf{r}})\hat{s}\rangle, \tag{2.5}$$

where $\hat{s}$ is an arbitrary component of the spin operator, and the bracket indicates quantum mechanical average over the wave packet with charge centroid ($\mathbf{r}_c, \mathbf{k}$). Further analysis of this distribution will be facilitated by making the analogy with the standard coarse graining of electrodynamics in material media [67]. Our wave packets play the role of 'molecules', whose size will be taken as much smaller than the length scale of the distribution function. We are thus allowed to view the δ-function

in the above definition of the spin density as a sampling function with a width somewhere between the microscopic scale of the wave packets and the macroscopic scale of the distribution function. We can therefore write it as $\delta[(\mathbf{r}-\mathbf{r}_c)-(\hat{\mathbf{r}}-\mathbf{r}_c)]$ and expand it around $\mathbf{r}_c$, keeping only the first order term. Performing the integral over $\mathbf{r}_c$, the spin density can be re-expressed in the following form:

$$S = \int d^3k f\langle\hat{s}\rangle - \nabla\cdot\int d^3k f\mathbf{p}^s \tag{2.6}$$

where f = f($\mathbf{r}$, $\mathbf{k}$, t), and $\mathbf{p}^s = \langle(\hat{\mathbf{r}}-\mathbf{r}_c)\hat{s}\rangle|_{\mathbf{r}_c=\mathbf{r}}$ is the spin dipole. The two terms can be regarded as monopole and dipole contributions. The second term is analogous to the contribution to the charge density in electrodynamics from the divergence of the polarization.

Spin is in general not conserved, and for what follows it will be useful to define a quantity, which we shall call the torque density, in order to include the rate of change of spin into our discussion of transport:

$$\mathcal{T}(\mathbf{r},t) \equiv \int\int d^3k d^3r_c f(\mathbf{r}_c,\mathbf{k},t)\langle\delta(\mathbf{r}-\hat{\mathbf{r}})\hat{\tau}\rangle, \tag{2.7}$$

in which $\hat{\tau}$ is understood as $\frac{i}{\hbar}[\hat{H},\hat{s}]$ and $\hat{H}$ is the Hamiltonian. Following the steps outlined above, the torque density becomes:

$$\mathcal{T} = \int d^3k f\langle\hat{\tau}\rangle - \nabla\cdot\int d^3k f\mathbf{p}^\tau, \tag{2.8}$$

with the torque dipole $\mathbf{p}^\tau = \langle(\hat{\mathbf{r}}-\mathbf{r}_c)\hat{\tau}\rangle|_{\mathbf{r}_c=\mathbf{r}}$.

We evaluate the spin current using the microscopic spin current operator and the semiclassical distribution function:

$$\mathbf{J}^s(\mathbf{r},t) \equiv \int\int d^3k d^3r_c\ f(\mathbf{r}_c,\mathbf{k},t)\langle\delta(\mathbf{r}-\hat{\mathbf{r}})\dot{\hat{\mathbf{r}}}\hat{s}\rangle. \tag{2.9}$$

Throughout this and the following chapters, symmetrization of products of non-commuting operators is implied. After expanding, the spin current takes the form:

$$\mathbf{J}^s = \int d^3k f \langle \dot{\hat{\mathbf{r}}} \hat{s} \rangle - \nabla \cdot \int d^3k f \langle (\hat{\mathbf{r}} - \mathbf{r}) \dot{\hat{\mathbf{r}}} \hat{s} \rangle \tag{2.10}$$

For a homogeneous system, where the distribution function is independent of position, the gradient term vanishes, and it is permissible to use Bloch states (which may be regarded as the limit of very wide wave packets) to evaluate the carrier spin current $\langle \dot{\hat{\mathbf{r}}} \hat{\mathrm{s}} \rangle$. Since the Bloch states contain a first order correction in the field, this can in general yield an overall linear-in-field spin current even with the equilibrium distribution function. This intrinsic spin current has been evaluated for a number of systems recently [50], and identical results are obtained with the semiclassical approach developed here.

To illuminate the underlying physics, we decompose the carrier spin current into a number of terms:

$$\mathbf{J}^s = \int d^3k f (\dot{\mathbf{r}}_c \langle \hat{s} \rangle + \frac{d\mathbf{p}^s}{dt} - \mathbf{p}^\tau). \tag{2.11}$$

The first contribution is convective, arising from the fact that the total spin is transported as the wave packet moves. The second comes from the rate of change of the spin dipole, while the third is from the torque dipole. This decomposition makes it possible to compare the Kubo formula result with those based on various heuristic arguments. The authors of [49] restricted their scope to the convective term and considered only the Berry phase contribution to the charge center velocity $\dot{\mathbf{r}}_\mathrm{c}$. The present semiclassical decomposition allows us to recognize the missing terms due to the spin and torque dipoles, as well as a field correction to the carrier spin in the convective term. The approach of [49] would give a zero result for the Rashba model, whereas the Kubo formula approach of [50], which agrees with (2.11), yields

a non-zero spin Hall current for this model. Interestingly, the spin Hall current in the Rashba model can be obtained exactly from a heuristic argument based on a velocity and field dependent correction to the carrier spin as discussed in [50]. This approach is, however, applicable only to single-band models with wave vector-dependent Zeeman coupling.

The spin density and current satisfy the following equation of continuity:

$$\frac{\partial S}{\partial t} + \nabla \cdot \mathbf{J}^s = \mathcal{T} + \int d^3k \frac{df}{dt} \langle \hat{s} \rangle. \tag{2.12}$$

It is seen that the torque density appears in the source, accounting for the spin non-conserving terms in the Hamiltonian, and acting as a bulk mechanism for spin generation. The second term accounts for the effect of collisions.

The source can be decomposed into intrinsic and extrinsic contributions, depending on the equilibrium and non-equilibrium parts of the distribution respectively. If we restrict our attention to homogeneous sytems the torque density is simply $f\langle \hat{\tau} \rangle$. We find that this term is always first order in the electric field, and is given by $f\frac{e}{\hbar}\mathbf{E} \cdot \frac{\partial \langle \hat{s} \rangle}{\partial \mathbf{k}}$. We are thus justified in replacing f by its equilibrium value f_0, in which case this term is purely intrinsic. The second term in the source, which depends on the non-equilibrium shift in the distribution function, is entirely extrinsic.

2.3 Spin current in the Luttinger model

Our formalism so far applies to independent non-degenerate bands, and for the Rashba model its predictions are in agreement with [50]. There exists a parallel formalism for coupled degenerate bands, which yields the same results as given above and will be outlined in Chapter 5. In this case, the distribution function becomes a density matrix, while $\langle \hat{s} \rangle$, $\langle \hat{\tau} \rangle$, $\dot{\mathbf{r}}_c$, $\mathbf{p}^s$ and $\mathbf{p}^\tau$ are replaced by matrices. To find the macroscopic expectation values, one traces over the density matrix.

This formalism can be applied, for example, to the spherical four-band Luttinger Hamiltonian,

$$H_0 = \frac{\hbar^2}{2m}[(\gamma_1 + \frac{5}{2}\gamma_2)k^2 - 2\gamma_2(\mathbf{k}\cdot\hat{\mathbf{J}})^2], \tag{2.13}$$

where $\hat{\mathbf{J}}$ is the operator for angular momentum 3/2 and γ_1, γ_2 are the Luttinger parameters. The Bloch states are eigenstates of the angular momentum projection in the $\mathbf{k}$ direction, $\hat{\mathrm{J}}_\mathrm{k}$. The four bands are split (for finite k) into two degenerate manifolds with $\hat{\mathrm{J}}_\mathrm{k} = \pm 3/2$ (heavy holes) and $\hat{\mathrm{J}}_\mathrm{k} = \pm 1/2$ (light holes).

Let us take a closer look at the source term, using the four-band model as an illustration. This discussion applies to either of the heavy and light-hole manifolds. In equilibrium, the density matrix is diagonal and equal to f_0 for each band. The mean spin in the z-direction is $\langle \hat{\mathrm{s}}_\mathrm{z}\rangle_{\pm 3/2} = \pm\frac{\hbar \mathrm{k}_\mathrm{z}}{2\mathrm{k}}$ for the heavy holes, and it is $\langle \hat{\mathrm{s}}_\mathrm{z}\rangle_{\pm 1/2} = \pm\frac{\hbar \mathrm{k}_\mathrm{z}}{6\mathrm{k}}$ for the light holes. The spin expectation values have opposite signs in the two bands, so that, when averaged over the equilibrium density matrix, the intrinsic term in the source, $\mathrm{f}\langle\hat{\tau}\rangle$, will vanish. The intrinsic source $\mathcal{T}$ therefore vanishes in the bulk for this system.

In the relaxation time approximation, the collision term in (2.4) is given by [68]:

$$\frac{df}{dt}|_{coll} = \frac{f_0 - \frac{1}{2}Trf}{\tau_p} I - \frac{Tr(f\sigma)\cdot\sigma}{\tau_s}, \tag{2.14}$$

where τ_p and τ_s are the momentum and spin relaxation times respectively, I is the identity matrix and σ is the vector of Pauli spin matrices. In the extrinsic term in the source, the part depending on the momentum relaxation time will also cancel between the two bands, leaving us with just the contribution coming from the second term on the right hand side of (2.14). The equation of continuity is then:

$$\frac{\partial S}{\partial t} + \nabla\cdot\mathbf{J}^s = \frac{-S}{\tau_s} - \nabla\cdot\mathbf{P}^\tau, \tag{2.15}$$

where $\mathbf{P}^\tau = \int \mathrm{d}^3\mathrm{k}\mathrm{f}\mathbf{p}^\tau$ is the torque dipole density. The two divergences will vanish

in the bulk if the sample is homogeneous.

We will now take a closer look at the spin current, making further use of the four-band model for the spin-orbit coupled valence bands of a zincblende semiconductor. In previous work [55, 56, 57], extensive discussions have been devoted to the extrinsic part of the spin current, which is given by the zero-field carrier velocity and spin integrated over the non-equilibrium part of the distribution. Here we will concentrate on the intrinsic part of the spin current, coming from the field correction to the carrier spin current integrated over the equilibrium distribution. In order for this term to be dominant, scattering must be strong enough to keep the distribution function close to its equilibrium value, and weak enough to limit inter-band mixing. This is therefore opposite to the limit of Dyakonov-Perel [63] relaxation of spin in the weakly spin-orbit split bands of crystals.

The two-fold degeneracy of both the heavy and light hole manifolds implies that, in the presence of an electric field, however weak it may be, mixing within the degenerate manifold will occur in general. Fortunately, for the heavy holes the $\hat{\mathrm{J}}_{\mathrm{k}} = \pm 3/2$ bands do not mix to first order in the electric field, and we can apply the nondegenerate band formalism. The $\mathrm{s_z}$ dipole moment for the heavy holes is found to be $\mathbf{p}_{\mathrm{h}}^{\mathrm{s_z}} = -\frac{\hbar \mathbf{k} \times \hat{\mathbf{z}}}{4\mathrm{k}^2}$. The torque dipole is, after an angular average, $\mathbf{p}_{\mathrm{h}}^{\tau_z} = -\frac{\mathrm{e}\mathbf{E} \times \hat{\mathbf{z}}}{6\mathrm{k}^2}$. The spin and torque dipoles are equal for both heavy hole bands. For the convective part of carrier spin current, in addition to a field correction to the carrier velocity due to the Berry phase [49], we obtain a term which is due to the change in the spin expectation value induced by the electric field and has the form $\frac{1}{\hbar}\frac{\partial \varepsilon}{\partial \mathbf{k}}\mathbf{E} \cdot \frac{\partial \langle \hat{\mathrm{s}} \rangle}{\partial \mathbf{E}}$. Using these results, we find the current for spin-z component of a heavy-hole carrier to be

$$\mathbf{j}_h^s = (\frac{1}{4k^2} - \frac{\hbar^2}{6m_h\Delta} - \frac{1}{12k^2} + \frac{1}{6k^2})e\mathbf{E} \times \hat{\mathbf{z}}, \tag{2.16}$$

where $\Delta = \epsilon_{\mathrm{h}} - \epsilon_{\mathrm{l}}$ is the energy difference between the heavy and light holes. The first two terms come from the convective part due to field corrections to the carrier

velocity and spin respectively. The third term comes from the rate of change of the spin dipole, while the last one comes from the torque dipole. The heavy-hole carrier spin current can be simplified to:

$$\mathbf{j}_h^s = \frac{e\mathbf{E}\times\hat{\mathbf{z}}}{3k^2}\frac{1}{1-\frac{m_l}{m_h}}. \tag{2.17}$$

For the light holes, we must consider field induced mixing between the two degenerate bands. The spin current per carrier in the light hole manifold has the same form as for the heavy holes, and differs only by a minus sign. Integrating over k and summing the contributions from all four bands, we arrive finally at the following expression for the total spin current: $\mathbf{J}^{\mathrm{s}} = \sigma_{\mathrm{SH}}\mathbf{E}\times\hat{\mathbf{z}}$, where the spin-Hall conductivity is given by:

$$\sigma_{SH} = \frac{e}{3\pi^2}\frac{k_h - k_l}{1-\frac{m_l}{m_h}} = \frac{e}{3\pi^2}\frac{k_h}{1+\sqrt{\frac{m_l}{m_h}}} \tag{2.18}$$

where $\mathrm{k_h}$ and $\mathrm{k_l}$ are the Fermi wavevectors for the heavy and light holes respectively. Separate calculations based on the Kubo formula by the present authors and by Murakami *et al.* [69] yield the same results.

Finally, we comment on the relationship between bulk spin currents and spin accumulation near the edge of the sample. A theory of spin accumulation must start from the spin density continuity equation (2.15). If the torque dipole density $\mathbf{P}^{\tau}$ were absent from this equation, then in the steady state the spin accumulation would be due only to the spin current. The presence of the torque dipole density modifies the expression for the spin accumulation, giving that

$$\int S dx = \tau_s(J_x^s + P_x^{\tau}). \tag{2.19}$$

We have already discussed the response of the spin current to an electric field, and

after a similar calculation for the torque density we find that

$$J_x^s + P_x^\tau = \frac{eE_y}{3\pi^2}\frac{k_h}{1+\sqrt{\frac{m_l}{m_h}}}(\frac{1}{2} - \frac{m_l}{2m_h} - \sqrt{\frac{m_l}{m_h}}). \tag{2.20}$$

Using n=5×10^{18} cm^{-3} and an electric field of 20000V/cm as typical values, the spin current is -5.4×10^{28} spins per unit area per second. We take the spin relaxation time to be $\tau_s = 30$ps [70] and a unit cell size of 6.3$\mathring{A}$, and we obtain a spin accumulation of 0.64 spins per *unit cell area*. This is a measurable effect as discussed in [49, 50].

Chapter 3

Electrical generation of spin in crystals with reduced symmetry

3.1 Introduction

The considerable amount of attention devoted in recent years to the field of semiconductor spintronics has been summarized in the previous chapter. Semiconductors can be used in the same spintronic devices as metals, and have the additional advantages brought about by the existence of an adjustable bandgap and by the ability to manipulate the carrier density over many orders of magnitude by doping, gating and heterojunction formation. Moreover, spin accumulation in a semiconductor will generate a much larger voltage because the density of states at the Fermi energy is lower than in a metal, leading to a larger spin splitting [71].

To this end it would be extremely desirable if a practical method existed for the efficient generation of spin polarization inside a semiconductor, as well as for the transport of spins over a sizable length scale. Although optical spin injection [72] has been known for decades, it is impractical for devices, since it is not sufficiently local for nanoelectronics. On the other hand, spin injection from a ferromagnetic

metal into a semiconductor requires long spin lifetimes [73] and is impeded by the resistivity mismatch between the two materials [40]. Progress has been achieved with ferromagnetic semiconductors [41, 42, 43, 44, 45, 47, 48, 74], but room temperature semiconductor ferromagnetism [75] has not yet been clearly established. Recent research has focused on the electrical control of spins in semiconductors [18, 38, 39], including theoretical work on the possibility of observing a spin-Hall effect [66, 76, 49], together with various efforts to generate a spin current [77, 78, 79, 80, 81, 82].

Following this energetic theoretical enterprise, recent experimental work [83, 84, 85, 86] has demonstrated the detection of a sizable spin accumulation in semiconductors to be a feasible task. Motivated by these findings, in this chapter we present a general theory of the intrinsic electrical spin generation that occurs generically in spin-orbit coupled systems that are not inversion-symmetric. We will concentrate on systems where the spin-orbit interactions are strong. The existence of the mechanism we discuss has been pointed out in the previous literature [87, 88, 89, 90]. Our approach is to be contrasted with earlier theories which consider directly the response of a spin polarization to an electric field, by Aronov *et al.* [89] for example. When calculating the spin density directly via linear response it is, in general, not possible to separate the intrinsic generation terms from the extrinsic scattering effects, thus physical transparency is often sacrificed. On the other hand, in the present work we recognize spin generation as an intrinsic effect so that it may be determined from first principles calculations. The interplay of the spin generation and relaxation terms, resulting in a finite spin polarization, emerges in the final analysis, as pointed out by Edelstein [88]. So far, experiment [83] has attempted to measure the spin generation rate separately from the spin relaxation time in the *weak* spin-orbit coupling limit. The idea we discuss has already been applied to the Rashba Hamiltonian [90], which we also examine below.

Spin-orbit interactions can be important in semiconductors for several rea-

sons. The first is the fact that the carriers are clustered near the band extrema around high symmetry points where there exist degeneracies, and the form of these degeneracies is determined by the spin-orbit interaction. Due to the fact that in semiconductors the carriers occupy a narrow width of k-space the spin-orbit interaction can therefore play a crucial role. Because of the fact that spin is not conserved, there exists a term which acts as a bulk source of spin generation, which has been introduced in the previous chapter. It represents the rate of change of the spin density in response to an external electric field.

Intrinsic spin-orbit effects have been shown to lead to a non-zero Berry curvature which gives a contribution to the anomalous Hall effect [15, 16, 64, 91], while a spin-orbit-induced metal-insulator transition has been detected by Koga *et al.* [92]. Aside from the references mentioned above, such intrinsic spin-orbit effects have not been taken into account previously, although several theories exist which account for the role intrinsic spin-orbit effects play in spin relaxation [60, 61, 62]. Moreover, a number of articles have considered an electric-field-induced rotation of the total spin polarization [59, 93, 94], which is distinct from the electrical spin generation we discuss in our work. In the situations we consider, the total spin polarization is initially zero, therefore it cannot undergo a rotation.

Because both the rate of change of the spin density and the electric field are even under time reversal, ferromagnetism is not required for spin generation. Nevertheless, because the electric field changes sign under spatial inversion while the rate of change of the spin density does not, spin generation occurs only in crystals with broken inversion symmetry. Further details of these symmetry arguments will be given in the last section.

This chapter is organized as follows. In section 3.2 we will present the formalism underlying our discussion of intrinsic spin generation. The mechanism we outline applies to a wide class of systems, and to illustrate this we discuss, in section

3.3, two-dimensional heterostructures described by the Rashba model, followed by a model of the conduction bands of unstrained bulk wurtzite structures and strained bulk n-GaAs. Finally, 3.4 is devoted to symmetry considerations.

3.2 Formalism

In the physical picture we consider, the spin-orbit interaction has been taken into account in the band structure. We adopt the Boltzmann-wave packet approach described in Chapter 2. For convenience, we will reproduce the semiclassical equations of motion for $(\mathbf{r}_c, \mathbf{k})$, in the presence of a constant uniform electric field $\mathbf{E}$:

$$\begin{aligned} \hbar\dot{\mathbf{k}}_c &= -e\mathbf{E} \\ \hbar\dot{\mathbf{r}}_c &= \frac{\partial\varepsilon_n}{\partial\mathbf{k}_c} + e\mathbf{E}\times\mathbf{\Omega}_n, \end{aligned} \tag{3.1}$$

where $\mathbf{\Omega}_\mathrm{n} = 2\Im\langle\frac{\partial \mathrm{u_n}}{\partial\mathbf{k}}| \times |\frac{\partial \mathrm{u_n}}{\partial\mathbf{k}}\rangle$ represents the Berry, or geometrical, curvature.

The theory presented in this chapter is a theory of spin accumulation in the strong spin-orbit coupling limit, implying that the splitting of the bands due to the spin-orbit interaction exceeds their broadening due to disorder at all wave vectors. As a result, the bands are well defined although, due to the presence of the spin-orbit interaction, they are not pure spin-up and spin-down bands. The distribution function corresponding to each band n, $\mathrm{f_n}$, can drift according to the semiclassical equations of motion, and can also change due to collisions. As mentioned in Chapter 2, the time evolution of the distribution function is governed by the Boltzmann equation,

$$\frac{\partial f_n}{\partial t} + \dot{\mathbf{r}}_c\cdot\frac{\partial f_n}{\partial\mathbf{r}_c} + \dot{\mathbf{k}}\cdot\frac{\partial f_n}{\partial\mathbf{k}} = \frac{df_n}{dt}|_{coll}. \tag{3.2}$$

For electrons in equilibrium, the solution of the Boltzmann equation is $\mathrm{f_n^{(0)}}$, the Fermi-Dirac distribution, while in a general nonequilibrium situation $\mathrm{f_n}$ can be written as $\mathrm{f_n^{(0)}} + \delta\mathrm{f_n}$. In the presence of a constant uniform electric field $\mathbf{E}$ the shift $\delta\mathrm{f_n}$

is given, in the relaxation time approximation, by the well-known result [65]:

$$\delta f_n = f_n - f_n^{(0)} = e\tau_p \mathbf{E} \cdot \mathbf{v}_n \frac{\partial f_n^{(0)}}{\partial \varepsilon}, \tag{3.3}$$

where $\mathbf{v}_{\mathrm{n}} = \frac{1}{\hbar}\frac{\partial \varepsilon_{\mathrm{n}}}{\partial \mathbf{k}}$, ε_{n} is the energy of band n, and τ_{p} is the momentum relaxation time. We refer to terms which depend only on the equilibrium value of the distribution function, $\mathrm{f}_{\mathrm{n}}^{(0)}$, as *intrinsic*, as opposed to *extrinsic* terms, depending on the non-equilibrium distribution, and therefore on scattering.

The study of spin generation necessarily relies on the spin equation of continuity. For the case we consider, we have shown in the previous chapter that this equation takes the form:

$$\frac{\partial S_n}{\partial t} + \nabla \cdot \mathbf{J}_n^s = \mathcal{T}_n + \int d^3k \frac{df_n}{dt} \langle \hat{s} \rangle_n. \tag{3.4}$$

The terms on the LHS represent the spin density and current in band n, while the last term in the equation takes into account collisions. The abbreviation $\langle \hat{\mathrm{s}} \rangle_{\mathrm{n}}$ stands for the expectation value in band n of the spin operator corresponding to any one component of the spin. We specialize henceforth in homogeneous systems, in which the divergence of the spin current in (3.4) will be zero and the equation of continuity may be written as:

$$\frac{\partial S_n}{\partial t} = \mathcal{T}_n + \int d^3k \frac{(f_n^{(0)} - f_n)}{\tau_p} \langle \hat{s} \rangle_n, \tag{3.5}$$

where τ_{p} is the relaxation time. The first term on the RHS, which accounts for spin generation in the absence of collisions, is the focus of this chapter. This term, which we shall call the torque density, exists due to the fact that spin is in general not conserved and therefore the average spin of a wave packet is not constant in time.

As discussed in Culcer *et al.* [66], this torque density is defined as:

$$\mathcal{T}_n(\mathbf{r},t) \equiv \int d^3r_c \int d^3k f_n(\mathbf{r}_c,\mathbf{k},t)\langle\hat{\tau}\delta(\hat{\mathbf{r}}-\mathbf{r})\rangle_n, \tag{3.6}$$

in which $\hat{\tau}$ is understood as $\frac{\mathrm{i}}{\hbar}[\hat{\mathrm{H}},\hat{\mathrm{s}}]$, $\hat{\mathrm{H}}$ is the Hamiltonian, $\hat{\mathbf{r}}$ is the quantum-mechanical position operator and $\langle\rangle_{\mathrm{n}}$ stands for the expectation value in a wave packet constructed starting from the eigenfunction corresponding to band n. We note that the equation of continuity (3.4) is derived directly from the first-principles definitions of $\mathrm{S_n}$, $\mathbf{J}^{\mathrm{s}}_{\mathrm{n}}$ and $\mathcal{T}_{\mathrm{n}}$. In homogeneous systems the torque density simplifies to

$$\mathcal{T}_n = \int d^3k f_n\langle\hat{\tau}\rangle_n, \tag{3.7}$$

which will be referred to as the *spin generation term.* The fact that we are considering homogeneous systems also implies, as in Chapter 2, that we may regard the wave packets as being wide in real space, thus sharp in k-space, and evaluate the expectation value $\langle\hat{\tau}\rangle_{\mathrm{n}}$ using Bloch wave functions. It should be pointed out that, although we arrive at our results semiclassically, one does not require a local description to obtain them, and they can be found using, for example, a Kubo formula approach. The connection to this latter approach will be detailed in what follows. Moreover, our theory is not restricted to the generation of spin, since $\langle\hat{\mathrm{s}}\rangle_{\mathrm{n}}$ may represent the wave packet expectation value of any component of any other non-conserved observable.

As mentioned in the previous chapter, we are concerned with a system in which only a constant uniform electric field is present. Making a convenient choice of gauge, this electric field can be included in the Hamiltonian through the electromagnetic vector potential $\mathbf{A}(\mathbf{r},\,\mathrm{t})$ only. This results in a nonadiabatic mixing of the

bands, so that the Bloch wavefunctions $|u_n\rangle$ have the following form:

$$|u_n\rangle = |\phi_n\rangle - \sum_{m \neq n} \frac{\langle \phi_m | i\hbar \frac{d}{dt} | \phi_n \rangle}{\epsilon_n - \epsilon_m} |\phi_m\rangle, \tag{3.8}$$

where the $\{|\phi_n\rangle\}$ are the unperturbed Bloch eigenstates. The only time dependence comes from the fact that $\mathbf{k}$ drifts under the action of the electric field, as in (3.1). Therefore it is legitimate to make the replacement $\frac{d}{dt} = -\frac{e\mathbf{E}}{\hbar} \cdot \frac{\partial}{\partial \mathbf{k}}$ in (3.8). In this way, the wave functions $|u_n\rangle$ depend on the electric field through a reactive term, in other words the field induces a change in the wave functions at each $\mathbf{k}$. The $\{|u_n\rangle\}$ form a complete set. Moreover, the expectation value $\langle \hat{s} \rangle_n \equiv \langle u_n | \hat{s} | u_n \rangle$ is a function of $\mathbf{k}$ only, and its time dependence arises implicitly through its dependence on the wave vector. We have thus included the effect of the electric field in mixing the bands but neglected inter-band scattering.

In the limit of wide wave packets, it is straightforward to prove, starting from Eq. (3.8), that $\langle \hat{\tau} \rangle_n$ evaluated in the $\{|u_n\rangle\}$ basis is equal to $\frac{d\langle \hat{s} \rangle_n}{dt} \equiv \dot{\mathbf{k}} \cdot \frac{\partial \langle \hat{s} \rangle}{\partial \mathbf{k}}$, where $\langle \hat{s} \rangle$ is evaluated in the unperturbed $\{|\phi_n\rangle\}$ basis and $\dot{\mathbf{k}} = -\frac{e\mathbf{E}}{\hbar}$ from (3.1). The former approach is equivalent to using the Kubo formula to find the response of $\hat{\tau}$ to an electric field. Following this line of thought, we find that the spin generation term is always at least first order in the electric field. Then, to first order in the electric field, we may replace f_n by its equilibrium value $f_n^{(0)}$. This spin generation term is then purely *intrinsic*, as defined in Chapter 2 and at the beginning of this section. The final form of this term is:

$$\mathcal{T}_n = -\frac{e\mathbf{E}}{\hbar} \cdot \int d^3k f_n^{(0)} \frac{\partial \langle \hat{s} \rangle_n}{\partial \mathbf{k}}. \tag{3.9}$$

Our theory thus shows that there exists a spin generation rate which can be interpreted as due to a displacement of the wave packet in k-space, but also as the expectation value of the operator $\hat{\tau}$ in a state which is not an eigenstate of the

crystal Hamiltonian. Experiment [83] has been attempting to measure this rate of generation, but for the time being success has been achieved only in the weak spin-orbit coupling limit [1].

In the steady state the equation of continuity becomes simply:

$$\int d^3k \frac{(f_n - f_n^{(0)})}{\tau_p} \langle \hat{s} \rangle_n = \mathcal{T}_n \tag{3.10}$$

Remembering that $\int \mathrm{d^3k f_n^{(0)}} \langle \hat{\mathrm{s}} \rangle_\mathrm{n} = 0$ and assuming a momentum relaxation time independent of wave vector, the LHS is simply $\frac{\mathrm{S_n}}{\tau_\mathrm{p}}$. We can then rewrite the above equation to express the steady state spin density as:

$$S_n = \mathcal{T}_n \tau_p \tag{3.11}$$

The characteristic time governing the relaxation of the spin density distribution is the momentum relaxation time. This is to be expected, since the theory describes non-degenerate, well defined bands.

We note that the torque term must be present even in a clean system if the Hamiltonian contains spin-non-conserving terms. In the presence of scattering mechanisms, the *intrinsic* spin *generation* term is balanced by the *extrinsic* spin *relaxation* term so that a net spin polarization can be reached in the steady state. In addition, in the systems we consider, we assume scattering is strong enough to keep the distribution function near equilibrium and the scattering time small, but not strong enough to make inter-band mixing important.

[1] The experiment described in Ref. [83] uses a magnetic field to rotate the spins in the z direction, then measures the accumulation of $\mathrm{s_z}$. This procedure is appropriate in the weak spin-orbit limit. In the limit of strong spin-orbit interaction, if one attempts to rotate the spins, one must also take into account the effect of the spin-orbit interaction on the spins.

3.3 Examples

In order to clarify the significance of the examples presented below, we start with some comments regarding the spin-orbit interaction and asymmetry. The terms in the spin-orbit interaction which are odd in k rely upon the inversion asymmetry of the system under study. This asymmetry can be of two kinds, depending on the dimensionality of the system. In three dimensions, the inversion asymmetry is a property of the underlying material, and is referred to as bulk inversion asymmetry (BIA). In two dimensions, an asymmetric confinement potential can provide an additional source of inversion asymmetry, known as structure inversion asymmetry (SIA). Moreover, the application of strain along a particular direction further reduces the symmetry of the structure, with important consequences which will be examined below.

3.3.1 Rashba-type spin-orbit interaction

Our theory allows us to treat bulk semiconductors and quantum wells on the same footing. We begin with a study of the Rashba Hamiltonian. This Hamiltonian describes, based on symmetry arguments, the SIA of quantum well or heterojunction based two dimensional electron systems and is usually the dominant source of spin-orbit coupling in these systems. The effective Hamiltonian has the form:

$$H = \frac{\hbar^2 k^2}{2m} + \alpha(\sigma \times \mathbf{k}) \cdot \hat{\mathbf{z}}, \tag{3.12}$$

in which α is the spin orbit constant and σ is the vector of Pauli spin matrices. As described in Chapter 1, the spin-orbit constant α is usually taken to have the form $a_{46}E_z$, where a_{46} is a material specific parameter while E_z is the component of the electric field in the z-direction. This electric field is in general a function of position in the quantum well, and there exists in principle an additional contribution if the

interface on one side of the quantum well is different from the interface on the other side. This effective Hamiltonian therefore provides only an approximate description of the Rashba effect. The magnitude of the Rashba interaction can be tuned by an external gate voltage by an amount which has been shown to be as much as 50% [95]. It is customary to view the term multiplying the spin as a momentum-dependent effective magnetic field [50] in which the spin precesses. In the absence of an external magnetic field, the bands in the Rashba model are degenerate at $\mathbf{k} = 0$, and each band contains the same number of spin up and spin down carriers [26]. The Hamiltonian has eigenvalues $\epsilon_{\pm} = \frac{\hbar^2 \mathrm{k}^2}{2\mathrm{m}} \pm \alpha \mathrm{k}$, which will be labeled by + and - respectively. The Berry curvatures are $\mathbf{\Omega}_{\pm} = \mp \frac{1}{2} \lim_{\mathrm{H}\to 0} \frac{\alpha^2 \mathrm{H}\hat{\mathbf{z}}}{(\alpha^2 \mathrm{k}^2 + \mathrm{H}^2)^{3/2}}$. The spin generation term takes the following form:

$$\langle \hat{\tau} \rangle_{\pm} = \mp \frac{e\mathbf{k}}{2k^3} (\mathbf{k} \times \mathbf{E}) \cdot \hat{\mathbf{z}}, \tag{3.13}$$

with $\langle \hat{\tau} \rangle$ defined in Eq.(3.7). Interestingly, the spin torque does not depend on the spin-orbit constant. However, the total torque term, summed over the two bands, depends on the difference in Fermi wave vectors, which is proportional to the spin-orbit constant. We find that

$$\mathcal{T} = \frac{e\alpha m}{4\pi\hbar^2} \mathbf{E} \times \hat{\mathbf{z}}, \tag{3.14}$$

which vanishes in the limit in which $\alpha \to 0$, is in agreement with Magarill [90]. This result does not depend on the number density and has a universal form, but it should be noted that its validity is not universal, rather it is restricted to systems in which disorder is weak as discussed at the end of the previous section.

Using symmetry arguments, we find that the spin-orbit coupling in the conduction band of bulk wurtzite structures is also described by a Rashba-type Hamiltonian, with a spin-orbit constant defined analogously. This conclusion is supported by group theory arguments [35]. The only terms linear in k (and therefore dominant

except at high densities) which are allowed by symmetry are $\beta(\sigma_x k_y - \sigma_y k_x)$, and the Hamiltonian is:

$$H = \frac{\hbar^2 k^2}{2m} + \beta(\sigma_x k_y - \sigma_y k_x) \tag{3.15}$$

with eigenvalues (labeled as before) $\frac{\hbar^2 k^2}{2m} \pm \beta k$. The spin generation term has a form very similar to (3.13):

$$\langle \hat{\tau} \rangle_{\pm} = \mp \frac{e\mathbf{k}_\perp}{2k_\perp^3} (\mathbf{k}_\perp \times \mathbf{E}) \cdot \hat{\mathbf{z}} \tag{3.16}$$

In the above, $\mathbf{k}_\perp = (k_x, k_y, 0)$. The total torque term is:

$$\mathcal{T} = \frac{em\beta}{4\pi^2\hbar^2} (3\pi^2 n)^{1/3} \mathbf{E} \times \hat{\mathbf{z}} \tag{3.17}$$

which again vanishes as $\beta \to 0$ and as the number density n$\to$ 0.

3.3.2 Cubic Dresselhaus spin-orbit interaction

Finally, we turn our attention to the conduction band of zincblende semiconductors, which has been the focus of experiment recently [83]. In order to be close to experiment, we consider a degenerate electron gas in an n-doped $In_xGa_{1-x}As$ heterostructure grown on GaAs, with x=0.07, with a strain of 0.46% directed along (001), as given in the recent paper of Kato *et al.* [83] and references therein. The Fermi surface is only slightly displaced from equilibrium, as revealed by the drift velocities and by the mobility measurements in the supplementary table. We will not, however, attempt to simulate experiment, as many aspects seem to remain incompletely understood. For example, it is known that the strain tensor acquires off-diagonal shear terms [96], but experiment does not so far provide an unambiguous way of determining their role [83, 84, 97]. Therefore, we will consider an idealized situation in which the x and y lattice constants are equal to their substrate values, while the lattice constant in the z-direction expands according to the elastic equations. Moreover, there is no convincing pattern in the variation of the BIA with

increasing strain in the experiment. In fact, the only sizable increase is observed at the largest value of the strain, which is 0.46%. Thus, it is not clear whether the linear or cubic term in the BIA is dominant in the samples investigated.

The symmetry of zincblende does not allow terms linear in k in the conduction band in the bulk. As a result, when strain is applied these linear-in-k terms will be first order in the strain. These terms will be important at small wave vectors, but we will concentrate on situations in which the number density makes the Fermi wave-vector k_F high enough that the cubic Dresselhaus term is dominant. We therefore neglect the effect of the terms linear in k in this calculation, although we take into account the effect of strain on the effective masses. We will take into account only the spin-orbit terms cubic in k which are present in the unstrained lattice [26], namely $\lambda\sigma_x[k_x(k_y^2 - k_z^2) + \text{c.p.}]$, where λ is the spin-orbit constant and c.p. stands for cubic permutations. To determine the range of validity of this approach, we estimate the doping density n at which the k^3 term dominates the linear-in-k BIA and SIA terms. Based on the data in Kato *et al.* [83, 84], we estimate that the term cubic in k will dominate if $n \geq 2.7 \times 10^{16}$ cm^{-3}, which puts the experiment, in which $n = 3 \times 10^{16}$ cm^{-3}, narrowly in the range in which this term is dominant. Our prediction is, however, consistent with the bulk BIA findings of Kato *et al.*

Our theory is valid in the limit of strong spin-orbit interactions or weak disorder. For our theory to be valid, the following must hold:

$$\frac{\hbar}{\tau_p} < \Delta_{so}(k_F), \tag{3.18}$$

where τ_p is the momentum relaxation time and $\Delta_{so}(k_F)$ is the spin-orbit splitting at the Fermi wave vector. From the mobility, τ_p is estimated at 0.22ps, yielding $\frac{\hbar}{\tau_p} = 2.9$ meV. In order for the spin-orbit to overwhelm this, the number density must exceed 3.5×10^{18} cm^{-3}, which is a more stringent requirement than the requirement that the cubic Dresselhaus term exceed the linear one. In other words, when the

system is in the strong spin-orbit limit it is already in the regime where the k^3 term dominates. This calculation also shows that the experiment of Kato *et al.* lies in the weak spin-orbit coupling limit, outside the validity limit of our theory. The experiment is performed in a doped semiconductor in the extrinsic regime, where extrinsic (as defined in section II) refers to situations in which scattering effects are dominant, that is the spin-orbit splitting at k_F is broadened beyond the point at which the bands overlap. By increasing n one increases k_F and therefore $\Delta_{so}(k_F)$, passing into the intrinsic regime.

The Hamiltonian for this system is:

$$H = \frac{\hbar^2 k_\perp^2}{2m_\perp} + \frac{\hbar^2 k_z^2}{2m_z} + \lambda \sigma_x [k_x(k_y^2 - k_z^2) + c.p.], \tag{3.19}$$

where $\mathbf{k}_\perp$ has been defined above. It has eigenvalues $\varepsilon_\pm = \frac{\hbar^2 k_\perp^2}{2m_\perp} + \frac{\hbar^2 k_z^2}{2m_z} \pm \lambda\Delta$ (labeled as before), with Δ given by $\sqrt{[k_x^2(k_y^2 - k_z^2)^2 + \text{c.p.}]}$. The Berry curvatures are given by $\mathbf{\Omega}_\pm = \pm \frac{(k_x^2 - k_y^2)(k_z^2 - k_x^2)(k_y^2 - k_z^2)\mathbf{k}}{2\Delta^3}$. In this model the x-component of spin takes the form:

$$\langle \hat{s}^x \rangle_\pm = \frac{\hbar k_x (k_y^2 - k_z^2)}{2\Delta}, \tag{3.20}$$

with the other components given by cubic permutations of this expression. The x-component of the spin generation term is:

$$\langle \hat{\tau}^x \rangle_\pm = \mp \frac{e\lambda E_x (k_y^2 - k_z^2)(k_y^2 k_z^4 + k_z^2 k_y^4 - k_x^4 k_y^2 - k_x^4 k_z^2)}{2\Delta^3}. \tag{3.21}$$

Again the other components can be found by cubic permutation. Note that, if strain were absent so that the effective mass would be isotropic, the band structure and thus the equilibrium distribution would have cubic symmetry, making $\langle \hat{\tau} \rangle$ zero after integration over k. As a result, we expect that in the strong spin-orbit coupling limit the spin generation will be proportional to the strain.

To facilitate comparison with experiment, it is more instructive to examine

the *electric torque response tensor*, which we define through the equation $\mathcal{T}_i = \chi^{\tau}_{ij} E_j$, and $\chi^{\tau}_{xx} = -\chi^{\tau}_{yy}$ have finite values, the other components being zero. In contrast to the Rashba model, the diagonal components of the tensor are finite, whereas the off-diagonals vanish. To obtain an explicit expression for the total spin torque, we must integrate over the Fermi surface, which in this case is ellipsoidal. We use the equilibrium distribution and the fact that $f(\varepsilon \pm \lambda\Delta) = f(\varepsilon) \pm \lambda\Delta \frac{\partial f(\varepsilon)}{\partial \varepsilon}$, with $\varepsilon = \frac{\hbar^2 k_\perp^2}{2m_\perp} + \frac{\hbar^2 k_z^2}{2m_z}$ and $-\frac{\partial f(\varepsilon)}{\partial \varepsilon} = \delta(\varepsilon - \varepsilon_F)$. We write the mass ratio as $\frac{m_z}{m_\perp} = 1 + \gamma$, where γ is a small quantity, and we find that for the system under study $\gamma = 0.023$ [98]. The total result is:

$$\chi^{\tau}_{xx} = \frac{3ne\lambda m_\perp \mathcal{I}}{8\pi\hbar^2}(1+\gamma)^{1/2}, \tag{3.22}$$

where $\mathcal{I}$ is a dimensionless angular integral, which contains the angular part of (3.21), given by:

$$\mathcal{I} = \int_0^{\pi} d\theta \int_0^{2\pi} d\phi \sin\theta [\sin^2\theta \sin^2\phi - (1+\gamma)\cos^2\theta] \times \frac{f(\theta,\phi)}{g(\theta,\phi)}, \tag{3.23}$$

where the functions f and g are as follows:

$$\begin{aligned} f(\theta,\phi) = (1+\gamma)^2 \cos^4\theta \sin^2\phi + (1+\gamma)\cos^2\theta \sin^2\theta \sin^4\phi - \\ - \sin^4\theta \cos^4\phi \sin^2\phi - (1+\gamma)\cos^2\theta \sin^2\theta \cos^4\phi \\ g(\theta,\phi) = \sin^4\theta \cos^2\phi \sin^2\phi + (1+\gamma)\sin^2\theta \cos^2\theta(\cos^4\phi + \sin^4\phi) + \\ + (1+\gamma)^2 \cos^4\theta - 6(1+\gamma)\cos^2\theta \sin^2\theta \cos^2\phi \sin^2\phi. \end{aligned} \tag{3.24}$$

Using a Monte Carlo integration method, we find that $\mathcal{I}$=-0.03037. Based on symmetry arguments, $\mathcal{I}$ vanishes when γ=0, i.e. when $m_z = m_\perp$. This can be seen from Eq. (3.21) by switching k_y and k_z. At small γ, $0 < \gamma < 0.1$, our calculations show that $\mathcal{I}$ is linear in γ and is given approximately by $\mathcal{I} = -1.28\gamma$. We expect that, for n=$10^{19}$ cm^{-3}, $\chi^{\tau}_{xx} = -1.15 \times 10^{-7}$ c/m^2. This result is four orders of magnitude

larger than that observed in the current experiments in the weak spin-orbit coupling limit, which should offer an incentive for doping the samples in order to move into the strong spin-orbit limit. Meanwhile, for the density n=3 $\times$ 10^{16} cm^{-3} used in experiment we find $\chi^{\tau}_{xx} = -3.83 \times 10^{-10}$ c/m^2, a number that is one order of magnitude higher than that observed. We stress again however that the experiment was performed in the weak spin-orbit coupling limit.

3.4 Symmetry considerations

The form of the electric torque response tensor in different systems can be understood based on symmetry arguments. In Table I we have listed a number of possible spatial transformations which are relevant to our problem as well as the behavior of the electric field and spin torque under these transformations. The transformations considered are *not* assumed to be symmetry operations of the materials under study. I_m refers to spatial inversion along the m-axis, that is m$\rightarrow$ -m, while R_m refers to a rotation of arbitrary magnitude *anticlockwise* about the m-axis. For simplicity and without loss of generality we take the electric field to be directed along x and consider the generation of spin along all three Cartesian axes in response to this electric field.

Referring to the first part of the table, first row, it can be seen that under spatial inversion in the x-direction the electric field changes sign while $\langle\hat{\tau}\rangle_x$ remains the same. On the other hand, under the same transformation $\langle\hat{\tau}\rangle_y$ and $\langle\hat{\tau}\rangle_z$ change sign, behaving in the same way as the electric field. As a result, if spatial inversion along x is a symmetry operation of the material the torque term $\langle\hat{\tau}\rangle_x$ will vanish, but no information can be inferred about $\langle\hat{\tau}\rangle_y$ and $\langle\hat{\tau}\rangle_z$. An analogous argument holds for inversion along the y and z-axes. Therefore if spatial inversion along all three axes is a symmetry of the material all components of the torque must vanish. This is consistent with the observation that $\langle\hat{\tau}\rangle$=0 under a three dimensional inversion.

Furthermore, this table helps explain the form of the torque in systems with Rashba spin-orbit interaction. The Rashba Hamiltonian has I_x and I_y as symmetries while I_z is broken. Examining Table I we see that in response to E_x only $\langle\hat{\tau}\rangle_y$ can be non-zero, consistent with our findings in the previous section.

Next let us consider the effect of rotations by π about the three Cartesian axes. In the second part of the table, first row, we examine a rotation about x. This rotation does not affect E_x and $\langle\hat{\tau}\rangle_x$, but $\langle\hat{\tau}\rangle_y$ and $\langle\hat{\tau}\rangle_z$ change sign. Therefore, if R_x^π is a symmetry operation $\langle\hat{\tau}\rangle_y$ and $\langle\hat{\tau}\rangle_z$ will be zero in response to E_x. By extension, if rotations by π about all Cartesian axes are symmetry operations then all the off-diagonal components of the electric torque response tensor are zero. For example, in GaAs, with or without strain applied along the z-axis, rotations by π about all three axes are symmetry operations so the off diagonal components of χ^τ are zero.

Let us also consider the combined effect of rotation and inversion, taking GaAs as an example. GaAs does not have inversion along any of the three Cartesian axes as a symmetry operation. Therefore we cannot infer the vanishing or survival of the electric torque response tensor in GaAs based on considerations of inversion alone. With the electric field still along x, we consider an *anticlockwise* rotation by

Table 3.1: Behavior of the electric field and torque under spatial transformations

Operation	E_x	$\langle\hat{\tau}\rangle_x$	$\langle\hat{\tau}\rangle_y$	$\langle\hat{\tau}\rangle_z$
I_x	$-$	$+$	$-$	$-$
I_y	$+$	$-$	$+$	$-$
I_z	$+$	$-$	$-$	$+$
R_x^π	$+$	$+$	$-$	$-$
R_y^π	$-$	$-$	$+$	$-$
R_z^π	$-$	$-$	$-$	$+$
$I_x R_x^{\pi/2}$	$-$	$+$	$-\langle\hat{\tau}\rangle_z$	$+\langle\hat{\tau}\rangle_y$
$I_y R_x^{\pi/2}$	$+$	$-$	$-\langle\hat{\tau}\rangle_z$	$-\langle\hat{\tau}\rangle_y$
$I_z R_x^{\pi/2}$	$+$	$-$	$+\langle\hat{\tau}\rangle_z$	$+\langle\hat{\tau}\rangle_y$

$\frac{\pi}{2}$ about the x-axis followed by inversion along x, y and z. The rotation does not affect E_x and $\langle\hat{\tau}\rangle_x$. Looking at the third part of the table, first row, the electric field changes sign under $I_x R_x$ while $\langle\hat{\tau}\rangle_x$ does not. Therefore, if $I_x R_x^{\pi/2}$ is a symmetry of the system, the x-component of the torque in response to E_x will vanish. A similar argument holds for the y and z-axes. Therefore if $I_m R_m^{\pi/2}$ is a symmetry for all Cartesian axes all the diagonal components of the electric torque response tensor will be zero. This operation is a symmetry of bulk GaAs and in order to remove the rotational symmetry we consider strain applied along the z-axis. In this case $I_x R_x^{\pi/2}$ and $I_y R_y^{\pi/2}$ are not symmetries anymore, allowing χ^{τ}_{xx} and χ^{τ}_{yy} to be non-zero, while χ^{τ}_{zz} is zero.

We have thus shown that in crystals with inversion asymmetry and strong spin-orbit interactions a spin accumulation will be generated in the presence of an electric field due to an intrinsic spin torque term. In the steady state, the spin accumulation is given simply by the product of this torque term with the momentum relaxation time. This term is expected to be observable both in systems with Rashba spin-orbit interactions and with Dresselhaus interactions. In the latter we have shown that by doping so as to bring the system into the intrinsic regime, the effect can be observed using currently available technology.

Finally, we would like to mention a novel experimental method which has been recently proposed. In condensed matter it is usually not possible to isolate a single charge and spin carrier in order to measure effects occurring on the scale of an individual wave packet. Therefore, to be able to verify the existence and properties of the spin generation term one may resort to atomic physics. Using a cold-atom system to construct an individual wave packet and mimic the spin-orbit interaction, an experiment can measure, for example, the size, wave vector and electric field dependence of $\langle\hat{\tau}\rangle$ for this wave packet. This method and the physics underlying it are described in detail by Dudarev *et al.* [99].

Chapter 4

Coherent wave packet evolution in coupled bands

4.1 Introduction

It often happens, in transport phenomena, that one has to consider carrier dynamics in bands which are coupled together. This coupling arises either through strong interband scattering or as a result of the bands being degenerate, or both. The nearly degenerate case is particularly relevant in transport theory as transitions often occur between bands even at relatively weak electromagnetic fields. Such situations include two-dimensional systems described by the Rashba Hamiltonian [17] with strong scattering, the doubly degenerate heavy and light hole bands in the Luttinger model [100], which is frequently used to model the valence bands of bulk zincblende semiconductors, and the conduction bands of wurtzite structures [35]. The case of nearly degenerate bands has not, to date, received the attention it deserves [59, 101, 102], despite the important role played by such bands in semiconductor spintronics systems [40, 103], whether in dealing with spin currents [49, 66, 69, 77, 76, 78, 79, 80], spin generation [83, 84, 85] and relaxation [93, 94], or spin injection across a

semiconductor interface [46, 104].

Spintronics systems lend themselves to a semiclassical treatment, as the external electromagnetic fields vary on scales that are considerably larger than atomic size. The semiclassical formalism has had much success in describing carrier dynamics and transport phenomena in condensed matter physics. In the non-degenerate case, the carrier dynamics can be obtained semiclassically then combined with the Boltzmann equation to produce accurate descriptions of the transport properties of many materials. This approximation is used in the descriptions of cyclotron orbits, conduction in solids, the Hall effect and magnetoresistance [65]. An essential application of the semiclassical model, which is specifically relevant to our discussion, is in treating external fields that are not represented by bounded operators, so that a perturbative expansion will not converge [105]. The most common example is provided by uniform electric and magnetic fields, where the potential is linear in position.

We therefore develop, in this chapter, a semiclassical description of transport in degenerate and nearly degenerate bands. One of our main purposes is to extend the semiclassical approach, as developed by Sundaram and Niu [10], to the case of coupled Bloch bands, in order to take into account the spin degree of freedom. We illustrate the underlying physics by treating two bands, without loss of generality. Two-band models are frequently an adequate description of the conduction bands of many semiconductors [24, 58, 91]. In experiments on spin transport in semiconductors the carriers have traditionally been electrons [43, 44, 106], as the strong spin-orbit coupling in the valence band causes holes to lose spin information much faster [40]. However, in recent years research has also focused on spin currents in the valence bands of semiconductors [49, 66], with a degeneracy which is usually greater than two, and the formalism we outline is straightforwardly extended to multiple bands.

To formulate a description of coherent transport in coupled bands we may no longer work with each band individually but must instead treat the coupled-band manifold as a whole. The condition for our theory to be valid, which in the one-band case states that there must be no transitions out of that band [65], translates into the requirement that there be no transitions out of the manifold under consideration. We will consider a wave packet made up of two bands, which is a suitable description of *coherent* transport, when the density matrix has off-diagonal terms and the relative phase of the two wave functions plays a crucial role. This approach allows us to retain the notion of the real-space center of the wave-packet, $\mathbf{r}_c$, which remains well defined. Moreover, in extending the formalism to two bands we are able, in the presence of a magnetic field, to recover Larmor precession, which is not possible from a one band picture. The additional degree of freedom of the two-band system can be taken into account by defining a wave function with the Bloch periodicity in such a way as to incorporate both bands, which allows us to derive the dynamics from a single-band point of view. The coefficients of the bands can then be grouped into a vector, the structure and dynamics of which makes clear the gauge structure of the problem. An interesting fact which will emerge from our analysis is that the effect of the external perturbations can be incorporated entirely into the Berry curvatures, which in turn are generated by a set of connections in real and reciprocal space as well as in time. The Berry curvatures acquire additional terms needed to ensure gauge covariance, and in the framework we present they take the form of field strength tensors associated with the connections.

The organization of this chapter is as follows. In Section 4.2 we develop the semiclassical formalism for coherent transport in the presence of electromagnetic fields, deriving the Lagrangian, based on a time-dependent variational principle, and the equations of motion. In Section 4.3 we use our formalism to show how coherent wave packet evolution under the action of an electric field leads to the separation of

up and down spins. This idea is similar in principle to the spin transistor proposed by Datta and Das [18]. We demonstrate that a large degree of tunability can be achieved by varying the gate field and number density. Finally, in section 4.4 we examine the case of a magnetic field. We show that the gradient correction to the energy can be interpreted as an intrinsic magnetic moment of the wave packet [9, 10, 107], and we calculate this angular momentum correction for the light holes in the spherical four-band model of the Luttinger Hamiltonian.

4.2 Formalism

The semiclassical model describes the dynamics of wave packets. The wave packet we consider is well localized in reciprocal space, and it is assumed it sees only a small part of the lattice at any one time. It is chosen in such a way that its spread in wave vector is much smaller than the size of the Brillouin zone, so that its motion at any moment is dependent only on the local properties of the band structure. In order for this to happen, the uncertainty principle dictates that the spread in real space must be greater than the size of the lattice constant.

We consider systems whose Hamiltonians are functions of slowly varying parameters, such as the potentials of weak external electromagnetic fields, which vary on larger length scales than that of the wave packet, and are treated classically. The periodic potential of the ions on the other hand, changing over dimensions small compared to the wave packet spread, must be treated quantum mechanically [65].

Given these conditions, we define the *local* Hamiltonian $\hat{\mathrm{H}}_{\mathrm{c}}(\mathbf{r}_{\mathrm{c}}, \mathrm{t})$ as the Hamiltonian with the slowly varying potentials evaluated at the center of the wave packet, which we denote by $\mathbf{r}_{\mathrm{c}}$, and time t. The Hamiltonian may be expanded [10] about $\mathbf{r}_{\mathrm{c}}$ and if the external fields vary on spatial scales much larger than that of the wave packet we may truncate the expansion at the gradient term, which we define

by $\Delta\hat{\mathrm{H}}$:

$$\Delta\hat{H} = \frac{1}{2}[(\hat{\mathbf{r}} - \mathbf{r}_c) \cdot \frac{\partial \hat{H}_c}{\partial \mathbf{r}_c} + c.c.]. \tag{4.1}$$

The gradient term gives rise to a correction to the energy, which will play an important role in our discussion below.

The energy spectrum of the local Hamiltonian $\hat{\mathrm{H}}_{\mathrm{c}}$ consists, as usual, of a series of bands, of which several are close together in energy and are separated from the others by larger gaps. It is the subset spanned by these bands that constitutes the focus of our attention. We regard the fields in this problem as small enough that Zener tunneling to the remote bands is negligible, but they may still be strong enough to induce transitions within the subset. For an energy spectrum with such a structure we may further decompose the local Hamiltonian into a degenerate part, $\hat{\mathrm{H}}_{\mathrm{d}}$, which, when restricted to the subset of bands closely spaced in energy is proportional to the identity matrix, and a non-degenerate part, $\hat{\mathrm{H}}_{\mathrm{n}}$, which is assumed small and treated perturbatively. The local Hamiltonian of (4.1) is then:

$$\hat{H}_c = \hat{H}_d + \hat{H}_n. \tag{4.2}$$

The gradient correction to $\hat{\mathrm{H}}_{\mathrm{c}}$ can also be expressed in terms of the degenerate and non-degenerate contributions:

$$\Delta\hat{H}_c = \Delta\hat{H}_d + \Delta\hat{H}_n. \tag{4.3}$$

Since $\hat{\mathrm{H}}_{\mathrm{n}}$ is treated as a perturbation the gradient correction to it, $\Delta\hat{\mathrm{H}}_{\mathrm{n}}$ will be second order in smallness. We will therefore neglect this correction henceforth.

When the external fields are smoothly varying the states move within the subset of bands which are close in energy and which henceforth, for simplicity and without loss of generality, we take to be two-dimensional. The subset is spanned by two basis functions, which are eigenstates of $\hat{\mathrm{H}}_{\mathrm{d}}$, the degenerate part of the local

Hamiltonian, evaluated at $\mathbf{r}_c$, which has the periodicity of the unperturbed crystal:

$$\hat{H}_d|\Psi_i(\mathbf{r}_c,\mathbf{q},t)\rangle = \epsilon|\Psi_i(\mathbf{r}_c,\mathbf{q},t)\rangle. \tag{4.4}$$

For a given $\mathbf{r}_c$, therefore, these eigenstates have the Bloch form, with the functions $|u_i\rangle$ representing the lattice periodic parts of the wave functions:

$$|\Psi_1(\mathbf{r}_c,\mathbf{q},t)\rangle = e^{i\mathbf{q}\cdot\hat{\mathbf{r}}}|u_1(\mathbf{r}_c,\mathbf{q},t)\rangle \tag{4.5}$$

$$|\Psi_2(\mathbf{r}_c,\mathbf{q},t)\rangle = e^{i\mathbf{q}\cdot\hat{\mathbf{r}}}|u_2(\mathbf{r}_c,\mathbf{q},t)\rangle \tag{4.6}$$

The wave functions $|u_i(\mathbf{r}_c,\mathbf{q},t)\rangle$ are spinors with the full periodicity of the lattice. Despite the fact that the two bands are spin split, it cannot be assumed that their local spin quantization axes are antiparallel, as the interactions with neighboring bands may affect the direction of quantization. Therefore, in principle, a finite overlap exists between the spinors corresponding to the two bands and it is not revealing to make a further decomposition of the eigenfunctions into an orbital and a spin part. Additionally, the Hamiltonian contains terms describing the spin-orbit interaction, which may depend on wave vector and position.

Employing the crystal momentum representation, the wave packet is expanded in the basis of Bloch eigenstates:

$$|w\rangle = \int d^3q\{a(\mathbf{q},t)[\eta_1(\mathbf{q},t)|\Psi_1\rangle + \eta_2(\mathbf{q},t)|\Psi_2\rangle]\}. \tag{4.7}$$

As the wave packet depends only on the local properties of the band structure, the basis functions $|\Psi_1\rangle$, $|\Psi_2\rangle$ are functions of the position of the wave packet center, $\mathbf{r}_c$, wave vector and time, although implicit in the ket notation is dependence on position. The function $a(\mathbf{q},t) = |a(\mathbf{q},t)|e^{-i\Gamma(\mathbf{q},t)/2}$, which incorporates the overall phase term, is a narrow distribution function describing the extent of the wave

packet in reciprocal space and is sharply peaked at the center of the wave packet, denoted by $\mathbf{q}_c$, as discussed by Sundaram and Niu [10]. The functions η_1 and η_2 describe the composition of the wave packet in terms of the two bands. The wave packet satisfies the normalization conditions:

$$\int d^3q|a|^2 = 1, |\eta_1|^2 + |\eta_2^2| = 1. \tag{4.8}$$

The wave packet can be rewritten by grouping together the coefficients in an overall wave function $|\mathrm{u}\rangle$, which retains the Bloch periodicity:

$$|w\rangle = \int d^3q|a|e^{-i\Gamma/2}e^{i\mathbf{q}\cdot\hat{\mathbf{r}}}|u\rangle. \tag{4.9}$$

Note that $|\mathrm{u}\rangle$ is not an eigenstate of the local Hamiltonian $\hat{\mathrm{H}}_\mathrm{c} = \hat{\mathrm{H}}_\mathrm{d} + \hat{\mathrm{H}}_\mathrm{n}$, but an expansion in eigenstates of $\hat{\mathrm{H}}_\mathrm{d}$, a crucial difference from the one-band situation. In addition, the wave vector and time dependence of $|\mathrm{u}\rangle$ come both from the time dependence of the Bloch states and that of the coefficients.

We require the real space center of the wave packet to be given by:

$$\mathbf{r}_c = \langle w|\hat{\mathbf{r}}|w\rangle = \frac{\partial\Gamma_c}{\partial\mathbf{q}_c} + \mathbf{R}_c \tag{4.10}$$

The subscript c signifies that the quantity is evaluated at the center of the wave packet in reciprocal space, that is $\mathbf{q} = \mathbf{q}_\mathrm{c}$. The vector $\mathbf{R}$, representing a connection in reciprocal space, is defined as follows:

$$\mathbf{R} = \langle u|i\frac{\partial}{\partial\mathbf{q}}|u\rangle. \tag{4.11}$$

The energy of the wave packet is given by the expectation value

$$\begin{aligned}\langle w|\hat{H}|w\rangle = \langle w|\hat{H}_d|w\rangle + \langle w|\hat{H}_n|w\rangle + \langle w|\Delta\hat{H}_d|w\rangle \\ \equiv \epsilon + \Delta_n + \Delta_d \equiv \varepsilon.\end{aligned} \tag{4.12}$$

Both Δ_{n} and Δ_{d} are expressible entirely in terms of the Bloch wave function $|\mathrm{u}\rangle$. Δ_{n} is given by

$$\Delta_n = \langle u|\tilde{H}_n|u\rangle = \eta_i^*\Delta_{ij}^n\eta_j, \Delta_{ij}^n\eta_j = \langle u_i|\tilde{H}_n|u_j\rangle, \tag{4.13}$$

while Δ_{d} is

$$\Delta_d = \frac{i}{2}(\langle u|\frac{\partial\tilde{H}_d}{\partial\mathbf{r}_c}\cdot|\frac{\partial u}{\partial\mathbf{q}}\rangle - c.c.) - \frac{\partial\hat{\epsilon}}{\partial\mathbf{r}_c}\cdot\mathbf{R}. \tag{4.14}$$

In the above $\tilde{\mathrm{H}}_{\mathrm{n}} = \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}\hat{\mathrm{H}}_{\mathrm{n}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}$ while $\tilde{\mathrm{H}}_{\mathrm{d}} = \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}\hat{\mathrm{H}}_{\mathrm{d}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}$. The energy correction Δ_{d} is identical to the expression obtained by Sundaram and Niu [10]. It takes on an additional significance when a magnetic field is present, as will be seen in the last section.

The Lagrangian $\mathcal{L}$ is obtained semiclassically by means of a variational principle:

$$\mathcal{L} = \langle w|(i\hbar\frac{d}{dt} - \hat{H})|w\rangle \tag{4.15}$$

Its use is justified by the fact that the Euler-Lagrange equation of motion for $|\mathrm{w}\rangle$ derived from it is the time-dependent Schrodinger equation. Following the method used by Sundaram and Niu, the following expression is found for the Lagrangian:

$$\begin{aligned}\mathcal{L} = \langle u|i\hbar\frac{du}{dt}\rangle + \hbar\dot{\mathbf{r}}_c\cdot\mathbf{q}_c - \varepsilon = \\ = i\hbar\eta_i^*\frac{d\eta_i}{dt} + \hbar\eta_i^*\langle u_i|i\frac{du_j}{dt}\rangle\eta_j + \hbar\dot{\mathbf{r}}_c\cdot\mathbf{q}_c - \epsilon - \eta_i^*(\Delta_{ij}^n + \Delta_{ij}^d)\eta_j.\end{aligned} \tag{4.16}$$

In the above $\frac{\mathrm{d}}{\mathrm{dt}}$ represents the total time derivative, including both the explicit time dependence and the implicit, which is due to dependence on $\mathbf{q}_\mathrm{c}$ and $\mathbf{r}_\mathrm{c}$. The Lagrangian depends only on the values of η_i and $\frac{\mathrm{d}\eta_\mathrm{i}}{\mathrm{dt}}$ along the trajectory $\mathbf{q} = \mathbf{q}_\mathrm{c}(\mathrm{t})$. Since $\mathbf{q}_\mathrm{c}$ is a function of time only, we may regard η_i in the Lagrangian as an independent variable, $\eta_\mathrm{i}(\mathrm{t})$. The equations of motion derived from the Lagrangian are:

$$\begin{aligned} \hbar\dot{\mathbf{q}}_c &= -\frac{\partial\epsilon}{\partial\mathbf{r}_c} + (\mathbf{\Omega_{rr}}\dot{\mathbf{r}}_c + \mathbf{\Omega_{rq}}\dot{\mathbf{q}}_c) - \mathbf{\Omega}_{t\mathbf{r}} \\ \hbar\dot{\mathbf{r}}_c &= \frac{\partial\epsilon}{\partial\mathbf{q}_c} - (\mathbf{\Omega_{qr}}\dot{\mathbf{r}}_c + \mathbf{\Omega_{qq}}\dot{\mathbf{q}}_c) + \mathbf{\Omega}_{t\mathbf{q}} \\ i\hbar\frac{d\eta_i}{dt} &= (\mathcal{H}_{ij} - \hbar\langle u_i|i\frac{du_j}{dt}\rangle)\eta_j. \end{aligned} \tag{4.17}$$

Here $\mathcal{H}_{\mathrm{ij}} = \langle \mathrm{u_i}|\hat{\mathrm{H}}_\mathrm{c}|\mathrm{u_j}\rangle$. The curvature tensor $\Omega^{\alpha\beta}_{\mathbf{rr}}$ is defined by:

$$\Omega^{\alpha\beta}_{\mathbf{rr}} = i(\langle\frac{\partial u}{\partial r_\alpha}|\frac{\partial u}{\partial r_\beta}\rangle - \langle\frac{\partial u}{\partial r_\beta}|\frac{\partial u}{\partial r_\alpha}\rangle) \tag{4.18}$$

and the vector $\mathbf{\Omega}_{\mathbf{tq}}$ by:

$$\Omega^{\alpha}_{t\mathbf{q}} = i(\langle\frac{\partial u}{\partial t}|\frac{\partial u}{\partial q_\alpha}\rangle - \langle\frac{\partial u}{\partial q_\alpha}|\frac{\partial u}{\partial t}\rangle) \tag{4.19}$$

The others can be deduced analogously. These quantities have exactly the same form as the curvatures defined in the paper by Sundaram and Niu [10].

We specialize in the case of an external electromagnetic field. The effect of such an external field is discussed thoroughly by Sundaram and Niu [10]. The wave vector $\mathbf{q}$ must be replaced by $\mathbf{k} = \mathbf{q} + \frac{\mathrm{e}}{\hbar}\mathbf{A}(\mathbf{r},\mathrm{t})$, which is the gauge invariant crystal momentum (for electrons with charge -e), and therefore the Hamiltonian will have the form $\hat{\mathrm{H}}(\mathbf{k}) + \mathrm{eV}(\mathbf{r},\mathrm{t})$. Provided the magnetic or exchange field is constant and uniform, so that the Zeeman term has no time or space dependence, the basis states $\{|\mathrm{u_i}\rangle\}$ will depend only on $\mathbf{k}$. The reason for this is that all the

spatial and time dependence of the wave functions will only come from the spatial and time dependence of the vector potential $\mathbf{A}(\mathbf{r},\mathrm{t})$. We will therefore restrict our attention to constant uniform magnetic fields, while the electric fields may be space- and time-dependent. As the electromagnetic fields vary on a spatial scale which is large compared to that of the wave packet, the local Hamiltonian will have the form $\hat{\mathrm{H}}[\mathbf{q}+\frac{\mathrm{e}}{\hbar}\mathbf{A}(\mathbf{r}_\mathrm{c},\mathrm{t})]+\mathrm{eV}(\mathbf{r}_\mathrm{c},\mathrm{t})$. The band eigenstates $\{|\Psi_{\mathrm{n}\mathbf{k}}\rangle\}$ take the form $|\psi_{\mathrm{n}\mathbf{k}}\rangle=\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}|\mathrm{u}_{\mathrm{n}\mathbf{k}}\rangle=\mathrm{e}^{\mathrm{i}(\mathbf{k}-\frac{\mathrm{e}\mathbf{A}}{\hbar})\cdot\mathbf{r}}|\mathrm{u}_{\mathrm{n}\mathbf{k}}\rangle$. The time dependence of $|\mathrm{u}\rangle$ comes both from the Bloch wave functions $\{|\mathrm{u_i}\rangle\}$, which depend only on $\mathbf{k}$, and from the coefficients, which depend only on time. Therefore, the Lagrangian in the presence of electromagnetic fields can be written as:

$$\mathcal{L}=\hbar\langle u|i\frac{d}{dt}|u\rangle+[\hbar\mathbf{k}_c-e\mathbf{A}(\mathbf{r}_c,t)]\cdot\dot{\mathbf{r}}_c-\epsilon-\Delta_n-\Delta_d-eV(\mathbf{r}_c,t). \tag{4.20}$$

The equations of motion now take the following form:

$$\begin{aligned}\hbar\dot{\mathbf{k}}_c&=-e(\mathbf{E}+\dot{\mathbf{r}}_c\times\mathbf{B})\\ \hbar\dot{\mathbf{r}}_c&=\frac{\partial}{\partial\mathbf{k}_c}\langle u|\hat{H}|u\rangle-\hbar\dot{\mathbf{k}}_c\times\boldsymbol{\Omega}+\boldsymbol{\Omega}_{t\mathbf{k}}\\ i\hbar\frac{d\eta_i}{dt}&=(\mathcal{H}_{ij}-\hbar\dot{\mathbf{k}}_c\langle u_i|i\frac{\partial u_j}{\partial\mathbf{k}_c}\rangle)\eta_j,\end{aligned}$$

where $\boldsymbol{\Omega}=\mathrm{i}\langle\frac{\partial\mathrm{u}}{\partial\mathbf{k}}|\times|\frac{\partial\mathrm{u}}{\partial\mathbf{k}}\rangle$. Note that the position vector equation of motion is very similar to the one band case [10], excepting the presence of the vector $\boldsymbol{\Omega}_{\mathrm{t}\mathbf{k}}$, which is non-zero due to the time dependence of $|\mathrm{u}\rangle$ through the coefficients. The equation of motion for $|\mathrm{u}\rangle$, if a magnetic field is present, leads to the formula for Larmor precession. The equations may be solved to any desired order in the external fields and are not limited to the linear response regime (the fields are weak enough that they do not induce transitions to remote bands).

4.3 The probability amplitudes

The treatment we have presented so far is an exact analogy with the single-band dynamics. The equations of motion (4.21) are complete. Nevertheless, the equations of motion can be made more explicit in terms of the coefficients η_i, and the non-Abelian quantities emerging in the process illustrate the gauge structure of the Hilbert space.

The coefficients η_1, η_2 give the composition of the wave packet in terms of the two bands, and it is natural to think of them as a vector, $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}$, which will be called η. The connection $\mathbf{R}$ can be expanded in terms of η:

$$R^\alpha = \eta^\dagger \mathcal{R}^\alpha \eta + i\eta^\dagger \frac{\partial \eta}{\partial q_\alpha} \ , \text{where } \mathcal{R}^\alpha_{ij} = \langle u_i | i \frac{\partial u_j}{\partial q_\alpha} \rangle. \tag{4.21}$$

We will also introduce the time connection $\mathcal{T}_{ij} = \langle u_i | i \frac{\partial u_j}{\partial t} \rangle$. The Lagrangian in this picture takes the form:

$$\mathcal{L} = i\hbar \eta^\dagger \frac{D\eta}{Dt} + \hbar \mathbf{q}_c \cdot \dot{\mathbf{r}}_c - \eta^\dagger \mathcal{H} \eta. \tag{4.22}$$

where the covariant derivative with respect to time, defined as $\frac{D}{Dt} = \frac{d}{dt} - i(\mathcal{T} + \dot{\mathbf{q}}_c \cdot \mathbf{R})$, has been introduced. Specializing in electromagnetic fields, we end up with the following Lagrangian:

$$\mathcal{L} = \eta^\dagger (i\hbar \frac{D}{Dt})\eta + [\hbar \mathbf{k}_c - e\mathbf{A}(\mathbf{r}_c, t)] \cdot \dot{\mathbf{r}}_c - \eta^\dagger \mathcal{H} \eta - eV(\mathbf{r}_c, t). \tag{4.23}$$

The equations of motion derived from this electromagnetic Lagrangian are as follows:

$$\hbar\dot{\mathbf{k}}_c = -e(\mathbf{E} + \dot{\mathbf{r}}_c \times \mathbf{B}) \tag{4.24}$$
$$\hbar\dot{\mathbf{r}}_c = \eta^\dagger[\frac{D}{D\mathbf{k}}, \mathcal{H}]\eta - \hbar\dot{\mathbf{k}}_c \times \eta^\dagger\mathcal{F}\eta$$
$$i\hbar\frac{D\eta}{Dt} = \mathcal{H}\eta$$

The covariant derivative with respect to the wave vector has been introduced, which has the form $\frac{\mathrm{D}}{\mathrm{Dk}_\alpha} = \frac{\partial}{\partial \mathrm{k}_\alpha} - \mathrm{i}\mathcal{R}^\alpha$. The non-Abelian Berry curvature matrix, $\mathcal{F}^\gamma_{\mathrm{ij}}$, is expressed in terms of the field strength tensor corresponding to the covariant wave vector derivatives:

$$\mathcal{F}^\gamma_{ij} = \frac{1}{2}\epsilon^{\alpha\beta\gamma}\mathcal{F}^{\alpha\beta}_{ij} \tag{4.25}$$

where

$$\mathcal{F}^{\alpha\beta}_{ij} = i[\frac{D}{Dk_\alpha}, \frac{D}{Dk_\beta}]_{ij} = \tag{4.26}$$
$$= \frac{\partial\mathcal{R}^\beta_{ij}}{\partial k_\alpha} - \frac{\partial\mathcal{R}^\alpha_{ij}}{\partial k_\beta} - i[\mathcal{R}^\alpha, \mathcal{R}^\beta]_{ij}$$

This form, which includes the non-Abelian correction from the commutator of the connection matrices, makes evident its gauge covariance with respect to unitary transformations of η. The curvature tensor is antisymmetric under interchange of α and β, while the indices i and j satisfy $\mathcal{F}^{\alpha\beta}_{\mathrm{ij}} = (\mathcal{F}^{\alpha\beta}_{\mathrm{ji}})^*$.

It is seen from the equations of motion that working in the coupled-band manifold entails the presence of non-Abelian quantities such as the modified Berry curvature and gauge covariant group velocity $\frac{1}{\hbar}[\frac{\mathrm{D}}{\mathrm{D}\mathbf{k}}, \mathcal{H}]$, which are corrections to the one band equations of motion needed to ensure gauge covariance. The matrix $\mathcal{H}$ is not necessarily diagonal, as it may include energy gradient corrections.

We note that equivalent results can be derived using an argument based on the Ehrenfest theorem, as in the extensive work of Shindou and Imura [108].

4.4 Constant electric field

We will examine first the case of a constant uniform electric field acting on two degenerate bands. We choose a gauge such that the scalar electric potential need not be included in the Hamiltonian, and the electric field is represented purely by the vector potential **A**. With experiment in mind, we take **E**=(0,0,E), modeling a gate field, and study its effect on transport in the xy-plane.

4.4.1 Electrical spin separation

We choose as an example the spherical four-band model:

$$\hat{H}_{Lutt} = \frac{\hbar^2}{2m}[(\gamma_1 + \frac{5}{2}\gamma_2)k^2 - 2\gamma_2(\mathbf{k}\cdot\hat{\mathbf{J}})^2], \tag{4.27}$$

where $\hat{\mathbf{J}}$ is the total angular momentum operator, m is the bare electron mass and γ_1 and γ_2 are material-specific parameters. The wave functions are eigenstates of the helicity operator $\mathbf{k}\cdot\hat{\mathbf{J}}$ and have the form $|\mathrm{u_m}\rangle = \mathrm{e}^{-\mathrm{i}\phi\mathrm{J_z}}\mathrm{e}^{-\mathrm{i}\theta\mathrm{J_y}}|\mathrm{m}\rangle$ where $|\mathrm{m}\rangle$ are eigenstates of the angular momentum operator $\mathrm{J_z}$ while θ and ϕ are the polar and azimuthal angles of the wave vector, respectively. We shall treat the two-fold degenerate heavy and light hole manifolds separately and we shall denote the probability amplitudes in the heavy hole subspace by η^{H} and those in the light hole subspace by η^{L}.

In these subspaces, the equations of motion for the probability amplitudes take the form

$$\begin{aligned} i\hbar\frac{d\eta^H}{dt} &= (\mathcal{H}^H - eE\mathcal{R}^{zH})\eta^H, \\ i\hbar\frac{d\eta^L}{dt} &= (\mathcal{H}^L - eE\mathcal{R}^{zL})\eta^L, \end{aligned} \tag{4.28}$$

where the superscripts H and L represent restrictions to the heavy and light hole subspaces, respectively. The reciprocal space connection matrix is given by the

following expression:

$$\mathcal{R} = \frac{\partial\theta}{\partial\mathbf{k}} J^y + \frac{\partial\phi}{\partial\mathbf{k}}(J^z\cos\theta - J^x\sin\theta). \tag{4.29}$$

In the heavy hole sector $\mathcal{R}^z$=0 and the bands decouple, therefore no spin separation can be achieved electrically in the heavy hole manifold. Henceforth we shall concentrate on the light hole manifold, where the connection matrix $\mathcal{R}^z = -\frac{k_\perp}{k^2}\sigma^y$ has off-diagonal elements only, with $\mathbf{k}_\perp = (k_x, k_y)$ and σ^y a Pauli spin matrix. We shall suppress the index L in what follows.

The equations of motion for the position and wave vector are:

$$\hbar\dot{\mathbf{k}} = e\mathbf{E}$$
$$\hbar\dot{\mathbf{r}}_c = \frac{\partial\varepsilon_l}{\partial\mathbf{k}} - e\mathbf{E}\times\eta^\dagger\mathcal{F}\eta,$$

in which $\mathbf{k}_0$ is the initial value of $\mathbf{k}$, $\varepsilon_l = \frac{\hbar^2 k^2}{2m_l}$ is the light hole energy, m_l is the light hole effective mass, and the curvature $\mathcal{F} = \frac{3}{2}\frac{\mathbf{k}}{k^3}\sigma^z$. The wave vector equation of motion is readily integrated to give $\mathbf{k} = \mathbf{k}_0 + \frac{eEt}{\hbar}$. Since the Berry curvature vector is parallel to $\mathbf{k}$, there are two limiting cases to consider: the case $\mathbf{k}_0//\mathbf{E}$ is trivial because the curvature correction vanishes and the bands decouple, so we will focus on the more interesting case $\mathbf{k}_0 \perp \mathbf{E}$.

The equations of motion can be solved exactly. η is given by:

$$\eta = \begin{pmatrix} \eta_1^{(0)}\cos\alpha + \eta_2^{(0)}\sin\alpha \\ \eta_2^{(0)}\cos\alpha - \eta_1^{(0)}\sin\alpha \end{pmatrix}, \tag{4.30}$$

with the angle $\alpha(\tau) = \arctan(\frac{\tau+\cos\theta_0}{\sin\theta_0}) - (\frac{\pi}{2} - \theta_0)$, where we have introduced the dimensionless time $\tau = \frac{eEt}{\hbar k_0}$ and θ_0 is the polar angle of $\mathbf{k}_0$, and where $\eta_i^{(0)}$ are the values of η at τ=0.

In this system, the contraction $\eta^\dagger\hat{\sigma}^i\eta$ (with i=1,2,3) is the expectation value

of the pseudo-spin. Its components evolve in time as:

$$\begin{aligned}\langle\hat{\sigma}^1\rangle &= \langle\hat{\sigma}^1\rangle_{\tau=0}\cos 2\alpha - \langle\hat{\sigma}^3\rangle_{\tau=0}\sin 2\alpha \\ \langle\hat{\sigma}^2\rangle &= \langle\hat{\sigma}^2\rangle_{\tau=0} \\ \langle\hat{\sigma}^3\rangle &= \langle\hat{\sigma}^3\rangle_{\tau=0}\cos 2\alpha + \langle\hat{\sigma}^1\rangle_{\tau=0}\sin 2\alpha.\end{aligned}$$

The electric field therefore only rotates the 1 and 3 components of the pseudo-spin into combinations of each other, while the 2 component remains unaffected. To understand the significance of these results we will examine a concrete example, taking initially a positive helicity eigenstate so that $\eta_1^{(0)} = 1, \eta_2^{(0)} = 0$, and fixing the initial wave vector along the x-axis such that $\mathbf{k}_0 = \mathrm{k}_0\hat{\mathbf{x}}$, which means that $\theta_0 = \frac{\pi}{2}$. The full time evolution of the pseudo-spin components is:

$$\begin{aligned}\langle\hat{\sigma}^1\rangle &= \langle\hat{\sigma}^1\rangle_{\tau=0}\frac{1-\tau^2}{1+\tau^2} - \langle\hat{\sigma}^3\rangle_{\tau=0}\frac{2\tau}{1+\tau^2} \\ \langle\hat{\sigma}^2\rangle &= \langle\hat{\sigma}^2\rangle_{\tau=0} \\ \langle\hat{\sigma}^3\rangle &= \langle\hat{\sigma}^3\rangle_{\tau=0}\frac{1-\tau^2}{1+\tau^2} + \langle\hat{\sigma}^1\rangle_{\tau=0}\frac{2\tau}{1+\tau^2}.\end{aligned} \tag{4.31}$$

As $\tau \to \infty$, α reaches the limiting value of $\frac{\pi}{2}$ and the components of the pseudo-spin become:

$$\begin{aligned}\langle\hat{\sigma}^1\rangle &= -\langle\hat{\sigma}^1\rangle_{\tau=0} \\ \langle\hat{\sigma}^2\rangle &= \langle\hat{\sigma}^2\rangle_{\tau=0} \\ \langle\hat{\sigma}^3\rangle &= -\langle\hat{\sigma}^3\rangle_{\tau=0}.\end{aligned} \tag{4.32}$$

Thus the 1 and 3 components of the pseudo-spin are reversed while the 2 component is conserved.

The time evolution of the wave vector is described entirely in terms of the

time evolution of the angle θ, which is most conveniently expressed as

$$\begin{aligned} \cos\theta &= \frac{k_z}{k} = \frac{\frac{eEt}{\hbar}}{\sqrt{k_0^2 + (\frac{eEt}{\hbar})^2}} = \frac{\tau}{\sqrt{1+\tau^2}}, \\ \sin\theta &= \frac{k_\perp}{k} = \frac{k_0}{\sqrt{k_0^2 + (\frac{eEt}{\hbar})^2}} = \frac{1}{\sqrt{1+\tau^2}}. \end{aligned} \tag{4.33}$$

Therefore, initially we have $\cos\theta$=0 and $\sin\theta$=1 while as $\tau \to \infty$, $\cos\theta \to$1 and $\sin\theta \to$0.

The expectation value of a spin component operator $\hat{s}^\alpha$ in the wave packet $|w\rangle$ is given by $\langle w|\hat{s}^\alpha|w\rangle = \eta^\dagger s^\alpha \eta$, where $s^\alpha_{ij} = \langle u_i|\hat{s}^\alpha|u_j\rangle$. The time evolution of the spin of one electron can thus be found by knowing the time evolution of its pseudo-spin. Since our goal is to separate spins of opposite orientations, it is sufficient to know the value of the pseudo-spin. The bands being spin-split holes with pseudo-spin up also have spin up and holes with pseudo-spin down have spin down. However, it is instructive to follow the motion of the spin as time progresses, as well as the time evolution of the helicity. The expectation values of $\hat{s}^x$, $\hat{s}^y$ and $\hat{s}^z$ are:

$$\begin{aligned} \langle \hat{s}^x \rangle &= \frac{\hbar}{3}(\frac{1}{2}\sin\theta\cos\phi\langle\hat{\sigma}^3\rangle + \cos\theta\cos\phi\langle\hat{\sigma}^1\rangle - \sin\phi\hat{\sigma^2}) \\ \langle \hat{s}^y \rangle &= \frac{\hbar}{3}(\frac{1}{2}\sin\theta\sin\phi\langle\hat{\sigma}^3\rangle + \cos\theta\sin\phi\langle\hat{\sigma}^1\rangle + \cos\phi\hat{\sigma^2}) \\ \langle \hat{s}^z \rangle &= \frac{\hbar}{3}(\frac{1}{2}\cos\theta\langle\hat{\sigma}^3\rangle - \sin\theta\langle\hat{\sigma}^1\rangle). \end{aligned} \tag{4.34}$$

We assume the carriers have been polarized (by optical means, for example as done in the experiments of Malajovich *et al.* [46, 109], although those utilized electrons) so that $\eta^{(0)}$ is either $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$. Therefore the initial expectation values of $\{\hat{\sigma}^i\}$

are:

$$\begin{aligned} \langle\hat{\sigma}^1\rangle_\uparrow = \langle\hat{\sigma}^1\rangle_\downarrow = 0 \\ \langle\hat{\sigma}^2\rangle_\uparrow = \langle\hat{\sigma}^2\rangle_\downarrow = 0 \\ \langle\hat{\sigma}^3\rangle_\uparrow = -\langle\hat{\sigma}^3\rangle_\downarrow = 1, \end{aligned} \tag{4.35}$$

and the initial spin expectation values are given by

$$\langle\hat{s}^x\rangle_0 = \pm\frac{\hbar}{6}, \langle\hat{s}^y\rangle_0 = \langle\hat{s}^z\rangle_0 = 0. \tag{4.36}$$

It can then easily be seen that the y-component of the spin is zero at all times. Substituting for θ and $\{\hat{\sigma}^{\mathrm{i}}\}$ in (4.34) we obtain the time evolution of the other two spin components:

$$\langle\hat{s}^x\rangle_0 = \pm\frac{\hbar}{3}\frac{1-5\tau^2}{2(1+\tau^2)^{3/2}}, \langle\hat{s}^z\rangle_0 = \pm\frac{\hbar}{3}\frac{\tau(5-\tau^2)}{2(1+\tau^2)^{3/2}}. \tag{4.37}$$

As $\tau \to \infty$, the expectation values of the spin components are:

$$\begin{aligned} \langle\hat{s}^x\rangle = \langle\hat{s}^y\rangle = 0 \\ \langle\hat{s}^z\rangle = \mp\frac{\hbar}{6}. \end{aligned} \tag{4.38}$$

The spin in this case is not conserved. However, a closer look at (4.34) reveals that $\langle\hat{\mathrm{s}}^{\mathrm{x}}\rangle$, $\langle\hat{\mathrm{s}}^{\mathrm{y}}\rangle$ and $\langle\hat{\mathrm{s}}^{\mathrm{z}}\rangle$ cannot be obtained from $\langle\hat{\sigma}^1\rangle$, $\langle\hat{\sigma}^2\rangle$ and $\langle\hat{\sigma}^3\rangle$ by a rotation, as the matrix describing the transformation is not unitary. Therefore one should not think of the projection of the spin onto the light-hole subspace as a vector.

Finally, the helicity is given by:

$$\chi = \frac{\mathbf{k}\cdot\langle\hat{\mathbf{s}}\rangle}{k} = \frac{k_x\langle\hat{s}^x\rangle + k_z\langle\hat{s}^z\rangle}{k} = \frac{\hbar}{6}\langle\hat{\sigma}^3\rangle = \pm\frac{\hbar}{6}\frac{1-\tau^2}{1+\tau^2}. \tag{4.39}$$

The helicity is proportional to the expectation value of the third component of the pseudo-spin. It is therefore not conserved for light holes in an electric field. This conclusion was also reached by Jiang *et al.* [110].

The $\mathbf{r}_c$ equation of motion can be integrated to give the trajectories of the carriers:

$$\mathbf{r}_c = \frac{\hbar^2 k_0^2}{eEm_l}(\tau\hat{\mathbf{x}} + \frac{\tau^2}{2}\hat{\mathbf{z}}) - \frac{\tau(3+\tau^2)(|\eta_1^{(0)}|^2 - |\eta_2^{(0)}|^2)}{2k_0(1+\tau^2)^{3/2}}\hat{\mathbf{y}}. \tag{4.40}$$

We have omitted a term proportional to $\eta_1^{(0)}\eta_2^{(0)}$ since in our setup either one of them will be zero . The second term in (4.40) will have opposite signs for the carriers with η initially up and those with η initially down. Therefore, these carriers will be separated in the y-direction. From the above and Fig. 4.1 it can be seen that the maximum separation in the y-direction occurs at $\tau = 1$ while as $\tau \to \infty$ this separation tends to $1/\mathrm{k}_0$.

4.4.2 Experimental observation

We discuss an experimental setup in which the effect we have described can be measured. We propose using a three-dimensional semiconductor slab containing a non-degenerate hole gas. The sample must be clean in order for the hole spin relaxation time to be long, specifically of the order of picoseconds. Carriers are excited optically from the conduction band into the valence bands by using a laser beam. Provided the laser beam is sharp, only a narrow range of k-space will be excited around $\mathbf{k}$=0. The optically excited holes will have wave vectors lying in a narrow spot about the origin. We assume they have been excited into a state of definite spin. Both light and heavy holes are excited but, as shown in the previous section, the heavy holes do not separate according to spin under the action of the electric field. A source and a drain will be positioned along the x-direction on the two faces of the sample while a gate terminal will be present on top. After the optical excitation, the magnitude of the holes' wave vector can be increased by applying

a source-drain field E_x in the form of a picosecond pulse, which will accelerate the carriers along the x-axis, its magnitude tuned to ensure k_0 has the desired value. This source-drain field provides an additional advantage. In the process of optical excitation electrons as well as holes will be excited in the sample and the field which drives the holes one way will drive the electrons the other way, ensuring the effect observed is indeed due to holes. By adjusting the magnitude of the source-drain electric field pulse the initial wave vector k_0 of the holes incident upon the interface is tunable over several orders of magnitude. We will choose a source-drain electric field in such a way that the wave vector $\mathbf{k}_0$ will have an x-component which overwhelms the y- and z-components. We will also choose the magnitude of k_0 to be approximately 1/b, where b is the real space thickness of the laser beam. The reason for this is that in the limit of large τ the spins are separated by a distance of approximately $1/k_0$, therefore the separation of the spins will be approximately the same as the width of the laser beam. Once excited, the carriers will be subjected to the action of the gate electric field $\mathbf{E}$ along $\hat{\mathbf{z}}$, which will lead to separation of spins as described above. The spin accumulation at the other end of the sample can be measured by Faraday or Kerr rotation. It will be position-dependent along y, that is, as one moves along $\hat{\mathbf{y}}$ the spin-z polarization will change sign.

We take the dimensions of the slab to to be 50 nm $\times$ 5 μm $\times$ 5 μm and the width of the laser beam is taken as 1 μm. The optically excited holes will be accelerated until their wave vector reaches the value $k_0 = 10^6$ m^{-1}. For a source-drain field E_x of 500 Vm^{-1} and a light hole mass of $0.1m_0$, where m_0 is the bare electron mass, the distance traveled by the light holes along the x-axis will be $\frac{\hbar^2 k_0^2}{2m_l eE_x}$ = 7.2 nm. This will happen after a time of 1.25 ps, which can be achieved in samples in which the holes have longer spin lifetimes. Therefore the source-drain field must be a 1.25 ps pulse of amplitude 500 V m^{-1}.

We will take the gate electric field E=25000 V m^{-1}. If one waits for the value

of τ to reach 50, then the magnitude of the spin polarization along the z-direction will be approximately $\hbar/6$ while along the x-direction it will be negligible. The separation between the carriers with spin-z up and spin-z down will be approximately 1 μm, which is the same as the real-space width of the laser beam and thus observable. The waiting time will be approximately 1.3 ps. Finally, the distances traveled in the x- and z-directions under the action of the gate electric field are 35 nm and 850 nm respectively.

This phenomenon is similar to effects such as the spin Hall effect since carriers with different helicities are separated in the xy-plane by the electric field normal to the plane.

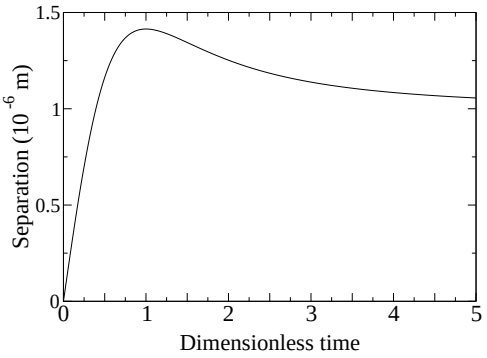

Figure 4.1: Separation l in the y-direction between light holes of opposite helicities as a function of τ, the dimensionless time.

4.5 Constant magnetic field

When a constant uniform magnetic field is present, the gradient correction to the degenerate part of the Hamiltonian takes the form:

$$\Delta_d = -\mathbf{M} \cdot \mathbf{B}. \tag{4.41}$$

$\mathbf{M}$, which is identified with the intrinsic magnetic moment of the wave packet [9, 10, 107], is given by the expression:

$$\mathbf{M} = \pm\frac{e}{2}\Re\langle u|\hat{\mathbf{v}} \times (i\frac{\partial}{\partial \mathbf{k}} - \mathbf{R})|u\rangle =$$
$$= \pm\frac{e}{2}\Re(\eta_i^*\langle u_i|\hat{\mathbf{v}} \times (i\frac{\partial}{\partial \mathbf{k}} - \mathbf{R})|u_j\rangle\eta_j),$$

where the sign is negative forelectrons and positive for holes. The operator $\mathbf{v} = \frac{1}{\hbar}\frac{\partial \hat{\mathrm{H}}_{\mathrm{d}}}{\partial \mathbf{k}}$ is the velocity operator corresponding to the degenerate part of the Hamiltonian.

Written explicitly in component form and restricting our attention to holes, the magnetic moment is:

$$M^\alpha = -\frac{e}{4}\epsilon^{\alpha\beta\gamma}\eta_i^*\langle u_i|\{\hat{v}_\beta, (i\frac{\partial}{\partial k_\gamma} - R^\gamma)\}|u_j\rangle\eta_j. \tag{4.42}$$

$\epsilon^{\alpha\beta\gamma}$ represents the antisymmetric tensor. It is straightforward to prove that

$$\eta_i^*\langle u_i|\hat{\mathbf{v}}|u_j\rangle\eta_j \equiv \eta^\dagger \mathbf{v}\eta = \frac{1}{\hbar}\frac{\partial \epsilon}{\partial \mathbf{k}}, \tag{4.43}$$

in which $\mathbf{v}_{\mathrm{ij}} = \langle \mathrm{u_i}|\hat{\mathbf{v}}|\mathrm{u_j}\rangle$. Therefore the second term in $\mathbf{M}$ is:

$$\mathbf{M}_2 = \frac{e}{2}\eta^\dagger \mathbf{v}\eta \times \mathbf{R} = -\frac{e}{2\hbar}\frac{\partial \epsilon}{\partial \mathbf{k}} \times \mathbf{R}. \tag{4.44}$$

The first term is

$$\mathbf{M}_1 = -\frac{e}{2\hbar}\frac{\partial \epsilon}{\partial \mathbf{k}} \times \mathbf{R} - \frac{e}{2}\Re\sum_{i,j}^{in}\sum_{l}^{out}\eta_i^*\mathbf{v}_{il} \times \mathcal{R}_{lj}\eta_j, \tag{4.45}$$

where 'out' means the sum runs over all bands outside the degenerate subspace,

that is l $\neq$ i, j. The first contribution exactly cancels $\mathbf{M}_2$, so the final result is

$$\mathbf{M} = -\frac{e}{2}\Re\sum_{i,j}^{in}\sum_{l}^{out}\eta_i^*\mathbf{v}_{il}\times\mathcal{R}_{lj}\eta_j. \tag{4.46}$$

Thus the magnetic moment can be expressed purely in terms of matrix elements connecting the degenerate subspace to bands outside the subspace.

We take as an example once again the light-hole manifold of the four-band Luttinger model in the spherical approximation in the presence of a constant uniform magnetic field. The Hamiltonian in this case is:

$$\hat{H} = \hat{H}_{Lutt} - \frac{ge}{m}\hat{\mathbf{S}}\cdot\mathbf{B}, \tag{4.47}$$

where $\hat{\mathrm{H}}_{\mathrm{Lutt}}$ has been defined in (4.27) and g is the Lande g-factor. The first part of the Hamiltonian is $\hat{\mathrm{H}}_{\mathrm{d}}$ while the Zeeman term is $\hat{\mathrm{H}}_{\mathrm{n}}$. The Zeeman interaction between the spin and the magnetic field does not contribute to the velocity operator and therefore it does not contribute to the magnetic moment. The light-hole intrinsic magnetic moment in the spherical four-band model is given by the following expression:

$$\mathbf{M} = \frac{3e\hbar\gamma_2\hat{\mathbf{k}}}{2m}\langle\hat{\sigma}^3\rangle. \tag{4.48}$$

The magnetic moment is proportional to the expectation value of the third component of the pseudo-spin and therefore to the helicity, as shown in (4.39). Depending on the weight of each band in the wave packet the intrinsic magnetic moment can be positive or negative and if the bands are equally represented it will be zero.

Chapter 5

Manifestations of the geometrical phase in the Wigner distribution of Bloch electrons

5.1 Introduction

Carrier dynamics in metals and semiconductors in the presence of external electromagnetic fields, the potentials of which usually vary on scales considerably large than the interatomic spacing, have been conveniently described by semiclassical transport theories. The semiclassical dynamics together with the Boltzmann equation produce accurate descriptions of electrical and thermal conduction [65]. In recent years efforts have been made to extend the semiclassical theory to include spin transport and generation in bulk and two dimensional systems [50, 66, 111]. The attempts to resolve the challenges inherent in treating the transport of non-conserved quantities constitute a vibrant ongoing effort [49, 69, 78, 79, 80, 112, 113, 114, 115, 116].

A fundamental feature of semiclassical transport is its accounting for the finite extent of particles in real and reciprocal space. This feature is most natu-

rally incorporated into the dynamics of wave packets [10, 105] where the notion of a wave packet center in real space and k-space is retained. The carrier dynamics are described in terms of the displacement of these points under the action of external fields. The finite extent of the wave packet has important consequences for transport theory. For example, our recent research on spin transport has shown that, due to the fact that the spin and charge centers of a wave packet do not coincide, the expressions for the spin density, torque and current distributions are, in the language of wave packets, expressed as series of multipole terms [66]. It has been demonstrated, in addition, that the wave packet formalism captures the physics connected with adiabatic motion and the Berry phase, in particular the Berry curvature correction to the semiclassical equations of motion [10]. The Berry phase, which until Berry's seminal article [14] was not taken into account, appears naturally as part of the wave packet distribution function. The Berry curvature correction to the wave packet equations of motion is believed to play an important role in the anomalous Hall effect [15, 64, 91] and in spin transport [49, 50, 66, 69, 111], among other phenomena.

Semiclassical transport theory is not restricted to wave packet dynamics. Wave packets, which may be constructed out of one band eigenstate [10] or out of a superposition of eigenstates of several degenerate bands [108, 117], represent pure states. In order to treat mixed states, or incoherent superpositions of eigenstates, one must resort to a more general formalism. This is the principal motivation of the current chapter.

The most general description of a quantum mechanical system is based on the density operator. In this chapter we start from the density operator and formulate a theory of carrier dynamics in metals and semiconductors. As a first step we focus on non-interacting systems in which disorder is weak, strong spin-orbit interactions are present and a weak slowly-varying electric field is acting. These systems have

come under intense scrutiny in recent years along with the take-off of spintronics [39, 82, 86, 118, 119]. In such non-interacting systems the formalism may be simplified by defining a reduced one-particle density operator [120]. Due to the fact that Bloch bands are clearly resolved the reduced density operator for this system may be expanded in a basis of Bloch wave functions. The density matrix which emerges from this expansion may be defined as a Wigner function. This function can be used to study the single particle dynamics of the carriers and to formulate definitions of the macroscopic quantities involved in the transport of conserved and non-conserved quantities. Four cases can be distinguished, namely a single particle in one band, many particles in one band, a single particle in multiple bands and many particles in multiple bands. Each of these cases appears in our work.

We focus on some fundamental aspects of adiabatic particle dynamics and demonstrate their relevance to transport phenomena. We pay particular attention to the band mixing induced by the electric field, which gives rise to a non-adiabatic correction to the wave functions. We also address several important questions concerning the relationship between the Wigner function formalism and the wave packet formalism [121]. We discuss the way the carrier position is to be found and, where possible, compare the result with the expression for the real space center of a wave packet. The requirement that the particle position be gauge invariant results in a gauge dependent shift in the position at which the Wigner function is evaluated. This shift was also found by Littlejohn and Flynn [122] in the study of coupled wave equations. This gauge dependent shift is equivalent to the freedom of choosing the real space center of a wave packet by changing the phase of the wave function [10, 123]. The gauge dependent shift in the position of the Wigner function must be taken into account when many particle distributions, such as the particle number density, are expressed in the crystal momentum representation. The carrier velocity may be derived directly from the particle position, recovering the Berry phase

physics which is known from previous work [10]. We also show that a spin velocity may be defined in such a way that spin transport may be described in an analogous fashion to charge transport.

We discuss, in addition, important consequences of finite particle size. The Wigner distribution is a quantum entity which takes into account the finite extent of the particles in real and reciprocal space. The distribution of, for example, velocity for a single carrier, may not coincide with that of charge and the macroscopic velocity distribution (and therefore the current distribution) will be composed of a series of multipoles. The spin, spin current and spin torque distributions are in turn composed of series of multipoles which we discuss and compare with those found in our previous work on spin transport [66].

The outline of this chapter is as follows. We present the fundamentals of the single particle density matrix formalism in section 5.2. We determine the position and velocity of the particles, emphasizing the gauge dependent shift in the former and the Berry phase correction in the latter. We calculate the particle spin, torque and spin velocity and define a modified spin velocity which satisfies an equation analogous to the charge velocity. In section 5.3 we demonstrate the modifications which must be made to extend the theory to many non-interacting particles. We define the charge and current densities and show the equation of continuity they satisfy. In the case of spin we define the spin, spin current and spin torque distributions and show the equation of continuity satisfied by them in the clean limit. In section 5.4 we demonstrate the effect of local gauge transformations and the modifications which must be made to the dipoles in the charge and spin related distributions in order to make them gauge covariant. The lengthier calculations are included in the appendix to this chapter.

5.2 Single-particle dynamics

We consider systems in which the Hamiltonian contains spin-orbit interaction terms involving the carrier spins and the lattice periodic potential. We restrict our attention to the limit in which this spin-orbit interaction is sizably stronger than the disorder broadening and the thermal broadening. In this limit the system may also be described in terms of well defined bands which we will take to be represented by a set of Bloch wave functions.

5.2.1 Density matrix

We take the system under study to be described by a density operator $\hat{\rho}$. Methods of finding the density operator have been studied extensively in the past [55, 59, 89, 124, 125, 126, 127, 128] and will not be our concern in this paper. We assume the density operator to be known and study its role in transport. The density operator may be expanded in the Hilbert space spanned by a complete orthonormal set of Bloch wave functions $|\psi_{\mathrm{n}\mathbf{k}}\rangle$ as:

$$\hat{\rho} = \sum_{n,n'} \sum_{\mathbf{k},\mathbf{k}'} \rho_{nn'}(\mathbf{k},\mathbf{k}') |\psi_{n\mathbf{k}}\rangle \langle \psi_{n'\mathbf{k}'}|. \tag{5.1}$$

In the approach we follow, all the time dependence of the density operator is contained in the wave functions, so that $\rho_{\mathrm{nn}'}(\mathbf{k},\mathbf{k}')$ does not have time dependence. In thermal equilibrium it is diagonal and its elements, $\rho_{\mathrm{nn}}(\mathbf{k},\mathbf{k})$, are equal to the Fermi-Dirac function $\mathrm{f}_0(\varepsilon_{\mathrm{n}})$, where ε_{n} is the band energy.

The expectation value of any operator $\hat{\mathrm{A}}$ is found from the formula:

$$\begin{aligned} \langle \hat{A} \rangle = \mathrm{Tr}\, \hat{\rho}\hat{A} \equiv \sum_{n,n'} \sum_{\mathbf{k},\mathbf{k}'} \langle \psi_{n\mathbf{k}}|\hat{\rho}|\psi_{n'\mathbf{k}'}\rangle \langle \psi_{n'\mathbf{k}'}|\hat{A}|\psi_{n\mathbf{k}}\rangle = \\ = \sum_{\mathbf{k},\mathbf{k}'} \mathrm{tr}\, \rho(\mathbf{k},\mathbf{k}') A(\mathbf{k}',\mathbf{k}). \end{aligned} \tag{5.2}$$

The notation $\mathrm{A}(\mathbf{k}', \mathbf{k})$ stands for the matrix elements of the operator $\hat{\mathrm{A}}$, namely $\mathrm{A_{n'n}} = \langle \psi_{\mathrm{n'k'}} | \hat{\mathrm{A}} | \psi_{\mathrm{nk}} \rangle$. The tr operation is simply the matrix trace, which does not include the wave vector summation. If the operator $\hat{\mathrm{A}}$ is replaced by the identity matrix we obtain $\mathrm{Tr}\,\hat{\rho} = \sum_{\mathbf{k}} \mathrm{tr}\,\rho(\mathbf{k}, \mathbf{k})$. For a single particle, the normalization condition on the density operator is $\mathrm{Tr}\,\hat{\rho} = 1$.

In order to make transparent the analogy with the language of wave packets, center of mass and relative coordinates may be defined in k-space such that $\mathbf{Q} = \mathbf{k} - \mathbf{k}'$ and $\mathbf{q} = \frac{\mathbf{k}+\mathbf{k}'}{2}$. The density operator as a function of these coordinates can be reexpressed as:

$$\hat{\rho} = \sum_{n,n'} \sum_{\mathbf{q},\mathbf{Q}} \rho_{nn'}(\mathbf{q}_+, \mathbf{q}_-) |\psi_{n+}\rangle \langle \psi_{n'-}|, \tag{5.3}$$

where $|\psi_{\mathrm{n}\pm}\rangle = \mathrm{e}^{\mathrm{i}\mathbf{q}_\pm \cdot \hat{\mathbf{r}}} |\mathrm{u}_{\mathrm{n}\pm}\rangle$ and $|\mathrm{u}_{\mathrm{n}\pm}\rangle$ is the periodic part of the Bloch wave at $\mathbf{q}_\pm = \mathbf{q} \pm \frac{\mathbf{Q}}{2}$.

We rewrite the one particle density matrix by means of a Wigner transformation:

$$\rho_{nn'}(\mathbf{q}_+, \mathbf{q}_-) = \int d^3r\, e^{-i\mathbf{Q}\cdot\mathbf{r}} \rho_{nn'}(\mathbf{q}, \mathbf{r}). \tag{5.4}$$

For the sake of concreteness the integrals are represented as three dimensional. Nevertheless, the theory applies to systems of any dimensionality. The inverse transformation is:

$$\rho_{nn'}(\mathbf{q}, \mathbf{r}) = \int \frac{d^3Q}{(2\pi)^3}\, e^{i\mathbf{Q}\cdot\mathbf{r}} \rho_{nn'}(\mathbf{q}_+, \mathbf{q}_-), \tag{5.5}$$

defining the Wigner function, which plays the role of a distribution function. Technically, however, it is not a distribution since it may take negative values [120]. Finally, replacing the vector summations by integrations we are able to represent

the density operator in the following form:

$$\hat{\rho} = \int\int d^3q\, d^3r \sum_{nn'} \rho_{nn'}(\mathbf{q},\mathbf{r}) \int \frac{d^3Q}{(2\pi)^3} \times \\ \times e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})} |u_{n+}\rangle\langle u_{n'-}| e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}. \tag{5.6}$$

Henceforth the Wigner function $\rho_{\mathrm{nn'}}(\mathbf{q},\mathbf{r})$ will be abbreviated to ρ. The variables $\mathbf{q}$ and $\mathbf{r}$ are simply labels for the carriers, not physical observables. In particular, the dummy variable $\mathbf{r}$ in the Fourier expansion of the density matrix is simply the Fourier dual of $\mathbf{Q}$ and must not be confused with the position operator $\hat{\mathbf{r}}$ appearing in the density operator. It does not correspond to an actual position.

5.2.2 Carrier position

If we consider a particle in only one band, labelled n, the Wigner function ρ reduces to a scalar. The Wigner function for band n will be denoted by ρ_{n}. In this limit, a clear connection can be made with the dynamics of wave packets. The position of the particle can be found by calculating the expectation value of the position operator:

$$\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \int\int \frac{d^3q\, d^3r}{(2\pi)^3}\, \rho_n(\mathbf{q},\mathbf{r})(\mathbf{r}+\mathcal{R}_n). \tag{5.7}$$

Here $\mathcal{R}_{\mathrm{n}} \equiv \mathcal{R}_{\mathrm{nn}} = \langle \mathrm{u_n} | \mathrm{i}\frac{\partial \mathrm{u_n}}{\partial \mathbf{q}} \rangle$. The integrand is not gauge invariant but the integral can be shown to be by changing the variable of integration $\mathbf{r}$ to $\mathbf{r'} = \mathbf{r} + \mathcal{R}_{\mathrm{n}}$. The connection $\mathcal{R}_{\mathrm{n}}$ has no position dependence, therefore the Jacobian of the transformation is unity. The expectation value of the position operator is:

$$\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \int\int \frac{d^3q\, d^3r'}{(2\pi)^3}\, \rho_n(\mathbf{q},\mathbf{r'}-\mathcal{R}_n)\,\mathbf{r'}. \tag{5.8}$$

The gauge invariance of this expression will emerge below. We conclude that $\mathbf{r}$ is to be interpreted as a label for the charge carrier, while the effective particle position

is $\mathbf{r}' = \mathbf{r} + \mathcal{R}_{\mathrm{n}}$. Neither the label $\mathbf{r}$ nor the gauge field $\mathcal{R}$ are by themselves gauge invariant, but together they form a gauge invariant quantity which represents the true position of the carrier. This result was also found earlier, in somewhat different circumstances, in the work of Littlejohn and Flynn [122]. It reflects the freedom of changing the gauge of the wave functions $|\psi_{\mathrm{n}\mathbf{k}}\rangle$, $|\psi_{\mathrm{n}'\mathbf{k}'}\rangle$ in (5.1) by altering their phase. It is the same freedom one has in defining the center of mass of a wave packet, as in the paper of Sundaram and Niu [10], by changing the overall phase of the wave function [123].

For multiple bands, the expression for the particle position is:

$$\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \int\int \frac{d^3q\,d^3r}{(2\pi)^3}\,\rho_{nn'}(\mathbf{q},\mathbf{r})(\mathbf{r}\delta_{n'n} + \mathcal{R}_{n'n}). \tag{5.9}$$

Henceforth, the Einstein summation convention will be used. One can write the above formally as:

$$\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \mathrm{tr}\int\int \frac{d^3q\,d^3r}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{r})\,(\mathbf{r} + \mathcal{R}). \tag{5.10}$$

and make the substitution $\mathbf{r}' = \mathbf{r} + \mathcal{R}$ as in the single band case. Although $\mathcal{R}$ is a matrix the Jacobian of this transformation is unity. Finally, the expectation value of the position operator can be expressed formally as:

$$\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \int\int \frac{d^3q\,d^3r'}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{r}' - \mathcal{R})\,\mathbf{r}'. \tag{5.11}$$

Unlike the single band case, the expression $\rho(\mathbf{q},\mathbf{r}' - \mathcal{R})$ is only a formal abbreviation for $\rho(\mathbf{q},\mathbf{r}') - \mathcal{R}\cdot\nabla_{\mathbf{r}'}\rho(\mathbf{q},\mathbf{r}')$ + higher order terms.

Expression (5.8) helps us identify the positions of the particles. In order to determine if a particle is localized at its position we must calculate the variance of the position operator, that is the expectation value $\langle\hat{\mathbf{r}}^2\rangle - \langle\hat{\mathbf{r}}\rangle^2$. It is shown in the appendix that the variance does not diverge. This result applies to both the single

band and the multiple band cases.

5.2.3 Carrier velocity in the presence of an electric field

We will investigate systems acted on by a weak external electric field. The effect of this electric field can be incorporated fully into the gauge invariant crystal momentum through the addition of the electromagnetic vector potential $\mathbf{A} = -\mathbf{E}\mathrm{t}$. Because the Bloch functions must retain translational symmetry the electromagnetic vector potential does not enter the travelling wave part of the wave functions, which have the form $|\psi_{\mathrm{n}\tilde{\mathbf{q}}}\rangle = \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}|\mathrm{u}_{\mathrm{n}\tilde{\mathbf{q}}}\rangle$. The wave functions $|\mathrm{u}_{\mathrm{n}\tilde{\mathbf{q}}}\rangle$ depend implicitly on time only through the crystal wave vector $\tilde{\mathbf{q}} = \mathbf{q} + \frac{\mathrm{e}\mathbf{A}}{\hbar}$. However, the presence of the external electric field results in a non-adiabatic mixing of the bands with the effect that the perturbed lattice periodic Bloch wave functions, $|\bar{\mathrm{u}}_{\mathrm{m}\tilde{\mathbf{q}}}\rangle$ have the following form [129] to first order in the electric field:

$$|\bar{u}_{m\tilde{\mathbf{q}}}\rangle = e^{i\phi_m}(|u_{m\tilde{\mathbf{q}}}\rangle - \sum_{n(\neq m)} \frac{\langle u_{n\tilde{\mathbf{q}}}|i\hbar\frac{d}{dt}|u_{m\tilde{\mathbf{q}}}\rangle}{\varepsilon_m - \varepsilon_n}|u_{n\tilde{\mathbf{q}}}\rangle). \tag{5.12}$$

The phase $\phi_{\mathrm{m}}(\tilde{\mathbf{q}}, \mathrm{t})$ includes the dynamical phase and the Berry phase. The differential $\frac{\mathrm{d}}{\mathrm{dt}}$ is equivalent to $\dot{\tilde{\mathbf{q}}} \cdot \frac{\partial}{\partial \tilde{\mathbf{q}}}$ since $\frac{\partial|\mathrm{u}_{\mathrm{n}\tilde{\mathbf{q}}}\rangle}{\partial \mathrm{t}} = 0$. The result expressed by (5.12) is general. Moreover, although its derivation relies on the assumption that the bands are non-degenerate it can be shown that, when calculating intrinsic contributions to transport (for the definition of intrinsic please see the appendix and our recent work [111]), the result also holds for degenerate bands with the difference that the sum must exclude *all* bands which are degenerate in energy with band m. The proof of this statement is given in the appendix. Henceforth, $|\mathrm{u}_{\mathrm{m}\tilde{\mathbf{q}}}\rangle$ and $|\bar{\mathrm{u}}_{\mathrm{m}\tilde{\mathbf{q}}}\rangle$ will be abbreviated to $|\mathrm{u}_{\mathrm{m}}\rangle$ and $|\bar{\mathrm{u}}_{\mathrm{m}}\rangle$ respectively. As stated above, given that the $|\bar{\mathrm{u}}_{\mathrm{m}}\rangle$ are functions of $\tilde{\mathbf{q}}$ only, one may replace $\frac{\mathrm{d}}{\mathrm{dt}}$ in (5.12) by $\dot{\tilde{\mathbf{q}}} \cdot \frac{\partial}{\partial \tilde{\mathbf{q}}}$, where $\dot{\tilde{\mathbf{q}}} = -\frac{\mathbf{eE}}{\hbar}$.

Equation (5.12) can then be written as:

$$|\bar{u}_m\rangle = e^{i\phi_m}(|u_m\rangle + e\mathbf{E}\cdot\sum_{n(\neq m)}\frac{\mathcal{R}_{nm}}{\varepsilon_m-\varepsilon_n}|u_n\rangle), \tag{5.13}$$

where the connection $\mathcal{R}_{\rm nm} = \langle {\rm u_n}|{\rm i}\frac{\partial}{\partial\tilde{\mathbf{q}}}|{\rm u_m}\rangle$. The $|\bar{\rm u}_{\rm m}\rangle$ form a complete set. They are, however, not eigenstates of the *time dependent* Hamiltonian of the system under study $\tilde{\rm H}\equiv\tilde{\rm H}(\tilde{\mathbf{q}})$.

We note that in the evaluations of matrix elements in this paper the only property of the basis functions that is used is their Bloch periodicity. Therefore the results which are expressed in terms of the lattice periodic Bloch functions hold as well for the $|\bar{\rm u}_{\rm n}\rangle$ as for the $|{\rm u_n}\rangle$.

It is customary to consider only a subset of the Hilbert space which contains the bands that are relevant to transport, which usually refers to the topmost filled valence bands and the lowest filled conduction bands. Since the gauge of the $|\bar{\rm u}_{\rm m}\rangle$ is not fixed we impose the following gauge fixing condition in the subspace under consideration:

$$\begin{aligned}\langle\bar{u}_n|i\hbar\frac{d}{dt}|\bar{u}_m\rangle &= \langle\bar{u}_n|\tilde{H}|\bar{u}_m\rangle,\ \text{ for } n \text{ in subspace}\\ &= 0,\ \text{ for } n \text{ out of subspace}\end{aligned} \tag{5.14}$$

This condition fixes the phase(s) $\phi_{\rm m}$ in (5.12). Henceforth, we shall work only with the basis set $\{|\bar{\rm u}_{\rm n}\rangle\}$.

In the absence of disorder, since all the time dependence is contained in the wave functions, the density matrix $\rho_{\rm nn'}(\mathbf{k},\mathbf{k'})$ in (5.1) can only depend on the wave vector $\mathbf{k}$, not on the crystal wave vector $\tilde{\mathbf{k}}$. As a result the Wigner function $\rho_{\rm nn'}(\mathbf{q},\mathbf{r})$ only depends on the wave vector $\mathbf{q}$, not on $\tilde{\mathbf{q}}$.

The particle velocity can be derived directly from (5.9) by evaluating the

time derivative:

$$\frac{d}{dt}\,\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \mathrm{tr}\int\int \frac{d^3q\,d^3r}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{r})\frac{d\bar{\mathcal{R}}}{dt}. \tag{5.15}$$

The differential $\frac{\mathrm{d}\bar{\mathcal{R}}}{\mathrm{dt}}$ becomes, using equation (5.14):

$$\begin{aligned}
\frac{d\bar{\mathcal{R}}_{n'n}}{dt} &= \langle \bar{u}_{n'}|i\frac{d}{dt}|\frac{\partial \bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle + \langle\frac{d\bar{u}_{n'}}{dt}|i\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle = \\
&= \langle \bar{u}_{n'}|i\frac{\partial}{\partial\tilde{\mathbf{q}}}|\frac{d\bar{u}_n}{dt}\rangle + \langle\frac{d\bar{u}_{n'}}{dt}|i\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle = \\
= \frac{\partial}{\partial\tilde{\mathbf{q}}}\langle \bar{u}_{n'}|i\frac{d\bar{u}_n}{dt}\rangle &- \langle\frac{\partial\bar{u}_{n'}}{\partial\tilde{\mathbf{q}}}|i\frac{d\bar{u}_n}{dt}\rangle + \langle\frac{d\bar{u}_{n'}}{dt}|i\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle = \\
&= \frac{1}{\hbar}\frac{\partial\bar{H}_{n'n}}{\partial\tilde{\mathbf{q}}} + i(\langle\frac{d\bar{u}_{n'}}{dt}|\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle - \langle\frac{\partial\bar{u}_{n'}}{\partial\tilde{\mathbf{q}}}|\frac{d\bar{u}_n}{dt}\rangle).
\end{aligned} \tag{5.16}$$

The abbreviation $\bar{\mathrm{H}}_{\mathrm{n'n}}$ stands for $\langle \bar{\mathrm{u}}_{\mathrm{n'}}|\tilde{\mathrm{H}}|\bar{\mathrm{u}}_{\mathrm{n}}\rangle$. In one band it is easily shown that (5.16) becomes:

$$\frac{d\bar{\mathcal{R}}_n}{dt} = \frac{1}{\hbar}\frac{\partial\varepsilon_n}{\partial\tilde{\mathbf{q}}} - \dot{\tilde{\mathbf{q}}}\times\bar{\mathbf{\Omega}}_{\mathbf{n}}. \tag{5.17}$$

The Berry curvature for band n is given by:

$$\bar{\mathbf{\Omega}}_n = 2\,\mathrm{Im}\,\langle\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}|\times|\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle = \nabla_{\tilde{\mathbf{q}}}\times\bar{\mathcal{R}}_n, \tag{5.18}$$

where $\nabla_{\tilde{\mathbf{q}}}\times$ represents the curl operator in wave vector space. Equation (5.17) reproduces the semiclassical equation of motion found by Sundaram and Niu. Writing $\dot{\tilde{\mathbf{q}}} = -\frac{\mathbf{eE}}{\hbar}$ we obtain for electrons in a single band:

$$\frac{d}{dt}\,\mathrm{Tr}\,\hat{\rho}\hat{\mathbf{r}} = \int\int\frac{d^3q\,d^3r}{(2\pi)^3}\,\rho_n(\frac{1}{\hbar}\frac{\partial\varepsilon_n}{\partial\tilde{\mathbf{q}}} + \frac{e\mathbf{E}}{\hbar}\times\bar{\mathbf{\Omega}}_n). \tag{5.19}$$

For multiple bands, it is useful to introduce the covariant derivatives $\frac{\mathrm{D}}{\mathrm{D}\tilde{\mathbf{q}}} = \frac{\partial}{\partial\tilde{\mathbf{q}}} - \mathrm{i}\mathcal{R}$ and $\frac{\mathrm{D}}{\mathrm{Dt}} = \frac{\partial}{\partial\mathrm{t}} - \mathrm{i}\mathcal{A}$, where $\mathcal{A}_{\mathrm{n'n}} = \langle \mathrm{u}_{\mathrm{n'}}|\mathrm{i}\frac{\mathrm{du_n}}{\mathrm{dt}}\rangle$, as done in [117]. Then equation (5.16)

can be written in a manifestly gauge covariant form:

$$\frac{d\bar{\mathcal{R}}^{\alpha}_{n'n}}{dt} = \frac{1}{\hbar}[\frac{D}{D\tilde{q}_{\alpha}}, \bar{H}]_{n'n} + i[\frac{D}{Dt}, \frac{D}{D\tilde{q}_{\alpha}}]_{n'n}. \tag{5.20}$$

5.2.4 Carrier spin, spin torque and spin velocity in the presence of an electric field

To evaluate the carrier spin consider the expectation value of the operator $\hat{\mathrm{s}}$, which stands for any one component of the spin operator. The result is:

$$\mathrm{Tr}\,\hat{\rho}\hat{s} = \mathrm{tr}\int\int \frac{d^3q\,d^3r}{(2\pi)^3}\,\rho\bar{s}, \tag{5.21}$$

where $\bar{\mathrm{s}}_{\mathrm{n'n}} = \langle \bar{\mathrm{u}}_{\mathrm{n'}}|\hat{\mathrm{s}}|\bar{\mathrm{u}}_{\mathrm{n}}\rangle$ are the matrix elements of the spin operator.

Associated with the spin of every carrier is a torque, which accounts for the non-conservation of spin. This torque can be found by evaluating the expectation value of $\hat{\tau} = \frac{\mathrm{i}}{\hbar}[\tilde{\mathrm{H}}, \hat{\mathrm{s}}]$. The result is:

$$\mathrm{Tr}\,\hat{\rho}\hat{\tau} = \mathrm{tr}\int\int \frac{d^3q\,d^3r}{(2\pi)^3}\,\rho\bar{\tau}, \tag{5.22}$$

where $\bar{\tau}_{\mathrm{n'n}} = \langle \bar{\mathrm{u}}_{\mathrm{n'}}|\hat{\tau}|\bar{\mathrm{u}}_{\mathrm{n}}\rangle$.

To determine the spin velocity we calculate the expectation value $\langle\hat{\mathrm{s}}\hat{\mathbf{v}}\rangle$, where products of non-commuting operators are assumed to be symmetrized. The result is:

$$\mathrm{Tr}\,\hat{\rho}\hat{s}\hat{\mathbf{v}} = \mathrm{tr}\int\int \frac{d^3q\,d^3r}{(2\pi)^3}\,\rho\bar{\mathbf{v}}^{s}. \tag{5.23}$$

The spin velocity $\bar{\mathbf{v}}^{\mathrm{s}}$ can be written as:

$$\bar{\mathbf{v}}^{s} = [\frac{D}{Dt}, \bar{\boldsymbol{\Xi}}^{s}] - \bar{\boldsymbol{\Xi}}^{\tau}. \tag{5.24}$$

The spin and spin torque gauge fields, $\bar{\Xi}^{\mathrm{s}}$ and $\bar{\Xi}^{\tau}$, are defined by:

$$\begin{aligned}\bar{\Xi}^{s}_{n'n} &= \frac{i}{2}(\langle \bar{u}_{n'}|\hat{s}|\frac{\partial \bar{u}_n}{\partial \mathbf{q}}\rangle - \langle \frac{\partial \bar{u}_{n'}}{\partial \mathbf{q}}|\hat{s}|\bar{u}_n\rangle) \\ \bar{\Xi}^{\tau}_{n'n} &= \frac{i}{2}(\langle \bar{u}_{n'}|\hat{\tau}|\frac{\partial \bar{u}_n}{\partial \mathbf{q}}\rangle - \langle \frac{\partial \bar{u}_{n'}}{\partial \mathbf{q}}|\hat{\tau}|\bar{u}_n\rangle).\end{aligned} \tag{5.25}$$

It will prove useful to make $\bar{\Xi}^{\mathrm{s}}$ and $\bar{\Xi}^{\tau}$ gauge covariant by defining $\bar{\mathbf{p}}^{\mathrm{s}} = \bar{\Xi}^{\mathrm{s}} - \frac{1}{2}\{\bar{\mathcal{R}}, \bar{\mathrm{s}}\}$ and $\bar{\mathbf{p}}^{\tau} = \bar{\Xi}^{\tau} - \frac{1}{2}\{\bar{\mathcal{R}}, \bar{\tau}\}$ respectively. The discussion of these terms is deferred to sections 5.3 and 5.4, where it will become evident that $\bar{\mathbf{p}}^{\mathrm{s}}$ and $\bar{\mathbf{p}}^{\tau}$ can be identified with a spin dipole and a torque dipole respectively. The spin velocity can then be written in a manifestly gauge covariant form as:

$$\bar{\mathbf{v}}^{s} = \frac{1}{2}\{[\frac{D}{Dt}, \bar{\mathcal{R}}], \bar{s}\} + [\frac{D}{Dt}, \bar{\mathbf{p}}^{s}] - \bar{\mathbf{p}}^{\tau}. \tag{5.26}$$

Equation (5.26) is the gauge covariant form of equation (10) in Ref. [66], the integrand of which represents the spin velocity. The first term is a convective contribution. It represents a moving electron transporting its spin along with it. The second term is the time derivative of the spin dipole. It is present because electrons occupy a finite fraction of real and reciprocal space and the spin distribution for an electron may not be the same as the charge distribution. When using the center of charge as the reference, multipole terms must be present in the spin distribution. The last term is the torque dipole which takes into account the non-conservation of spin.

We may incorporate the torque dipole term into a modified velocity, which we shall call $\bar{\mathbf{v}}^{\mathrm{t}}$, by $\bar{\mathbf{v}}^{\mathrm{t}} = \bar{\mathbf{v}}^{\mathrm{s}} + \bar{\mathbf{p}}^{\tau}$. The modified spin velocity is given simply by:

$$\bar{\mathbf{v}}^{t} = \frac{1}{2}\{[\frac{D}{Dt}, \bar{\mathcal{R}}], \bar{s}\} + [\frac{D}{Dt}, \bar{\mathbf{p}}^{s}]. \tag{5.27}$$

This velocity will play an important role in the following section, when the spin current is introduced.

5.3 Many particle distributions

We will consider the distributions of charge and spin starting from the formalism we have developed. The particle number density is defined by the formula:

$$n(\mathbf{R},t) = \mathrm{Tr}[\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})]. \tag{5.28}$$

Whereas the operator $\hat{\rho}$ is a one particle density operator, the quantity $\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})$ plays the role of a many particle density operator, the trace of which yields the particle number density. Because the variance of the position operator is finite, the δ-function may be expanded in a Taylor series about the dummy variable $\mathbf{r}$ as $\delta(\mathbf{R}-\hat{\mathbf{r}}) = \delta(\mathbf{R}-\mathbf{r}) - \nabla_{\mathbf{R}}\cdot(\hat{\mathbf{r}}-\mathbf{r})\delta(\mathbf{R}-\mathbf{r}) + \mathcal{O}[(\hat{\mathbf{r}}-\mathbf{r})^2]$. The expansion is truncated at the first order for simplicity but in the final result we will recover all the terms. The number density can be expressed as:

$$n(\mathbf{R},t) = \mathrm{tr}\int d^3q\,[\rho(\mathbf{q},\mathbf{R}) - \nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\mathcal{R}}]. \tag{5.29}$$

Note that the gauge field $\bar{\mathcal{R}}$ plays the role of a dipole. This equation can be formally written the following way:

$$n(\mathbf{R},t) = \mathrm{tr}\int d^3q\,\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}}). \tag{5.30}$$

The above is tantamount to making, in the density matrix, the replacement $\rho(\mathbf{q},\mathbf{r}) \to \rho(\mathbf{q},\mathbf{r}-\bar{\mathcal{R}})$. To ensure the gauge covariance of the number density the Fourier dual $\mathbf{r}$ is replaced by the effective position $\mathbf{r}+\bar{\mathcal{R}}$. This is the same position as found in section II. If the subspace contains one band expression (5.30) is not just a formal way of writing (5.29).

5.3.1 Electrical charge and current densities

The charge density is defined by the formula:

$$n^q(\mathbf{R},t) = q\, n(\mathbf{R},t) = q\,\mathrm{Tr}[\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})]. \tag{5.31}$$

The charge q is not to be confused with the wave vector $\mathbf{q}$. The charge current density $\mathbf{J}^{\mathrm{q}}$ is defined by the equation:

$$\mathbf{J}^q(\mathbf{R},t) = q\,\mathrm{Tr}[\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})\hat{\mathbf{v}}]. \tag{5.32}$$

The charge equation of continuity in the absence of external sources,

$$\frac{\partial n^q}{\partial t} + \nabla\cdot\mathbf{J}^q = 0, \tag{5.33}$$

is readily verified from the first principles definitions of the charge and current densities.

When the δ-function is expanded the current density takes the form:

$$\mathbf{J}^q(\mathbf{R},t) = q\,\mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{v}} - \nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{\Theta}}]. \tag{5.34}$$

The velocity matrix elements $\bar{\mathbf{v}}_{\mathrm{n'n}} = \langle \bar{\mathrm{u}}_{\mathrm{n'}}|\tilde{\mathbf{v}}|\bar{\mathrm{u}}_{\mathrm{n}}\rangle$ as shown in the appendix, where $\tilde{\mathbf{v}} = \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}\hat{\mathbf{v}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\hat{\mathbf{r}}}$. The velocity gauge field $\bar{\mathbf{\Theta}}$, which plays the role of a velocity dipole, is defined by:

$$\bar{\mathbf{\Theta}}^{\alpha}_{n'n} = \frac{i}{2}(\langle\bar{u}_{n'}|\tilde{v}^{\alpha}|\frac{\partial\bar{u}_n}{\partial\mathbf{q}}\rangle - \langle\frac{\partial\bar{u}_{n'}}{\partial\mathbf{q}}|\tilde{v}^{\alpha}|\bar{u}_n\rangle). \tag{5.35}$$

It plays the same role as the gauge field $\bar{\mathcal{R}}$ in the analogous term in the charge density.

In homogeneous systems the gradient term in (5.34) drops out and, based on

equations (5.15) and (5.34), the charge current can be expressed as the time rate of change of the gauge field $\bar{\mathcal{R}}$:

$$\mathbf{J}^q(\mathbf{R},t) = q\,\mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{R})\frac{d\bar{\mathcal{R}}}{dt}. \tag{5.36}$$

5.3.2 Spin, spin current and spin torque densities

The spin density is defined as

$$S(\mathbf{R},t) = \mathrm{Tr}[\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})\hat{s}] \tag{5.37}$$

while the torque density takes the form:

$$\mathcal{T}(\mathbf{R},t) = \mathrm{Tr}[\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})\hat{\tau}] \tag{5.38}$$

and the spin current is defined as

$$\mathbf{J}^s(\mathbf{R},t) = \mathrm{Tr}[\hat{\rho}\delta(\mathbf{R}-\hat{\mathbf{r}})\hat{s}\hat{\mathbf{v}}]. \tag{5.39}$$

In the absence of disorder the spin equation of continuity is:

$$\frac{\partial S}{\partial t} + \nabla\cdot\mathbf{J}^s = \mathcal{T}, \tag{5.40}$$

which is verified from the first principles definitions introduced above.

The δ-functions are expanded in the same way as for the particle number density, whereupon the spin density can be expressed as:

$$S(\mathbf{R},t) = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R})\bar{s} - \nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\Xi}^s], \tag{5.41}$$

and the torque density is:

$$\mathcal{T}(\mathbf{R},t) = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R})\bar{\tau} - \nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\Xi}^{\tau}]. \tag{5.42}$$

The matrices $\bar{\mathrm{s}}$, $\bar{\tau}$, $\bar{\Xi}^{\mathrm{s}}$ and $\bar{\Xi}^{\tau}$ have been defined in the previous section. Note that $\bar{\Xi}^{\mathrm{s}}$ and $\bar{\Xi}^{\tau}$ play the role of spin dipole and torque dipole terms in the corresponding distributions.

Ignoring the gradient term, the spin current is:

$$\mathbf{J}^s(\mathbf{R},t) = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{v}}^s. \tag{5.43}$$

In analogy with the modified spin velocity of the previous section, a modified spin current $\mathbf{J}^{\mathrm{t}}$ may be defined by:

$$\begin{aligned}\mathbf{J}^t(\mathbf{R},t) &= \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{v}}^t = \\ &= \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{R})(\frac{1}{2}\{[\frac{D}{Dt},\bar{\mathcal{R}}],\bar{s}\} + [\frac{D}{Dt},\bar{\mathbf{p}}^s]).\end{aligned} \tag{5.44}$$

The equation of continuity satisfied by this current in the clean limit is:

$$\frac{\partial S}{\partial t} + \nabla\cdot\mathbf{J}^t = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,\rho(\mathbf{q},\mathbf{R})\bar{\tau}. \tag{5.45}$$

5.4 Gauge transformations

We discuss the way the density matrix and connection matrix change under gauge transformations. A general local gauge transformation is represented by $|\bar{\mathrm{u}}_{\mathrm{m}}\rangle \rightarrow \mathrm{O}_{\mathrm{mn}}|\bar{\mathrm{u}}_{\mathrm{n}}\rangle$, where $\mathrm{O}_{\mathrm{mn}} = \mathrm{O}_{\mathrm{mn}}(\tilde{\mathbf{q}})$. Under this operation we have that:

$$\bar{\mathcal{R}} \rightarrow \tilde{\mathcal{R}} + i(O^{-1}\frac{\partial O}{\partial \tilde{\mathbf{q}}}), \tag{5.46}$$

in which $\tilde{\mathcal{R}}$ stands for $\mathrm{O}\bar{\mathcal{R}}\mathrm{O}^{-1}$. The Wigner function itself also changes under this transformation. Using (5.4), we find that, to first order in O, the Wigner function changes as

$$\rho \to \tilde{\rho} + \frac{i}{2}\nabla_{\mathbf{R}} \cdot \{\tilde{\rho}, O^{-1}\frac{\partial O}{\partial \tilde{\mathbf{q}}}\}, \tag{5.47}$$

where $\tilde{\rho} = \mathrm{O}^{-1}\rho\mathrm{O}$ (the opposite of the gauge connection $\bar{\mathcal{R}}$).

The Berry curvature for one band $\bar{\mathbf{\Omega}}_{\mathrm{n}}$ is discussed extensively in Sundaram and Niu [10]. One important additional detail which emerges from the above and equation (5.18) is the fact that, if the curvature is non-zero, one cannot make a gauge transformation to eliminate $\bar{\mathcal{R}}_{\mathrm{n}}$. It can be easily shown that, whereas $\bar{\mathcal{R}}_{\mathrm{n}}$ is gauge dependent, the Berry curvature is gauge covariant. Therefore if $\bar{\mathbf{\Omega}}_{\mathrm{n}}$ is non-zero in one gauge it is non-zero in all gauges and there is no gauge in which $\bar{\mathcal{R}}_{\mathrm{n}}$ can be zero.

All the macroscopic densities defined in the previous sections are covariant under a local gauge transformation. This will be demonstrated for the spin density. Under the above gauge transformation:

$$\begin{aligned}\bar{\mathbf{\Xi}}^{s}_{n'n} \to \frac{i}{2}(\langle \bar{u}_{n'}|\tilde{s}|\frac{\partial \bar{u}_n}{\partial \tilde{\mathbf{q}}}\rangle - \langle \frac{\partial \bar{u}_{n'}}{\partial \tilde{\mathbf{q}}}|\tilde{s}|\bar{u}_n\rangle) + \\ + \frac{i}{2}\{\tilde{s}, O^{-1}\frac{\partial O}{\partial \tilde{\mathbf{q}}}\}_{n'n}\end{aligned} \tag{5.48}$$

where the transformed operator $\tilde{\mathrm{s}} = \mathrm{O}^{-1}\hat{\mathrm{s}}\mathrm{O}$ and the abbreviation $\{\}_{\mathrm{n'n}}$ stands for the matrix elements of the operator product in brackets between the states $\langle \bar{\mathrm{u}}_{\mathrm{n'}}|$ and $|\bar{\mathrm{u}}_{\mathrm{n}}\rangle$. To first order the extra terms acquired by the spin density under a gauge transformation are:

$$\mathrm{tr}(\tilde{\rho}\{\tilde{s}, O^{-1}\frac{\partial O}{\partial \tilde{\mathbf{q}}}\} - \{\tilde{\rho}, O^{-1}\frac{\partial O}{\partial \tilde{\mathbf{q}}}\}\tilde{s}) = 0, \tag{5.49}$$

so that the spin density remains gauge covariant. The cancellation remains true for

all orders in the expansion.

Similarly, the charge and current densities do not acquire additional terms under the local gauge transformation introduced above. The extra terms appearing as a result of the transformation cancel when the trace is taken. When the change in the Wigner function under a gauge transformation is taken into account the overall expressions are gauge covariant. However, if one takes a closer look at the charge current it is evident that, while the current itself is gauge covariant its individual constituents are not. In particular the vector $\bar{\mathbf{\Theta}}$ is not gauge covariant. We will define a gauge covariant charge velocity dipole by the equation $\bar{\mathbf{p}}^{\mathrm{q}} = \mathrm{q}(\bar{\mathbf{\Theta}} - \frac{1}{2}\{\bar{\mathcal{R}}, \bar{\mathbf{v}}\})$ and re-express the current in such a way that each individual term is gauge covariant. Then the current is:

$$\begin{aligned} \mathbf{J}^q(\mathbf{R},t) = q\,\mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{v}} - \\ -\nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\mathcal{R}}\bar{\mathbf{v}} - \nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{p}}^q]. \end{aligned} \tag{5.50}$$

Following the same steps as for the charge density, we can formally rewrite the current density as:

$$\begin{aligned} \mathbf{J}^q(\mathbf{R},t) = q\,\mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}})\bar{\mathbf{v}} - \\ -\nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}})\bar{\mathbf{p}}^q]. \end{aligned} \tag{5.51}$$

The gauge covariant velocity dipole provides an analogy with wave packets. The argument has been made for spin in the discussion above, but it will be reiterated here for the sake of clarity. A wave packet will in general possess a charge distribution and a velocity distribution. The centers of the two may not coincide. The real space center of the wave packet is the center of charge and mass, therefore, if one is calculating the velocity distribution while using the center of charge as the reference one must include multipole terms in the velocity distribution to compensate for this.

Let us take a closer look at the spin density, as given in (5.41). As in the case of charge transport if one were to consider, for example, the integrand of the dipole term in (5.41) without the distribution function, this quantity would not by itself be gauge covariant. We have already constructed a gauge covariant spin dipole $\bar{\mathbf{p}}^{\mathrm{s}} = \bar{\mathbf{\Xi}}^{\mathrm{s}} - \frac{1}{2}\{\bar{\mathcal{R}}, \bar{\mathrm{s}}\}$. In terms of the gauge covariant spin dipole, the spin density can be re-expressed as:

$$S(\mathbf{R},t) = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R})\bar{s} - \nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\mathcal{R}}\bar{s} - \\ -\nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R})\bar{\mathbf{p}}^{s}]. \tag{5.52}$$

It can be formally written the following way:

$$S(\mathbf{R},t) = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}})\bar{s} - \\ -\nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}})\bar{\mathbf{p}}^{s}]. \tag{5.53}$$

The gauge covariant spin dipole $\bar{\mathbf{p}}^{\mathrm{s}}$ is a result of a carrier's center of spin being different from its center of charge.

In the same way the gauge covariant torque dipole has been defined by $\bar{\mathbf{p}}^{\tau} = \bar{\mathbf{\Xi}}^{\tau} - \frac{1}{2}\{\bar{\mathcal{R}}, \bar{\tau}\}$. Carrying out an identical manipulation to that for the spin density, use of the gauge invariant torque dipole allows us to rewrite the torque density formally as:

$$\mathcal{T}(\mathbf{R},t) = \mathrm{tr}\int \frac{d^3q}{(2\pi)^3}\,[\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}})\bar{\tau} - \\ -\nabla_{\mathbf{R}}\cdot\rho(\mathbf{q},\mathbf{R}-\bar{\mathcal{R}})\bar{\mathbf{p}}^{\tau}]. \tag{5.54}$$

A similar formal expression exists for the spin current. Evidently, since these expressions are simply a rewriting of the spin, torque and current densities, the equation of continuity satisfied by these is unaltered.

Finally, if the subspace under consideration contains all the bands in the system, it is easy to show that the gauge covariant dipoles $\bar{\mathbf{p}}^{\mathrm{q}}$, $\bar{\mathbf{p}}^{\mathrm{s}}$ and $\bar{\mathbf{p}}^{\tau}$ vanish. This is because for the spin dipole, for example, by inserting unity in the form $\sum_{\mathrm{m}} |\bar{\mathrm{u}}_{\mathrm{m}}\rangle\langle\bar{\mathrm{u}}_{\mathrm{m}}|$, the term $\bar{\mathbf{\Xi}}^{\mathrm{s}}_{\mathrm{n'n}}$ can be written as $\frac{1}{2}\sum_{\mathrm{m}}(\mathcal{R}_{\mathrm{n'm}}\mathrm{s}_{\mathrm{mn}} + \mathrm{s}_{\mathrm{n'm}}\mathcal{R}_{\mathrm{mn}})$ which exactly cancels the second term in $\bar{\mathbf{p}}^{\mathrm{s}}$. In that case the spin density reduces to:

$$S(\mathbf{R}, t) = \mathrm{tr} \int \frac{d^3q}{(2\pi)^3}\, \rho(\mathbf{q}, \mathbf{R} - \bar{\mathcal{R}})\bar{s}, \tag{5.55}$$

with analogous expressions holding for the torque, spin current and charge current densities. For the purpose of practical calculations, however, when one cannot deal with an infinite eigenspace, the dipole corrections must be taken into account.

We have thus formulated a theory based on the density matrix which is suitable for describing the dynamics of particles in both single and multiple bands. The one band results known from the wave packet formalism, including the terms connected to the Berry phase, emerge from our theory. The formalism can be applied to any clean system regardless of the dimensionality of the Hilbert space under consideration. As a result of the gauge degree of freedom of the basis functions, the position vector which is used as a label of the carrier must be modified by a gauge dependent shift in order to obtain the true particle position. We have shown the way to define macroscopic density distributions for conserved and non-conserved operators. In the case of charge we have derived the equation of continuity and emphasized the dipoles which appear in the current density. In the case of spin we have shown that an additional distribution is present, which takes the form of a torque density, and have also highlighted the correspondence between the multipole terms which appear in the density matrix formalism and those found earlier in the wave packet formalism.

5.5 Appendix

We will present in this appendix some of the proofs and evaluations which require lengthier calculations and would have interrupted the flow if incorporated into the main text. For all derivations except the first, the $|\bar{u}_m\rangle$ and $|u_m\rangle$ are interchangeable.

5.5.1 Proof of Eq. (5.12) for degenerate bands

The perturbed eigenfunctions are given by Eq.(5.12):

$$|\bar{u}_m\rangle = |u_m\rangle - \sum_{n\neq m} \frac{\langle u_n|i\hbar\frac{d}{dt}|u_m\rangle}{\varepsilon_m - \varepsilon_n}|u_n\rangle.$$

For a set of degenerate bands the equilibrium part of the density matrix is proportional to the identity matrix I and will have the form $f_0\,I$, where f_0 is usually the Fermi-Dirac distribution. Transport theory often distinguishes between intrinsic effects, which are due to the equilibrium part of the Boltzmann distribution, and extrinsic effects, which are due to the non-equilibrium correction to it. If we wish to evaluate the expectation value of an operator $\hat{A}$ using the equilibrium part of the density matrix for a set of degenerate bands, which for simplicity here we will take as being two dimensional, the following quantity must be evaluated:

$$\langle\hat{A}\rangle = f_0(\langle\bar{u}_1|\hat{A}|\bar{u}_1\rangle + \langle\bar{u}_2|\hat{A}|\bar{u}_2\rangle)$$

The perturbed wave functions are given by:

$$|\bar{u}_1\rangle = |u_1\rangle - \frac{\langle u_2|i\hbar\frac{d}{dt}|u_1\rangle}{\varepsilon_1 - \varepsilon_2}|u_2\rangle - \sum_{n\neq 1,2} \frac{\langle u_n|i\hbar\frac{d}{dt}|u_1\rangle}{\varepsilon_m - \varepsilon_n}|u_n\rangle$$

$$|\bar{u}_2\rangle = |u_2\rangle - \frac{\langle u_1|i\hbar\frac{d}{dt}|u_2\rangle}{\varepsilon_2 - \varepsilon_1}|u_1\rangle - \sum_{n\neq 1,2} \frac{\langle u_n|i\hbar\frac{d}{dt}|u_m\rangle}{\varepsilon_m - \varepsilon_n}|u_n\rangle.$$

The expectation values are:

$$\langle \bar{u}_1|\hat{A}|\bar{u}_1\rangle = A_{11} - \frac{\langle u_1|i\hbar\frac{du_2}{dt}\rangle}{\varepsilon_1-\varepsilon_2}A_{21} - \frac{\langle u_2|i\hbar\frac{du_1}{dt}\rangle}{\varepsilon_1-\varepsilon_2}A_{12} + \sum_1^{out}$$

$$\langle \bar{u}_2|\hat{A}|\bar{u}_2\rangle = A_{22} - \frac{\langle u_1|i\hbar\frac{du_2}{dt}\rangle}{\varepsilon_2-\varepsilon_1}A_{21} - \frac{\langle u_2|i\hbar\frac{du_1}{dt}\rangle}{\varepsilon_2-\varepsilon_1}A_{12} + \sum_2^{out}.$$

In the above $\sum_1^{\text{out}}$ and $\sum_2^{\text{out}}$ stand for the sums involving the bands outside the degenerate manifold and $\mathrm{A_{ij}} = \langle \mathrm{u_i}|\hat{\mathrm{A}}|\mathrm{u_j}\rangle$. The denominators of the terms involving A_{12} and A_{21} have opposite signs. Therefore, adding up the expectation values we obtain:

$$(\langle \bar{u}_1|\hat{A}|\bar{u}_1\rangle + \langle \bar{u}_2|\hat{A}|\bar{u}_2\rangle) = A_{11} + A_{22} + \sum_1^{out} + \sum_2^{out}.$$

Therefore the terms with diverging denominators cancel out.

5.5.2 Evaluation of position matrix elements

In the general case, the expectation value of the position operator is found as follows:

$$\langle\hat{\mathbf{r}}\rangle = \int\int d^3q d^3r \rho_{nn'} \int \frac{d^3Q}{(2\pi)^3}\times$$
$$\times\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{r}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle$$

The bracket on the second line is:

$$\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{r}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle =$$
$$= \langle \bar{u}_{n'-}|\hat{\mathbf{r}}e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}} =$$
$$= \langle \bar{u}_{n'-}|(-i\frac{\partial}{\partial\mathbf{Q}}e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}})|\bar{u}_{n+}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}}.$$

Taking the differential outside, we find:

$$\begin{aligned}
\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{r}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle = \\
= -i\frac{\partial}{\partial\mathbf{Q}}\langle \bar{u}_{n'-}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}} + \\
+\langle i\frac{\partial\bar{u}_{n'-}}{\partial\mathbf{Q}}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}} + \\
+\langle \bar{u}_{n'-}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|i\frac{\partial\bar{u}_{n+}}{\partial\mathbf{Q}}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}} + \\
+\mathbf{r}\langle \bar{u}_{n'-}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}}.
\end{aligned}$$

All four brackets represent integrals of products of lattice periodic functions and exponentials. They will all eventually be proportional to $\delta(\mathbf{Q})$. Because of this fact the first term above integrates to zero. The partial derivatives in the second and third terms are evaluated by expanding the lattice periodic Bloch wave functions:

$$\begin{aligned}
|\bar{u}_{n+}\rangle &= |\bar{u}_{n\tilde{\mathbf{q}}}\rangle + \frac{\mathbf{Q}}{2}\cdot|\frac{\partial\bar{u}_{n\tilde{\mathbf{q}}}}{\partial\tilde{\mathbf{q}}}\rangle \\
|\bar{u}_{n-}\rangle &= |\bar{u}_{n\tilde{\mathbf{q}}}\rangle - \frac{\mathbf{Q}}{2}\cdot|\frac{\partial\bar{u}_{n\tilde{\mathbf{q}}}}{\partial\tilde{\mathbf{q}}}\rangle
\end{aligned} \tag{5.56}$$

In the limit in which $\mathbf{Q}\to 0$:

$$|\frac{\partial\bar{u}_{n\pm}}{\partial\mathbf{Q}}\rangle = \pm\frac{1}{2}|\frac{\partial\bar{u}_{n\tilde{\mathbf{q}}}}{\partial\tilde{\mathbf{q}}}\rangle.$$

Consequently:

$$\begin{aligned}
\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{r}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle = \\
= \frac{1}{2}(\langle \bar{u}_{n'}|i\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle - \langle i\frac{\partial\bar{u}_{n'}}{\partial\mathbf{q}}|\bar{u}_n\rangle) + \mathbf{r}\delta_{n'n} = \\
= \mathbf{r}\delta_{n'n} + \bar{\mathcal{R}}_{n'n}.
\end{aligned}$$

The expectation value of the position vector yields, finally:

$$\langle \hat{\mathbf{r}} \rangle = \int\int \frac{d^3q d^3r}{(2\pi)^3} \rho_{nn'}(\mathbf{r}\delta_{n'n} + \bar{\mathcal{R}}_{n'n}).$$

5.5.3 Finite variance of carrier position

To see that the variance of the particle position does not contain any divergences, we need to evaluate the expectation value $\langle \hat{\mathbf{r}}^2 \rangle$, or

$$\begin{aligned}
\langle \bar{u}_{n'-} | e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})} \hat{r}^2 e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})} | \bar{u}_{n+} \rangle = \\
= \langle \bar{u}_{n'-} | \hat{r}^2 e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}} | \bar{u}_{n+} \rangle e^{-i\mathbf{Q}\cdot\mathbf{r}} = \\
= -\langle \bar{u}_{n'-} | (\frac{\partial^2}{\partial Q^2} e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}) | \bar{u}_{n+} \rangle e^{-i\mathbf{Q}\cdot\mathbf{r}}.
\end{aligned}$$

Since the carrier position has been shown above to contain no divergences, if $\langle \hat{\mathbf{r}}^2 \rangle$ is not divergent then the variance $\langle \hat{\mathbf{r}}^2 \rangle - \langle \hat{\mathbf{r}} \rangle^2$ is also finite.

First note that an expression of the form $-\mathrm{A}\frac{\partial^2 \mathrm{B}}{\partial \mathrm{Q}^2}\mathrm{CD}$ can be written as:

$$\begin{aligned}
-\frac{\partial^2}{\partial Q^2}(ABCD) + 2\frac{\partial}{\partial \mathbf{Q}} \cdot [\frac{\partial A}{\partial \mathbf{Q}} BCD + AB\frac{\partial}{\partial \mathbf{Q}}(CD)] - \\
-\frac{\partial^2 A}{\partial Q^2} BCD - 2\frac{\partial A}{\partial \mathbf{Q}} \cdot B\frac{\partial}{\partial \mathbf{Q}}(CD) - AB\frac{\partial^2 B}{\partial Q^2}(CD).
\end{aligned}$$

In the case we are considering, all the products involve brackets which are proportional to $\delta(\mathbf{Q})$. As a result, all the terms on the first line vanish under integration with respect to $\mathbf{Q}$ and only the terms on the second line need to be evaluated.

Writing them out explicitly,

$$[-\langle\frac{\partial^2\bar{u}_{n'-}}{\partial Q^2}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle - \langle\bar{u}_{n'-}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\frac{\partial^2\bar{u}_{n+}}{\partial Q^2}\rangle +$$
$$+2i\mathbf{r}\cdot(\langle\frac{\partial\bar{u}_{n'-}}{\partial\mathbf{Q}}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle + \langle\bar{u}_{n'-}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\frac{\partial\bar{u}_{n+}}{\partial\mathbf{Q}}\rangle) -$$
$$-2\langle\frac{\partial\bar{u}_{n'-}}{\partial\mathbf{Q}}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\frac{\partial\bar{u}_{n+}}{\partial\mathbf{Q}}\rangle + r^2\langle\bar{u}_{n'-}|e^{i\mathbf{Q}\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle)]e^{-i\mathbf{Q}\cdot\mathbf{r}}.$$

In order to evaluate the differentials, the wave functions are expanded as in (5.56) except that now the expansion must be made to second order in $\mathbf{Q}$. The final result is:

$$\langle\hat{\mathbf{r}}^2\rangle = \int\int\frac{d^3qd^3r}{(2\pi)^3}\rho_{nn'}[r^2\delta_{n'n} + 2\mathbf{r}\cdot\bar{\mathcal{R}}_{n'n} +$$
$$+\frac{1}{2}\langle\frac{\partial\bar{u}_{n'}}{\partial\tilde{\mathbf{q}}}|\cdot|\frac{\partial\bar{u}_n}{\partial\tilde{\mathbf{q}}}\rangle - \frac{1}{4}(\langle\frac{\partial^2\bar{u}_{n'}}{\partial\tilde{q}^2}|\bar{u}_n\rangle + \langle\bar{u}_{n'}|\frac{\partial^2\bar{u}_n}{\partial\tilde{q}^2}\rangle)].$$

This is clearly finite so the variance of the position operator is finite.

5.5.4 Evaluation of velocity matrix elements

The velocity matrix element:

$$\langle\bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle =$$
$$= \langle\bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot\hat{\mathbf{r}}}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle e^{-i\mathbf{Q}\cdot\mathbf{r}}$$

is easily evaluated by writing out the ket in the second line explicitly:

$$\begin{aligned}
\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot\hat{\mathbf{r}}}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot\hat{\mathbf{r}}}|\bar{u}_{n+}\rangle = \\
= \int d^3r'\,\bar{u}_{n'-}(\mathbf{r}')e^{-i\mathbf{q}_-\cdot\mathbf{r}'}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot\mathbf{r}'}\bar{u}_{n+}(\mathbf{r}') = \\
= \int d^3r'\,e^{i\mathbf{Q}\cdot\mathbf{r}'}\bar{u}_{n'-}(\mathbf{r}')(e^{-i\mathbf{q}_+\cdot\hat{\mathbf{r}}'}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot\hat{\mathbf{r}}'})\bar{u}_{n+}(\mathbf{r}') = \\
= \delta(\mathbf{Q})\int d^3r'\,\bar{u}_{n'}(\mathbf{r}')\tilde{\mathbf{v}}\bar{u}_n(\mathbf{r}') = \\
= \delta(\mathbf{Q})\langle \bar{u}_{n'}|\tilde{\mathbf{v}}|\bar{u}_n\rangle.
\end{aligned}$$

In the above we have abbreviated $\tilde{\mathbf{v}} = e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}'}\hat{\mathbf{v}}e^{i\mathbf{q}\cdot\hat{\mathbf{r}}'}$.

We also need to evaluate the expression $\frac{1}{2}\langle \bar{\mathrm{u}}_{\mathrm{n}'-}|\mathrm{e}^{-\mathrm{i}\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\{\hat{\mathbf{v}},(\hat{\mathbf{r}}-\mathbf{r})\}\mathrm{e}^{\mathrm{i}\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{\mathrm{u}}_{\mathrm{n}+}\rangle$. The matrix element:

$$\begin{aligned}
\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}(\hat{\mathbf{r}}-\mathbf{r})e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle = \\
= \langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}[-i\frac{\partial}{\partial\mathbf{q}_+}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}]|\bar{u}_{n+}\rangle = \\
= -i\frac{\partial}{\partial\mathbf{q}_+}\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle + \\
+\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\frac{\partial\hat{\mathbf{v}}}{\partial\mathbf{q}_+}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle + \\
+\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|i\frac{\partial\bar{u}_{n+}}{\partial\mathbf{q}_+}\rangle.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}(\hat{\mathbf{r}}-\mathbf{r})\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle = \\
= \langle \bar{u}_{n'-}|[i\frac{\partial}{\partial\mathbf{q}_-}e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}]\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle = \\
= i\frac{\partial}{\partial\mathbf{q}_-}\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle + \\
-\langle i\frac{\partial\bar{u}_{n-}}{\partial\mathbf{q}_-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\hat{\mathbf{v}}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n'+}\rangle - \\
-\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\frac{\partial\hat{\mathbf{v}}}{\partial\mathbf{q}_-}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle
\end{aligned}$$

All the brackets in the above two equations are proportional to $\delta(\mathbf{Q})$, causing most of the terms to cancel. We are left with:

$$\frac{1}{2}\langle \bar{u}_{n'-}|e^{-i\mathbf{q}_-\cdot(\hat{\mathbf{r}}-\mathbf{r})}\{\tilde{\mathbf{v}},(\hat{\mathbf{r}}-\mathbf{r})\}e^{i\mathbf{q}_+\cdot(\hat{\mathbf{r}}-\mathbf{r})}|\bar{u}_{n+}\rangle = \\ = \frac{i}{2}\delta(\mathbf{Q})(\langle \bar{u}_n|\tilde{\mathbf{v}}|\frac{\partial \bar{u}_{n'}}{\partial \mathbf{q}}\rangle - \langle \frac{\partial \bar{u}_n}{\partial \mathbf{q}}|\tilde{\mathbf{v}}|\bar{u}_{n'}\rangle).$$

5.5.5 Effect of gauge transformation on $\rho(\mathbf{q},\mathbf{r})$

The operator $\hat{\rho}$ must be invariant under gauge transformations. From the expansion of the density operator in Bloch eigenstates,

$$\hat{\rho} = \sum_{n,n'}\sum_{\mathbf{k},\mathbf{k}'} \rho_{nn'}(\mathbf{k},\mathbf{k}')|\bar{\psi}_{n\tilde{\mathbf{k}}}\rangle\langle\bar{\psi}_{n'\tilde{\mathbf{k}}'}|,$$

it is evident that, if the wave functions change according to $|\bar{\psi}_{\mathrm{m}\tilde{\mathbf{k}}}\rangle \rightarrow \mathrm{O}_{\mathrm{mn}}(\tilde{\mathbf{k}})|\bar{\psi}_{\mathrm{n}\tilde{\mathbf{k}}}\rangle$, then $\rho_{\mathrm{nn}'}(\mathbf{k},\mathbf{k}')$ must transform to $\mathrm{O}^{-1}\rho\mathrm{O}$. Using the definition of $\rho_{\mathrm{nn}'}(\mathbf{k},\mathbf{k}')$ in terms of the Wigner function,

$$\rho(\mathbf{q},\mathbf{r}) = \int \frac{d^3Q}{(2\pi)^3}e^{i\mathbf{Q}\cdot\mathbf{r}}\rho(\mathbf{q}_+,\mathbf{q}_-),$$

and remembering that $\mathbf{k} \equiv \mathbf{q}_+$ and $\mathbf{k}' \equiv \mathbf{q}_-$, we obtain

$$\rho(\mathbf{q},\mathbf{r}) \rightarrow \int \frac{d^3Q}{(2\pi)^3}e^{i\mathbf{Q}\cdot\mathbf{r}}O^{-1}(\tilde{\mathbf{q}}_+)\rho(\mathbf{q}_+,\mathbf{q}_-)O(\tilde{\mathbf{q}}_-).$$

The matrices $\mathrm{O}^{-1}(\tilde{\mathbf{q}}_+)$, $\mathrm{O}(\tilde{\mathbf{q}}_-)$ may be expanded about their arguments as:

$$O^{-1}(\tilde{\mathbf{q}}_+) = O^{-1}(\tilde{\mathbf{q}}) + \frac{\mathbf{Q}}{2}\cdot\frac{\partial O^{-1}(\tilde{\mathbf{q}})}{\partial\tilde{\mathbf{q}}} + O(Q^2) \\ O(\tilde{\mathbf{q}}_-) = O(\tilde{\mathbf{q}}) - \frac{\mathbf{Q}}{2}\cdot\frac{\partial O(\tilde{\mathbf{q}})}{\partial\tilde{\mathbf{q}}} + O(Q^2).$$

Then, to first order in $\mathbf{Q}$, abbreviating $\mathrm{O}(\tilde{\mathbf{q}})$ to O and $\rho(\mathbf{q}_+, \mathbf{q}_-)$ to ρ_q,

$$\begin{aligned} O^{-1}(\tilde{\mathbf{q}}_+)\rho_q(\mathbf{q}_+,\mathbf{q}_-)O(\tilde{\mathbf{q}}_-) = \\ = O^{-1}\rho_q O + \frac{\mathbf{Q}}{2}\cdot(\frac{\partial O^{-1}}{\partial\tilde{\mathbf{q}}}\rho_q O - O^{-1}\rho_q\frac{\partial O}{\partial\tilde{\mathbf{q}}}) = \\ = \tilde{\rho}_q + \frac{\mathbf{Q}}{2}\cdot(\frac{\partial O^{-1}}{\partial\tilde{\mathbf{q}}}O\tilde{\rho}_q - \tilde{\rho}_q O^{-1}\frac{\partial O}{\partial\tilde{\mathbf{q}}}), \end{aligned}$$

where $\tilde{\rho}_\mathrm{q} = \mathrm{O}^{-1}\rho_\mathrm{q}\mathrm{O}$. The last line is obtained by inserting $\mathrm{O}\mathrm{O}^{-1}$ or $\mathrm{O}^{-1}\mathrm{O}$ as appropriate. Since $\frac{\partial}{\partial\tilde{\mathbf{q}}}(\mathrm{O}\mathrm{O}^{-1}) = 0$, we have $\frac{\partial \mathrm{O}^{-1}}{\partial\tilde{\mathbf{q}}}\mathrm{O} = -\mathrm{O}^{-1}\frac{\partial \mathrm{O}}{\partial\tilde{\mathbf{q}}}$ and

$$\rho \to \tilde{\rho} - \frac{\mathbf{Q}}{2}\cdot\{\tilde{\rho}, O^{-1}\frac{\partial O}{\partial\tilde{\mathbf{q}}}\}.$$

Finally, writing $-\frac{\mathbf{Q}}{2} = \frac{\mathrm{i}}{2}\nabla_{\mathbf{r}}\mathrm{e}^{\mathrm{i}\mathbf{Q}\cdot\mathbf{r}}$, we recover formula (5.47) for the transformation of $\rho(\mathbf{q}, \mathbf{r})$.

Bibliography

[1] P. Matl, N. P. Ong, Y. F. Yan, Y. Q. Li, D. Studebaker, T. Baum and G. Doubinina, Phys. Rev. B **57** 10252 (1998).

[2] S. H. Chun, M. B. Salamon and P. D. Han, Phys. Rev. B **59**, 11 155 (1999).

[3] S. Sonoda, S. Shimizu, T. Sasaki, Y. Yamamoto and H. Hori, cond-mat/0108159 (2001).

[4] M. Onoda, N. Nagaosa, cond-mat/0110504 (2001).

[5] R. Karplus and J.M. Luttinger, Phys. Rev. **95**, 1154 (1954).

[6] J. Smit, Physica (Amsterdam) **21**, 877 (1955); ibid. **23**, 39 (1958).

[7] J. M. Luttinger, Phys. Rev. **112** 739 (1958).

[8] L. Berger, Phys. Rev. B **2**, 4559 (1970).

[9] M. Chang and Q. Niu, Phys. Rev B **53**, 7010 (1996).

[10] G. Sundaram and Q. Niu, Phys. Rev B **59**, 14915 (1999).

[11] Q. Niu, D. J. Thouless and Y. Wu, Phys. Rev. B **31**, 6 3372 (1985).

[12] *The Hall effect and its applications,* edited by C. L. Chien and C. R. Westgate (Plenum Press, New York, 1979).

[13] J. Ye, Y. B. Kim, A.J. Millis, B.I. Shraiman, P. Majumdar and Z. Tesanovic, Phys. Rev. Lett. **83**, 18 3737 (1999).

[14] M. V. Berry, Proc. R. Soc. Lond. A **392**, 45 (1984).

[15] T. Jungwirth, Q. Niu, A. Macdonald, Phys. Rev. Lett. **88**, 207208 (2002).

[16] T. Jungwirth, J. Sinova, K. Y. Wang, K. W. Edmonds, R. P. Campion, B. L. Gallagher, C. T. Foxon, Q. Niu, and A. H. MacDonald, Appl. Phys. Lett. **83**, 320 (2003).

[17] Yu. A. Bychkov and E. I. Rashba, JETP Lett. (Engl. transl.) **39**, 66 (1984).

[18] B. Das, S. Datta and R. Reifenberger, Phys. Rev. B **41**, 12 8278 (1990).

[19] S.J. Papadakis, E. P. De Poortere, H. C. Manoharan, M. Shayegan and R. Winkler, Science **283**, 2056 (1999).

[20] G. Lommer, F. Malcher, and U. Rossler , Phys. Rev. Lett. **60**, 8 728 (1988).

[21] G. Lommer, F. Malcher, and U. Rossler , Phys. Rev. B **32**,10 6965 (1985).

[22] X. C. Zhang, A. Pfeuffer-Jeschke, K. Ortner, V. Hock, H. Buhmann, C. R. Becker, and G. Landwehr, Phys. Rev. B **63**, 245305 (2001).

[23] B. Jusserand, D. Richards, G. Allan, C. Priester, and B. Etienne, Phys. Rev. B **51**, 7 4707 (1995).

[24] Y. S. Gui, C. R. Becker, J. Liu, V. Daumer, V. Hock, H. Buhmann, and L. W. Molenkamp, Europhys. Lett. **65** (3), 393 (2004).

[25] R. Lassnig, Phys. Rev. B **31** 8076 (1985).

[26] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer Tracts in Modern Physics Vol. *191*, Springer, Berlin, 2003).

[27] V. F. Radantsev, A. M. Yafyasov, and V. B. Bogevolnov, Semicond. Sci. Technol. **16**, 320 (2001).

[28] T. Matsuyama, R. Kursten, C. Meissner, and U. Merkt , Phys. Rev. B **61**, 23 15 588 (2000).

[29] D. Grundler, Phys. Rev. Lett. **84**, 26 6074 (2000).

[30] V. F. Radantsev, A. M. Yafyasov, V. B. Bogevolnov, I. M. Ivankiv, and O. Yu. Shevchenko, Surf. Sci. **482-485**, 989 (2001).

[31] M. Schultz, F. Heinrichs, U. Merkt, T. Colin, T. Skauli, and S. Lovold, Semicond. Sci. Technol. **11**, 1168 (1996).

[32] Y. Sato, T. Kita, S. Gozu, and S. Yamada , J. Appl. Physics **89**, 12 8017 (2001).

[33] Th. Schapers, G. Engels, J. Lange, Th. Klocke, M. Hollfelder, and H. Luth, J. Appl. Physics **81**, 8 4324 (1998).

[34] G. Engels, J. Lange, Th. Schapers, and H. Luth, Phys. Rev. B **55**, 4 1958 (1997).

[35] L. C. Lew Yan Voo, M. Willatzen, M. Cardona, and N. E. Christensen, Phys. Rev. B **53**, 16 10703 (1996).

[36] R. Knobel, I. P. Smorchkova, and N. Samarth, J. Vac. Sci. Technol. **B17** (3), 1147 (1999).

[37] J. K. Furdyna, J. Appl. Phys. **64** (4), R29 (1988).

[38] G.A. Prinz, Science **282**, 1660 (1998).

[39] S.A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnar, M. L. Roukes, A. Y. Chtchelkanova, and D. M. Treger, Science **294**, 1488 (2001).

[40] G. Schmidt and L.W. Molenkamp, Semicond. Sci. Technol. **17**, 310 (2002).

[41] G. Schmidt, D. Ferrand, L. W. Molenkamp, A. T. Filip, and B. J. van Wees, Phys. Rev. B **62**, R4790(2000).

[42] G. Schmidt, G. Richter, P. Grabs, C. Gould, D. Ferrand, and L. W. Molenkamp, Phys. Rev. Lett. **87**, 227203 (2001).

[43] Y. Ohno, R. Terauchi, T. Adachi, F. Matsukura, and H. Ohno, Phys. Rev. Lett. **83**, 4196 (1999).

[44] Y. Ohno, D. K. Young, B. Beschoten, F. Matsukura, H. Ohno, and D. D. Awschalom, Nature **402**, 790 (1999).

[45] R. Fiederling, M. Keim, G. Reuscher, W. Ossau, G. Schmidt, A. Waag, and L. W. Molenkamp, Nature **402**, 787 (1999).

[46] I. Malajovich, J. M. Kikkawa, D. D. Awschalom, J. J. Berry, and N. Samarth, Phys. Rev. Lett **84**, 1015 (2000).

[47] X. Jiang, R. Wang, S. van Dijken, R. Shelby, R. Macfarlane, G. S. Solomon, J. Harris, and S. S. P. Parkin, Phys. Rev. Lett. **90**, 256603 (2003).

[48] B.T. Jonker, Proceedings IEEE **91**, 727 (2003).

[49] S. Murakami *et al.*, Science **301**, 1348 (2003).

[50] J. Sinova *et al.*, Phys. Rev. Lett. **92**, 126603 (2004).

[51] Florian Schuetz, M. Kollar, and P. Kopietz, Phys. Rev. B **69**, 035313 (2004).

[52] A.M. Bratkovsky and V.V. Osipov, J. Appl. Phys **96** (8), 4525 (2004).

[53] N.M. Chtchelkatchev, JETP. Lett. **78**, 230 (2003).

[54] A. Slobodskyy, C. Gould, T. Slobodskyy, C. R. Becker, G. Schmidt, and L. W. Molenkamp, Phys. Rev. Lett. **90**, 246601 (2003).

[55] M.I. Dyakonov and V. I. Perel, Zh. Eksp. Ter. Fiz. **60**, 1954 (1971) [JETP **33**, 1053 (1971)].

[56] J.E. Hirsch, Phys. Rev. Lett **83**, 1834 (1999).

[57] S. Zhang, Phys. Rev. Lett **85**, 393 (2000). .

[58] E.G. Mishchenko and B.I. Halperin, Phys. Rev. B **68** 045317 (2003).

[59] Y. Qi and S. Zhang, Phys. Rev. B **65**, 214407 (2002), ibid. **67**, 052407 (2003).

[60] V. K. Dugaev, J. Barnas, A. Lusakowski, and L. A. Turski, Phys. Rev. B **65**, 224419 (2002).

[61] J. M. V. Lopes, J. M. B. Lopes dos Santos, and Y. G. Pogorelov , J. Mag. Mag. Mat. **242**, 482-484 (2002).

[62] J. Fabian, S. Das Sarma, Phys. Rev. B **66**, 024436 (2002).

[63] M. I. Dyakonov and V. I. Perel, ZhETF Pis. Red. **13**, 657 (1971) [JETP Lett. **13**, 467 (1971)].

[64] Yugui Yao, Leonard Kleinman, A. H. MacDonald, Jairo Sinova, T. Jungwirth, Ding-sheng Wang, Enge Wang, and Qian Niu, Phys. Rev. Lett. **92**, 037204 (2004).

[65] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 2^{nd} edition (Saunders College Publishing, 1976).

[66] D. Culcer, J. Sinova, N. A. Sinitsyn, T. Jungwirth, A. H. MacDonald, and Q. Niu, Phys. Rev. Lett. **93**, 046602 (2004).

[67] J.D. Jackson, *Classical Electrodynamics*, 3^{rd} edition (John Wiley & Sons, 1999), section 6.6 (pp. 248-258).

[68] We have defined the collision term in analogy with Eq. (2) in [59]. However, we have made some simplifications which are valid for degenerate bands.

[69] S. Murakami, N. Nagaosa, and S.-C. Zhang, Phys. Rev. B **69**, 235206 (2004).

[70] Y. Mitsumori *et al.*, cond-mat/0307268.

[71] A. Fert, J. M. George, H. Jaffres, and G. Faini, J. Phys. D **35**, 2443 (2002).

[72] D.K. Young, J. A. Gupta, E. Johnston-Halperin, R. Epstein, Y. Kato, and D. D. Awschalom, Semicond. Sci. Technol. **17**, 275 (2002).

[73] J.M. Kikkawa and D.D. Awschalom, Phys. Rev. Lett. **80**, 4313 (1998).

[74] H.J. Zhu, M. Ramsteiner, H. Kostial, M. Wassermeier, H.-P. Schnherr, and K. H. Ploog, Phys. Rev. Lett. **87**, 016601 (2001).

[75] H. Ohno, Science **281**, 951 (1998).

[76] T. Damker, H. Bottger, and V. V. Bryksin, Phys. Rev. B **69**, 205327 (2004).

[77] S. Zhang and Z. Yang, Phys. Rev. Lett. **94**, 066602 (2005).

[78] L.Hu, J. Gao, and S. Q. Shen, Phys. Rev. B **70**, 235323 (2004).

[79] E.G.Mishchenko, A.V.Shytov, and B. I. Halperin, Phys. Rev. Lett. **93**, 226602.

[80] J Inoue, G.W.Bauer, L. W. Molenkamp, Phys. Rev. B **67**, 033104 (2003).

[81] Y. K. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Science **306**, 1910 (2004).

[82] J. Wunderlich, B. Kaestner, J. Sinova, and T. Jungwirth, Phys. Rev. Lett. **94**, 047204 (2005).

[83] Y. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Phys. Rev. Lett. **93**, 176601 (2004).

[84] Y. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Nature **427**, 50 (2004).

[85] J. Stephens, J. Berezovsky, J. P. McGuire, L. J. Sham, A. C. Gossard, and D. D. Awschalom, Phys. Rev. Lett. **93**, 097602 (2004).

[86] P.R. Hammar and M. Johnson, Phys. Rev. Lett. **88**, 066806 (2002).

[87] L. S. Levitov, Y. V. Nazarov, and G. M. Eliashberg, Sov. Phys. JETP **61** (1), 133 (1985).

[88] V. M. Edelstein, Solid State Comm. **73**, 233 (1990);.

[89] A.G.Aronov, Yu.B.Lyanda-Geller, and G.E.Pikus, Sov.Phys. JETP **73**, 537 (1991).

[90] L. I. Magarill, A. V. Chaplik, and M. V. Entin, Semiconductors **35** (9), 1081 (2001).

[91] D. Culcer, A. H. MacDonald, and Q. Niu, Phys. Rev. B **68**, 045327 (2003).

[92] T. Koga, J. Nitta, T. Akazaki, and H. Takayanagi, Phys. Rev. Lett. **89**, 046801 (2002).

[93] O.Bleibaum, Phys. Rev. B **69**, 205202, (2004).

[94] M.Q.Weng, M.W.Wu, and L. Jiang, Phys. Rev. B **69**, 245320 (2004).

[95] J. Nitta, T. Akazaki, and H. Takayanagi, Phys. Rev. Lett. **78**, 1335 (1997).

[96] K. L. Kavanagh, M. A. Capano, L. W. Hobbs, J. C. Barbour, P. M. J. Maree, W. Schaff, J. W. Mayer, D. Pettit, J. M. Woddall, J. A. Stroscio, and R. M. Feenstra, J. Appl. Phys. **64** (10), 4843 (1988).

[97] B. Bernevig and S.-C. Zhang, cond-mat/0408442.

[98] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89** (11), Part 1, 5815 (2001).

[99] A. Dudarev, R. B. Diener, I. Carusotto, and Q. Niu, Phys. Rev. Lett. **92** (15), 153005 (2004).

[100] See Chapter 5 of *Semiconductor Laser Fundamentals: Physics of the Gain Materials* by W. W. Chow and S. W. Koch (Springer, Berlin, 1999).

[101] M.Q.Weng and M.W.Wu, J. Appl. Phys. **93**, 410 (2003).

[102] M.Q.Weng, M.W.Wu, and Q. W. Shi, Phys. Rev. B **69**, 125310 (2004).

[103] For a review of spin injection in metals and semiconductors, see M. Johnson, Semicond. Sci. Technol. **17**, 298 (2002).

[104] R. Fiederling *et al.*, Appl. Phys. Lett. **82**, 2160 (2003).

[105] J. Callaway, *Quantum Theory of the Solid State* (Academic Press, New York, 1976).

[106] V. Daumer *et al.*, Appl. Phys. Lett. **83**, 1376 (2003).

[107] R. B. Diener *et al.*, cond-mat/0306184.

[108] R. Shindou and K. Imura, Nucl. Phys. B, **720**, 399 (2005).

[109] I. Malajovich, J. J. Berry, N. Samarth, and D. D. Awschalom, Nature **411**, 770 (2001).

[110] Z. F. Jiang, R. D. Li, S.-C. Zhang, and W. M. Liu, Phys Rev. B **72**, 045201 (2005).

[111] D. Culcer, Y. G. Yao, A. H. MacDonald, and Q. Niu, Phys. Rev. B (2005).

[112] S. Zhang, Phys. Rev. Lett **85**, 393 (2000).

[113] S. Zhang and Z. Yang, Phys. Rev. Lett. **94**, 066602 (2005).

[114] O. V. Dimitrova, Phys. Rev. B **71**, 245327 (2005).

[115] G. Y. Guo, Y. Yao, and Q. Niu, Phys. Rev. Lett. **94**, 226601 (2005).

[116] Y. Yao and Z. Fang, cond-mat/0502351.

[117] D. Culcer, Y. G. Yao, and Q. Niu, Phys. Rev. B (2005).

[118] Y. K. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Science **306**, 1910 (2004).

[119] Y. Kato, R. C. Myers, A. C. Gossard, and D. D. Awschalom, Phys. Rev. Lett. **93**, 176601 (2004).

[120] L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980).

[121] An interesting recent article, G. A. Fiete and E. G. Heller, Phys. Rev. A **68**, 022112 (2003), discusses this topic from a somewhat different point of view.

[122] R. G. Littlejohn and W. G. Flynn, Phys. Rev. A **44**, 5239 (1991).

[123] M. Marder, *Condensed Matter Physics* (John Wiley, New York, 2000).

[124] M.I. Dyakonov and V. I. Perel, JETP Lett. **35**A, 459 (1971).

[125] M.I. Dyakonov and V. I. Perel, Sov. Phys. Solid State **13**, 3023 (1972).

[126] M.I. Dyakonov and V. Y. Kachorovskii, Sov. Phys. Semicond. **20**, 110 (1986).

[127] E. L. Ivchenko, Y. B. Lyanda-Geller, and G. E. Pikus, Sov. Phys. JETP **71**, 550 (1990).

[128] G. E. Pikus and A. N. Titkov in *Optical Orientation* (Elsevier Science Publications, Amsterdam, 1984).

[129] A. Messiah, *Quantum Mechanics* (Dover Publications, New York, 2000).

Vita

Dimitrie Max Culcer was born in the glamorous city of Bucharest, Romania, on December 27th, 1977. Having attended school in Romania and the UK, he enrolled at Oxford University in Fall 1996, pursuing a physics course which led directly to an M. Phys. degree. Outside physics, interests include languages, travel, literature and history, not to mention soccer. He currently lives in Austin with no wives and no children.

Permanent Address: 10 A Delamere Road, Wimbledon,
London SW20 8PS, UK

This dissertation was typeset with LaTeX 2_ε[1] by the author.

[1] LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The macros used in formatting this dissertation were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay, James A. Bednar, and Ayman El-Khashab.