Berry phase effects on electronic properties

Di Xiao

Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

Ming-Che Chang

Department of Physics, National Taiwan Normal University, Taipei 11677, Taiwan

Qian Niu

Department of Physics, The University of Texas at Austin, Austin, Texas 78712, USA

(Published 6 July 2010)

Ever since its discovery the notion of Berry phase has permeated through all branches of physics. Over the past three decades it was gradually realized that the Berry phase of the electronic wave function can have a profound effect on material properties and is responsible for a spectrum of phenomena, such as polarization, orbital magnetism, various (quantum, anomalous, or spin) Hall effects, and quantum charge pumping. This progress is summarized in a pedagogical manner in this review. A brief summary of necessary background is given and a detailed discussion of the Berry phase effect in a variety of solid-state applications. A common thread of the review is the semiclassical formulation of electron dynamics, which is a versatile tool in the study of electron dynamics in the presence of electromagnetic fields and more general perturbations. Finally, a requantization method is demonstrated that converts a semiclassical theory to an effective quantum theory. It is clear that the Berry phase should be added as an essential ingredient to our understanding of basic material properties.

PACS number(s): 71.70.Ej, 72.10.Bg, 73.43.−f, 77.84.−s

CONTENTS

I. INTRODUCTION

A. Topical overview

In 1984, Michael Berry wrote a paper that has generated immense interest throughout the different fields of physics including quantum chemistry (Berry, 1984). This paper is about the adiabatic evolution of an eigenenergy state when the external parameters of a quantum system change slowly and make up a loop in the parameter space. In the absence of degeneracy, the eigenstate will surely come back to itself when finishing the loop, but there will be a phase difference equal to the time integral of the energy (divided by $\hbar$) plus an extra, which is now commonly known as the Berry phase.

The Berry phase has three key properties that make the concept important (Shapere and Wilczek, 1989; Bohm *et al.*, 2003). First, it is gauge invariant. The eigenwave function is defined by a homogeneous linear equation (the eigenvalue equation), so one has the gauge freedom of multiplying it with an overall phase factor which can be parameter dependent. The Berry phase is unchanged (up to integer multiple of 2π) by such a phase factor, provided the eigenwave function is kept to be single valued over the loop. This property makes the Berry phase physical, and the early experimental studies were focused on measuring it directly through interference phenomena.

Second, the Berry phase is geometrical. It can be written as a line integral over the loop in the parameter space and does not depend on the exact rate of change along the loop. This property makes it possible to express the Berry phase in terms of local geometrical quantities in the parameter space. Indeed, Berry himself showed that one can write the Berry phase as an integral of a field, which we now call the Berry curvature, over a surface suspending the loop. A large class of applications of the Berry phase concept occur when the parameters themselves are actually dynamical variables of slow degrees of freedom. The Berry curvature plays an essential role in the effective dynamics of these slow variables. The vast majority of applications considered in this review are of this nature.

Third, the Berry phase has close analogies to gauge field theories and differential geometry (Simon, 1983). This makes the Berry phase a beautiful, intuitive, and powerful unifying concept, especially valuable in today's ever specializing physical science. In primitive terms, the Berry phase is like the Aharonov-Bohm phase of a charged particle traversing a loop including a magnetic flux, while the Berry curvature is like the magnetic field. The integral of the Berry curvature over closed surfaces, such as a sphere or torus, is known to be topological and quantized as integers (Chern numbers). This is analogous to the Dirac monopoles of magnetic charges that must be quantized in order to have a consistent quantum-mechanical theory for charged particles in magnetic fields. Interestingly, while the magnetic monopoles are yet to be detected in the real world, the topological Chern numbers have already found correspondence with the quantized Hall conductance plateaus in the spectacular quantum Hall phenomenon (Thouless *et al.*, 1982).

This review concerns applications of the Berry phase concept in solid-state physics. In this field, we are typically interested in macroscopic phenomena which is slow in time and smooth in space in comparison with the atomic scales. Not surprisingly, the vast majority of applications of the Berry phase concept are discussed here. This field of physics is also extremely diverse, and we have many layers of theoretical frameworks with different degrees of transparency and validity (Ashcroft and Mermin, 1976; Marder, 2000). Therefore, a unifying organizing principle such as the Berry phase concept is particularly valuable.

We focus our attention on electronic properties, which play a dominant role in various aspects of material properties. The electrons are the glue of materials and they are also the agents responding swiftly to external fields and giving rise to strong and useful signals. A basic paradigm of the theoretical framework is based on the assumption that electrons are in Bloch waves traveling more or less independently in periodic potentials of the lattice, except that the Pauli exclusion principle has to be satisfied and electron-electron interactions are taken care of in some self-consistent manner. Much of our intuition on electron transport is derived from the semiclassical picture that electrons behave almost as free particles in response to external fields provided one uses the band energy in place of the free-particle dispersion. Because of this, first-principles studies of electronic properties tend to document only the energy band structures and various density profiles.

There has been overwhelming evidence that such a simple picture cannot give complete account of effects to first order in electromagnetic fields. A prime example is the anomalous velocity, a correction to the usual quasiparticle group velocity from the band energy dispersion. This correction can be understood from a linear response analysis: the velocity operator has off-diagonal elements and electric field mixes the bands so that the expectation value of the velocity acquires an additional term to first order in the field (Karplus and Luttinger, 1954; Kohn and Luttinger, 1957). The anomalous velocity can also be understood as due to the Berry curvature of the Bloch states, which exist in the absence of the external fields and manifest in the quasiparticle velocity when the crystal momentum is moved by external forces

(Chang and Niu, 1995, 1996; Sundaram and Niu, 1999). This understanding enabled us to make a direct connection with the topological Chern number formulation of the quantum Hall effect (Thouless *et al.*, 1982; Kohmoto, 1985), providing motivation as well as confidence in our pursuit of the eventually successful intrinsic explanation of the anomalous Hall effect (Jungwirth *et al.*, 2002; Nagaosa *et al.*, 2010).

Interestingly enough, the traditional view cannot even explain some basic effects to zeroth order of the fields. The two basic electromagnetic properties of solids as a medium are the electric polarization and magnetization, which can exist in the absence of electric and magnetic fields in ferroelectric and ferromagnetic materials. Their classical definitions were based on the picture of bound charges and currents, but these are clearly inadequate for the electronic contribution and it was known that the polarization and orbital magnetization cannot be determined from the charge and current densities in the bulk of a crystal at all. A breakthrough on electric polarization was made in the early 1990s by linking it with the phenomenon of adiabatic charge transport and expressing it in terms of the Berry phase[1] across the entire Brillouin zone (Resta, 1992; King-Smith and Vanderbilt, 1993). Based on the Berry phase formula, one can now routinely calculate polarization related properties using first-principles methods, with a typical precision of the density functional theory. The breakthrough on orbital magnetization came only recently, showing that it not only consists of the orbital moments of quasiparticles but also contains a Berry curvature contribution of topological origin (Thonhauser *et al.*, 2005; Xiao *et al.*, 2005; Shi *et al.*, 2007).

In this article, we follow the traditional semiclassical formalism of quasiparticle dynamics, only to make it more rigorous by including the Berry curvatures in the various facets of the phase space including the adiabatic time parameter. All of the above-mentioned effects are transparently revealed with complete precision of the full quantum theory. A number of new and related effects, such as anomalous thermoelectric, valley Hall, and magnetotransport, are easily predicted, and other effects due to crystal deformation and order parameter inhomogeneity can also be explored without difficulty. Moreover, by including Berry phase induced anomalous transport between collisions and "side jumps" during collisions (which is itself a kind of Berry phase effect), the semiclassical Boltzmann transport theory can give complete account of linear response phenomena in weakly disordered systems (Sinitsyn, 2008). On a microscopic level, although the electron wave-packet dynamics is yet to be directly observed in solids, the formalism has been replicated for light transport in photonic crystals, where the associated Berry phase effects are vividly exhibited in experiments (Bliokh *et al.*, 2008). Finally, it is possible to generalize the semiclassical dynamics in a single band into one with degenerate or nearly degenerate bands (Culcer *et al.*, 2005; Shindou and Imura, 2005), and one can study transport phenomena where interband coherence effects as in spin transport, only to realize that the Berry curvatures and quasiparticle magnetic moments become non-Abelian (i.e., matrices).

The semiclassical formalism is certainly amendable to include quantization effects. For integrable dynamics, such as Bloch oscillations and cyclotron orbits, one can use the Bohr-Sommerfeld or Einstein-Brillouin-Keller quantization rule. The Berry phase enters naturally as a shift to the classical action, affecting the energies of the quantized levels, e.g., the Wannier-Stark ladders and the Landau levels. A high point of this kind of application is the explanation of the intricate fractal-like Hofstadter spectrum (Chang and Niu, 1995, 1996). A recent breakthrough has also enabled us to find the density of quantum states in the phase space for the general case (including nonintegrable systems) (Xiao *et al.*, 2005), revealing Berry curvature corrections which should have broad impacts on calculations of equilibrium as well as transport properties. Finally, one can execute a generalized Peierls substitution relating the physical variables to the canonical variables, turning the semiclassical dynamics into a full quantum theory valid to first order in the fields (Chang and Niu, 2008). Spin-orbit coupling and anomalous corrections to the velocity and magnetic moment are all found from a simple explanation of this generalized Peierls substitution.

Therefore, it is clear that Berry phase effects in solid-state physics are not something just nice to be found here and there; the concept is essential for a coherent understanding of all the basic phenomena. It is the purpose of this review to summarize a theoretical framework which continues the traditional semiclassical point of view but with a much broader range of validity. It is necessary and sufficient to include the Berry phases and gradient energy corrections in addition to the energy dispersions in order to account for all phenomena up to first order in the fields.

B. Organization of the review

This review can be divided into three main parts. In Sec. II we consider the simplest example of Berry phase in crystals: the adiabatic transport in a band insulator. In particular, we show that induced adiabatic current due to a time-dependent perturbation can be written as a Berry phase of the electronic wave functions. Based on this understanding, the modern theory of electric polarization is reviewed. In Sec. III the electron dynamics in the presence of an electric field is discussed as a specific example of the time-dependent problem, and the basic formula developed in Sec. II can be directly applied. In this case, the Berry phase is made manifest as transverse velocity of the electrons, which gives rise to a Hall current. We then apply this formula to study the quantum, anomalous, and valley Hall effect.

---

[1]Also called Zak's phase, it is independent of the Berry curvature which only characterizes Berry phases over loops continuously shrinkable to zero (Zak, 1989).

To study the electron dynamics under spatial-dependent perturbations, we turn to the semiclassical formalism of Bloch electron dynamics, which has proven to be a powerful tool to investigate the influence of slowly varying perturbations on the electron dynamics. In Sec. IV we discuss the construction of the electron wave packet and show that the wave packet carries an orbital magnetic moment. Two applications of the wave-packet approach, the orbital magnetization and anomalous thermoelectric transport in ferromagnet, are discussed. In Sec. V the wave-packet dynamics in the presence of electromagnetic fields is studied. We show that the Berry phase not only affects the equations of motion but also modifies the electron density of states in the phase space, which can be changed by applying a magnetic field. The formula for orbital magnetization is rederived using the modified density of states. We also present a comprehensive study of the magnetotransport in the presence of the Berry phase. The electron dynamics under more general perturbations is discussed in Sec. VI. We also present two applications: electron dynamics in deformed crystals and polarization induced by inhomogeneity.

In the remaining part of the review we focus on the requantization of the semiclassical formulation. In Sec. VII the Bohr-Sommerfeld quantization is reviewed in detail. With its help, one can incorporate the Berry phase effect into the Wannier-Stark ladders and the Landau levels very easily. In Sec. VIII we show that the same semiclassical approach can be applied to systems subject to a very strong magnetic field. One only has to separate the field into a quantization part and a perturbation. The former should be treated quantum mechanically to obtain the magnetic Bloch band spectrum while the latter is treated perturbatively. Using this formalism, the cyclotron motion, the splitting into magnetic subbands, and the quantum Hall effect are discussed. In Sec. IX we review recent advances on the non-Abelian Berry phase in degenerate bands. The Berry connection now plays an explicit role in spin dynamics and is deeply related to the spin-orbit interaction. The relativistic Dirac electrons and the Kane model in semiconductors are cited as two applied examples. Finally, we discuss the requantization of the semiclassical theory and the hierarchy of effective quantum theories.

We do not attempt to cover all of the Berry phase effects in this review. Interested readers can consult the following: Shapere and Wilczek (1989); Nenciu (1991); Resta (1994, 2000); Thouless (1998); Bohm et al. (2003); Teufel (2003); Chang and Niu (2008). In this review, we focus on a pedagogical and self-contained approach, with the main focus on the semiclassical formalism of Bloch electron dynamics (Chang and Niu, 1995, 1996; Sundaram and Niu, 1999). We start with the simplest case, then gradually expand the formalism as more complicated physical situations are considered. Whenever a new ingredient is added, a few applications are provided to demonstrate the basic ideas. The vast number of applications we discuss is a reflection of the universality of the Berry phase effect.

C. Basic concepts of the Berry phase

In this section we introduce the basic concepts of the Berry phase. Following Berry's original paper (Berry, 1984), we first discuss how the Berry phase arises during the adiabatic evolution of a quantum state. We then introduce the local description of the Berry phase in terms of the Berry curvature. A two-level model is used to demonstrate these concepts as well as some important properties, such as the quantization of the Berry phase. Our aim is to provide a minimal but self-contained introduction. For a detailed account of the Berry phase, including its mathematical foundation and applications in a wide range of fields in physics, see Shapere and Wilczek (1989) and Bohm et al. (2003), and references therein.

1. Cyclic adiabatic evolution

Consider a physical system described by a Hamiltonian that depends on time through a set of parameters, denoted by $\boldsymbol{R} = (R_1, R_2, \ldots)$, i.e.,

$$H = H(\boldsymbol{R}), \quad \boldsymbol{R} = \boldsymbol{R}(t). \tag{1.1}$$

We are interested in the adiabatic evolution of the system as $\boldsymbol{R}(t)$ moves slowly along a path $\mathcal{C}$ in the parameter space. For this purpose, it is useful to introduce an instantaneous orthonormal basis from the eigenstates of $H(\boldsymbol{R})$ at each value of the parameter $\boldsymbol{R}$, i.e.,

$$H(\boldsymbol{R})|n(\boldsymbol{R})\rangle = \varepsilon_n(\boldsymbol{R})|n(\boldsymbol{R})\rangle. \tag{1.2}$$

However, Eq. (1.2) alone does not completely determine the basis function $|n(\boldsymbol{R})\rangle$; it still allows an arbitrary $\boldsymbol{R}$-dependent phase factor of $|n(\boldsymbol{R})\rangle$. One can make a phase choice, also known as a gauge, to remove this arbitrariness. Here we require that the phase of the basis function is smooth and single valued along the path $\mathcal{C}$ in the parameter space.[2]

According to the quantum adiabatic theorem (Kato, 1950; Messiah, 1962), a system initially in one of its eigenstates $|n(\boldsymbol{R}(0))\rangle$ will stay as an instantaneous eigenstate of the Hamiltonian $H(\boldsymbol{R}(t))$ throughout the process. (A derivation can be found in the Appendix.) Therefore the only degree of freedom we have is the phase of the quantum state. We write the state at time t as

---

[2]Strictly speaking, such a phase choice is guaranteed only in finite neighborhoods of the parameter space. In the general case, one can proceed by dividing the path into several such neighborhoods overlapping with each other, then use the fact that in the overlapping region the wave functions are related by a gauge transformation of form (1.7).

$$|\psi_n(t)\rangle = e^{i\gamma_n(t)}\exp\left[-\frac{i}{\hbar}\int_0^t dt'\,\varepsilon_n(\boldsymbol{R}(t'))\right]|n(\boldsymbol{R}(t))\rangle,$$

$$(1.3)$$

where the second exponential is known as the dynamical phase factor. Inserting Eq. (1.3) into the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi_n(t)\rangle = H(\boldsymbol{R}(t))|\psi_n(t)\rangle \qquad (1.4)$$

and multiplying it from the left by $\langle n(\boldsymbol{R}(t))|$, one finds that γ_n can be expressed as a path integral in the parameter space

$$\gamma_n = \int_{\mathcal{C}} d\boldsymbol{R}\cdot\mathcal{A}_n(\boldsymbol{R}), \qquad (1.5)$$

where $\mathcal{A}_n(\boldsymbol{R})$ is a vector-valued function

$$\mathcal{A}_n(\boldsymbol{R}) = i\langle n(\boldsymbol{R})|\frac{\partial}{\partial\boldsymbol{R}}|n(\boldsymbol{R})\rangle. \qquad (1.6)$$

This vector $\mathcal{A}_n(\boldsymbol{R})$ is called the Berry connection or the Berry vector potential. Equation (1.5) shows that, in addition to the dynamical phase, the quantum state will acquire an additional phase γ_n during the adiabatic evolution.

Obviously, $\mathcal{A}_n(\boldsymbol{R})$ is gauge dependent. If we make a gauge transformation

$$|n(\boldsymbol{R})\rangle \rightarrow e^{i\zeta(\boldsymbol{R})}|n(\boldsymbol{R})\rangle, \qquad (1.7)$$

with $\zeta(\boldsymbol{R})$ an arbitrary smooth function and $\mathcal{A}_n(\boldsymbol{R})$ transforms according to

$$\mathcal{A}_n(\boldsymbol{R}) \rightarrow \mathcal{A}_n(\boldsymbol{R}) - \frac{\partial}{\partial\boldsymbol{R}}\zeta(\boldsymbol{R}). \qquad (1.8)$$

Consequently, the phase γ_n given by Eq. (1.5) will be changed by $\zeta(\boldsymbol{R}(0)) - \zeta(\boldsymbol{R}(T))$ after the transformation, where $\boldsymbol{R}(0)$ and $\boldsymbol{R}(T)$ are the initial and final points of the path $\mathcal{C}$. This observation has led Fock (1928) to conclude that one can always choose a suitable $\zeta(\boldsymbol{R})$ such that γ_n accumulated along the path $\mathcal{C}$ is canceled out, leaving Eq. (1.3) with only the dynamical phase. Because of this, the phase γ_n has long been deemed unimportant and it was usually neglected in the theoretical treatment of time-dependent problems.

This conclusion remained unchallenged until Berry (1984) reconsidered the cyclic evolution of the system along a *closed* path $\mathcal{C}$ with $\boldsymbol{R}(T) = \boldsymbol{R}(0)$. The phase choice we made earlier on the basis function $|n(\boldsymbol{R})\rangle$ requires $e^{i\zeta(\boldsymbol{R})}$ in the gauge transformation [Eq. (1.7)] to be single valued, which implies

$$\zeta(\boldsymbol{R}(0)) - \zeta(\boldsymbol{R}(T)) = 2\pi\times\text{integer}. \qquad (1.9)$$

This shows that γ_n can be only changed by an integer multiple of 2π under the gauge transformation [Eq. (1.7)] and it cannot be removed. Therefore for a closed path, γ_n becomes a gauge-invariant physical quantity,

now known as the Berry phase or geometric phase in general; it is given by

$$\gamma_n = \oint_{\mathcal{C}} d\boldsymbol{R}\cdot\mathcal{A}_n(\boldsymbol{R}). \qquad (1.10)$$

From the above definition, we can see that the Berry phase only depends on the geometric aspect of the closed path and is independent of how $\boldsymbol{R}(t)$ varies in time. The explicit time dependence is thus not essential in the description of the Berry phase and will be dropped in the following discussion.

2. Berry curvature

It is useful to define, in analogy to electrodynamics, a gauge-field tensor derived from the Berry vector potential:

$$\Omega_{\mu\nu}^n(\boldsymbol{R}) = \frac{\partial}{\partial R^\mu}\mathcal{A}_\nu^n(\boldsymbol{R}) - \frac{\partial}{\partial R_\nu}\mathcal{A}_\mu^n(\boldsymbol{R})$$

$$= i\left[\left\langle\frac{\partial n(\boldsymbol{R})}{\partial R^\mu}\middle|\frac{\partial n(\boldsymbol{R})}{\partial R^\nu}\right\rangle - (\nu\leftrightarrow\mu)\right]. \qquad (1.11)$$

This field is called the Berry curvature. Then according to Stokes's theorem the Berry phase can be written as a surface integral

$$\gamma_n = \int_{\mathcal{S}} dR^\mu\wedge dR^\nu\tfrac{1}{2}\Omega_{\mu\nu}^n(\boldsymbol{R}), \qquad (1.12)$$

where $\mathcal{S}$ is an arbitrary surface enclosed by the path $\mathcal{C}$. It can be verified from Eq. (1.11) that, unlike the Berry vector potential, the Berry curvature is gauge invariant and thus observable.

If the parameter space is three dimensional, Eqs. (1.11) and (1.12) can be recast into a vector form

$$\boldsymbol{\Omega}_n(\boldsymbol{R}) = \boldsymbol{\nabla}_{\boldsymbol{R}}\times\mathcal{A}_n(\boldsymbol{R}), \qquad (1.11')$$

$$\gamma_n = \int_{\mathcal{S}} d\boldsymbol{S}\cdot\boldsymbol{\Omega}_n(\boldsymbol{R}). \qquad (1.12')$$

The Berry curvature tensor $\Omega_{\mu\nu}^n$ and vector $\boldsymbol{\Omega}_n$ are related by $\Omega_{\mu\nu}^n = \epsilon_{\mu\nu\xi}(\boldsymbol{\Omega}_n)_\xi$ with $\epsilon_{\mu\nu\xi}$ the Levi-Cività antisymmetry tensor. The vector form gives us an intuitive picture of the Berry curvature: it can be viewed as the magnetic field in the parameter space.

Besides the differential formula given in Eq. (1.11), the Berry curvature can be also written as a summation over the eigenstates:

$$\Omega_{\mu\nu}^n(\boldsymbol{R}) = i\sum_{n'\neq n}\frac{\langle n|\partial H/\partial R^\mu|n'\rangle\langle n'|\partial H/\partial R^\nu|n\rangle - (\nu\leftrightarrow\mu)}{(\varepsilon_n - \varepsilon_{n'})^2}.$$

$$(1.13)$$

The curvature can be obtained from Eq. (1.11) using $\langle n|\partial H/\partial\boldsymbol{R}|n'\rangle = \langle\partial n/\partial\boldsymbol{R}|n'\rangle(\varepsilon_n - \varepsilon_{n'})$ for $n'\neq n$. The summation formula has the advantage that no differentiation on the wave function is involved, therefore it can be evaluated under any gauge choice. This property is par-

ticularly useful for numerical calculations, in which the condition of a smooth phase choice of the eigenstates is not guaranteed in standard diagonalization algorithms. It has been used to evaluate the Berry curvature in crystals with the eigenfunctions supplied from first-principles calculations (Fang *et al.*, 2003; Yao *et al.*, 2004).

Equation (1.13) offers further insight on the origin of the Berry curvature. The adiabatic approximation adopted earlier is essentially a projection operation, i.e., the dynamics of the system is restricted to the nth energy level. In view of Eq. (1.13), the Berry curvature can be regarded as the result of the "residual" interaction of those projected-out energy levels. In fact, if all energy levels are included, it follows from Eq. (1.13) that the total Berry curvature vanishes for each value of $\boldsymbol{R}$,

$$\sum_n \Omega_{\mu\nu}^n(\boldsymbol{R}) = 0. \tag{1.14}$$

This is the local conservation law for the Berry curvature.[3] Equation (1.13) also shows that $\Omega_{\mu\nu}^n(\boldsymbol{R})$ becomes singular if two energy levels $\varepsilon_n(\boldsymbol{R})$ and $\varepsilon_{n'}(\boldsymbol{R})$ are brought together at certain value of $\boldsymbol{R}$. This degeneracy point corresponds to a monopole in the parameter space; an explicit example is given below.

So far we have discussed the situation where a single energy level can be separated out in the adiabatic evolution. However, if the energy levels are degenerate, then the dynamics must be projected to a subspace spanned by those degenerate eigenstates. Wilczek and Zee (1984) showed that in this situation non-Abelian Berry curvature naturally arises. Culcer *et al.* (2005) and Shindou and Imura (2005) discussed the non-Abelian Berry curvature in the context of degenerate Bloch bands. In the following we focus on the Abelian formulation and defer the discussion of the non-Abelian Berry curvature to Sec. IX.

Compared to the Berry phase which is always associated with a closed path, the Berry curvature is truly a *local* quantity. It provides a local description of the geometric properties of the parameter space. Moreover, so far we have treated the adiabatic parameters as passive quantities in the adiabatic process, i.e., their time evolution is given from the outset. Later we will show that the parameters can be regarded as dynamical variables and the Berry curvature will directly participate in the dynamics of the adiabatic parameters (Kuratsuji and Iida, 1985). In this sense, the Berry curvature is a more fundamental quantity than the Berry phase.

3. Example: The two-level system

Consider a concrete example: a two-level system. The purpose to study this system is twofold. First, as a simple model, it demonstrates the basic concepts as well as several important properties of the Berry phase. Second, it will be repeatedly used later in this article, although in different physical context. It is therefore useful to go through the basis of this model.

The generic Hamiltonian of a two-level system takes the following form:

$$H = \boldsymbol{h}(\boldsymbol{R}) \cdot \boldsymbol{\sigma}, \tag{1.15}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices. Despite its simple form, the above Hamiltonian describes a number of physical systems in condensed-matter physics for which the Berry phase effect has been discussed. Examples include spin-orbit coupled systems (Culcer *et al.*, 2003; Liu *et al.*, 2008), linearly conjugated diatomic polymers (Su *et al.*, 1979; Rice and Mele, 1982), one-dimensional ferroelectrics (Vanderbilt and King-Smith, 1993; Onoda *et al.*, 2004b), graphene (Semenoff, 1984; Haldane, 1988), and Bogoliubov quasiparticles (Zhang *et al.*, 2006).

Parametrize $\boldsymbol{h}$ by its polar angle θ and azimuthal angle ϕ, $\boldsymbol{h} = h(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, the two eigenstates, with energies $\pm h$, are

$$|u_-\rangle = \begin{pmatrix} \sin\frac{\theta}{2}e^{-i\phi} \\ -\cos\frac{\theta}{2} \end{pmatrix}, \quad |u_+\rangle = \begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2} \end{pmatrix}. \tag{1.16}$$

We are, of course, free to add an arbitrary phase to these wave functions. Consider the lower energy level. The Berry connection is given by

$$\mathcal{A}_\theta = \langle u|i\partial_\theta u\rangle = 0, \tag{1.17a}$$

$$\mathcal{A}_\phi = \langle u|i\partial_\phi u\rangle = \sin^2\frac{\theta}{2}, \tag{1.17b}$$

and the Berry curvature is

$$\Omega_{\theta\phi} = \partial_\theta \mathcal{A}_\phi - \partial_\phi \mathcal{A}_\theta = \tfrac{1}{2}\sin\theta. \tag{1.18}$$

However, the phase of $|u_-\rangle$ is not defined at the south pole ($\theta = \pi$). We can choose another gauge by multiplying $|u_-\rangle$ by $e^{i\phi}$ so that the wave function is smooth and single valued everywhere except at the north pole. Under this gauge we find $\mathcal{A}_\theta = 0$ and $\mathcal{A}_\phi = -\cos^2\theta/2$, and the same expression for the Berry curvature as in Eq. (1.18). This is not surprising because the Berry curvature is a gauge-independent quantity and the Berry connection is not.[4]

---

[3]The conservation law is obtained under the condition that the full Hamiltonian is known. However, in practice one usually deals with effective Hamiltonians which are obtained through some projection process of the full Hamiltonian. Therefore there will always be some "residual" Berry curvature accompanying the effective Hamiltonian [see Chang and Niu (2008) and discussions in Sec. IX].

[4]One can verify that $|u\rangle = (\sin(\theta/2)e^{-i\beta\phi}, -\cos(\theta/2)e^{i(\beta-1)\phi})^T$ is also an eigenstate. The phase accumulated by such a state along the loop defined by $\theta = \pi/2$ is $\Gamma = 2\pi(\beta - \frac{1}{2})$, which seems to imply that the Berry phase is gauge dependent. This is because for an arbitrary β the basis function $|u\rangle$ is not single valued; one must also trace the phase change in the basis function. For integer value of β the function $|u\rangle$ is single valued along the loop and the Berry phase is well defined up to an integer multiple of 2π.

If $\boldsymbol{h}(\boldsymbol{R})$ depends on a set of parameters $\boldsymbol{R}$, then

$$\Omega_{R_1 R_2} = \frac{1}{2} \frac{\partial(\phi, \cos \theta)}{\partial(R_1, R_2)}. \tag{1.19}$$

Several important properties of the Berry curvature can be revealed by considering the specific case of $\boldsymbol{h} = (x, y, z)$. Using Eq. (1.19), we find the Berry curvature in its vector form

$$\boldsymbol{\Omega} = \frac{1}{2} \frac{\boldsymbol{h}}{h^3}. \tag{1.20}$$

One recognizes that Eq. (1.20) is the field generated by a monopole at the origin $\boldsymbol{h} = 0$ (Dirac, 1931; Wu and Yang, 1975; Sakurai, 1993), where the two energy levels become degenerate. Therefore the degeneracy points act as sources and drains of the Berry curvature flux. Integrate the Berry curvature over a sphere containing the monopole, which is the Berry phase on the sphere; we find

$$\frac{1}{2\pi} \int_{S^2} d\theta d\phi \, \Omega_{\theta\phi} = 1. \tag{1.21}$$

In general, the Berry curvature integrated over a closed manifold is quantized in the units of 2π and equals to the net number of monopoles inside. This number is called the Chern number and is responsible for a number of quantization effects discussed below.

D. Berry phase in Bloch bands

Above we introduced the basic concepts of the Berry phase for a generic system described by a parameter-dependent Hamiltonian. We now consider its realization in crystalline solids. As we shall see, the band structure of crystals provides a natural platform to investigate the occurrence of the Berry phase effect.

Within the independent electron approximation, the band structure of a crystal is determined by the following Hamiltonian for a single electron:

$$H = \frac{\hat{p}^2}{2m} + V(\boldsymbol{r}), \tag{1.22}$$

where $V(\boldsymbol{r} + \boldsymbol{a}) = V(\boldsymbol{r})$ is the periodic potential with $\boldsymbol{a}$ the Bravais lattice vector. According to Bloch's theorem, the eigenstates of a periodic Hamiltonian satisfy the following boundary condition:[5]

$$\psi_{n\boldsymbol{q}}(\boldsymbol{r} + \boldsymbol{a}) = e^{i\boldsymbol{q}\cdot\boldsymbol{a}} \psi_{n\boldsymbol{q}}(\boldsymbol{r}), \tag{1.23}$$

where n is the band index and $\hbar\boldsymbol{q}$ is the crystal momentum, which resides in the Brillouin zone. Thus the system is described by a $\boldsymbol{q}$-independent Hamiltonian with a $\boldsymbol{q}$-dependent boundary condition [Eq. (1.23)]. To comply with the general formalism of the Berry phase, we make

_____

[5]Through out this article, $\boldsymbol{q}$ refers to the canonical momentum and $\boldsymbol{k}$ is reserved for mechanical momentum.

the following unitary transformation to obtain a $\boldsymbol{q}$-dependent Hamiltonian:

$$H(\boldsymbol{q}) = e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} H e^{i\boldsymbol{q}\cdot\boldsymbol{r}} = \frac{(\hat{\boldsymbol{p}} + \hbar\boldsymbol{q})^2}{2m} + V(\boldsymbol{r}). \tag{1.24}$$

The transformed eigenstate $u_{n\boldsymbol{q}}(\boldsymbol{r}) = e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \psi_{n\boldsymbol{q}}(\boldsymbol{r})$ is just the cell-periodic part of the Bloch function. It satisfies the strict periodic boundary condition

$$u_{n\boldsymbol{q}}(\boldsymbol{r} + \boldsymbol{a}) = u_{n\boldsymbol{q}}(\boldsymbol{r}). \tag{1.25}$$

This boundary condition ensures that all the eigenstates live in the same Hilbert space. We can thus identify the Brillouin zone as the parameter space of the transformed Hamiltonian $H(\boldsymbol{q})$ and $|u_n(\boldsymbol{q})\rangle$ as the basis function.

Since the $\boldsymbol{q}$ dependence of the basis function is inherent to the Bloch problem, various Berry phase effects are expected in crystals. For example, if $\boldsymbol{q}$ is forced to vary in the momentum space, then the Bloch state will pick up a Berry phase:

$$\gamma_n = \oint_{\mathcal{C}} d\boldsymbol{q} \cdot \langle u_n(\boldsymbol{q}) | i\boldsymbol{\nabla}_{\boldsymbol{q}} | u_n(\boldsymbol{q}) \rangle. \tag{1.26}$$

We emphasize that the path $\mathcal{C}$ must be closed to make γ_n a gauge-invariant quantity with physical significance.

Generally speaking, there are two ways to generate a closed path in the momentum space. One can apply a magnetic field, which induces a cyclotron motion along a closed orbit in the $\boldsymbol{q}$ space. This way the Berry phase can manifest in various magneto-oscillatory effects (Mikitik and Sharlai, 1999, 2004, 2007), which have been observed in metallic compound LaRhIn$_5$ (Goodrich et al., 2002), and most recently graphene systems (Novoselov et al., 2005, 2006; Zhang et al., 2005). Such a closed orbit is possible only in two- or three-dimensional (3D) systems (see Sec. VII.A). Following our discussion in Sec. I.C, we define the Berry curvature of the energy bands by

$$\boldsymbol{\Omega}_n(\boldsymbol{q}) = \boldsymbol{\nabla}_{\boldsymbol{q}} \times \langle u_n(\boldsymbol{q}) | i\boldsymbol{\nabla}_{\boldsymbol{q}} | u_n(\boldsymbol{q}) \rangle. \tag{1.27}$$

The Berry curvature $\boldsymbol{\Omega}_n(\boldsymbol{q})$ is an intrinsic property of the band structure because it only depends on the wave function. It is nonzero in a wide range of materials, in particular, crystals with broken time-reversal or inversion symmetry. In fact, once we have introduced the concept of the Berry curvature, a closed loop is not necessary because the Berry curvature itself is a local gauge-invariant quantity. It is now well recognized that information on the Berry curvature is essential in a proper description of the dynamics of Bloch electrons, which has various effects on transport and thermodynamic properties of crystals.

One can also apply an electric field to cause a linear variation in $\boldsymbol{q}$. In this case, a closed path is realized when $\boldsymbol{q}$ sweeps the entire Brillouin zone. To see this, we note that the Brillouin zone has the topology of a torus: the two points $\boldsymbol{q}$ and $\boldsymbol{q} + \boldsymbol{G}$ can be identified as the same point, where $\boldsymbol{G}$ is the reciprocal lattice vector. This can

be seen by noting that $|\psi_n(\boldsymbol{q})\rangle$ and $|\psi_n(\boldsymbol{q}+\boldsymbol{G})\rangle$ satisfy the same boundary condition in Eq. (1.23); therefore, they can at most differ by a phase factor. The torus topology is realized by making the phase choice $|\psi_n(\boldsymbol{q})\rangle = |\psi_n(\boldsymbol{q}+\boldsymbol{G})\rangle$. Consequently, $|u_n(\boldsymbol{q})\rangle$ and $|u_n(\boldsymbol{q}+\boldsymbol{G})\rangle$ satisfy the following phase relation:

$$u_{n\boldsymbol{q}}(\boldsymbol{r}) = e^{i\boldsymbol{G}\cdot\boldsymbol{r}} u_{n\boldsymbol{q}+\boldsymbol{G}}(\boldsymbol{r}). \tag{1.28}$$

This gauge choice is called the periodic gauge (Resta, 2000).

In this case, the Berry phase across the Brillouin zone is called Zak's phase (Zak, 1989)

$$\gamma_n = \int_{\mathrm{BZ}} d\boldsymbol{q} \cdot \langle u_n(\boldsymbol{q})|i\boldsymbol{\nabla}_{\boldsymbol{q}}|u_n(\boldsymbol{q})\rangle. \tag{1.29}$$

This phase plays an important role in the formation of Wannier-Stark ladders (Wannier, 1962); see Sec. VII.B. We note that this phase is entirely due to the torus topology of the Brillouin zone, and it is the only way to realize a closed path in a one-dimensional parameter space. By analyzing the symmetry properties of Wannier functions (Kohn, 1959) of a one-dimensional crystal, Zak (1989) showed that γ_n is either 0 or π in the presence of inversion symmetry; a simple argument is given in Sec. II.C. If the crystal lacks inversion symmetry, γ_n can assume any value. Zak's phase is also related to macroscopic polarization of an insulator (King-Smith and Vanderbilt, 1993; Resta, 1994; Sipe and Zak, 1999); see Sec. II.C.

II. ADIABATIC TRANSPORT AND ELECTRIC POLARIZATION

One of the earlier examples of the Berry phase effect in crystals is the adiabatic transport in a one-dimensional band insulator, first considered by Thouless (1983). He found that if the potential varies slowly in time and returns to itself after some time, the particle transport during the time cycle can be expressed as a Berry phase and it is an integer. This idea was later generalized to many-body systems with interactions and disorder provided that the Fermi energy always lies in a bulk energy gap during the cycle (Niu and Thouless, 1984). Avron and Seiler (1985) studied the adiabatic transport in multiply connected systems. The scheme of adiabatic transport under one or several controlling parameters has proven very powerful and opened the door to the field of parametric charge pumping (Niu, 1990; Talyanskii et al., 1997; Brouwer, 1998; Switkes et al., 1999; Zhou et al., 1999). It also provides a firm foundation to the modern theory of polarization developed in the early 1990s (King-Smith and Vanderbilt, 1993; Ortiz and Martin, 1994; Resta, 1994).

A. Adiabatic current

Consider a one-dimensional band insulator under a slowly varying time-dependent perturbation. We assume the perturbation is periodic in time, i.e., the Hamiltonian satisfies

$$H(t + T) = H(t). \tag{2.1}$$

Since the time-dependent Hamiltonian still has the translational symmetry of the crystal, its instantaneous eigenstates have the Bloch form $e^{iqx}|u_n(q,t)\rangle$. It is convenient to work with the q-space representation of the Hamiltonian $H(q,t)$ [see Eq. (1.24)] with eigenstates $|u_n(q,t)\rangle$. We note that under this parametrization the wave vector q and time t are put on an equal footing as both are independent coordinates of the parameter space.

We are interested in the adiabatic current induced by the variation in external potentials. Apart from an unimportant overall phase factor and up to first order in the rate of the change in the Hamiltonian, the wave function is given by (see the Appendix)

$$|u_n\rangle - i\hbar \sum_{n'\neq n} \frac{|u_{n'}\rangle\langle u_{n'}|\partial u_n/\partial t\rangle}{\varepsilon_n - \varepsilon_{n'}}. \tag{2.2}$$

The velocity operator in the q representation has the form $v(q,t) = \partial H(q,t)/\partial(\hbar q)$.[6] Hence, the average velocity in a state of given q is found to first order as

$$\begin{aligned} v_n(q) = &\frac{\partial \varepsilon_n(q)}{\hbar \partial q} \\ &- i \sum_{n'\neq n} \left\{ \frac{\langle u_n|\partial H/\partial q|u_{n'}\rangle\langle u_{n'}|\partial u_n/\partial t\rangle}{\varepsilon_n - \varepsilon_{n'}} - \mathrm{c.c.} \right\}, \end{aligned} \tag{2.3}$$

where c.c. denotes the complex conjugate. Using the fact that $\langle u_n|\partial H/\partial q|u_{n'}\rangle = (\varepsilon_n - \varepsilon_{n'})\langle \partial u_n/\partial q|u_{n'}\rangle$ and the identity $\sum_{n'}|u_{n'}\rangle\langle u_{n'}| = 1$, we find

$$v_n(q) = \frac{\partial \varepsilon_n(q)}{\hbar \partial q} - i\left[\left\langle \frac{\partial u_n}{\partial q} \middle| \frac{\partial u_n}{\partial t} \right\rangle - \left\langle \frac{\partial u_n}{\partial t} \middle| \frac{\partial u_n}{\partial q} \right\rangle \right]. \tag{2.4}$$

The second term is exactly the Berry curvature Ω_{qt}^n defined in the parameter space (q,t) [see Eq. (1.11)]. Therefore Eq. (2.4) can be recast into a compact form

$$v_n(q) = \frac{\partial \varepsilon_n(q)}{\hbar \partial q} - \Omega_{qt}^n. \tag{2.5}$$

Upon integration over the Brillouin zone, the zeroth-order term given by the derivative of the band energy vanishes, and only the first-order term survives. The induced adiabatic current is given by

---

[6]The velocity operator is defined by $\boldsymbol{v} \equiv \dot{\boldsymbol{r}} = (i/\hbar)[H,\boldsymbol{r}]$. In the q representation, it becomes $\boldsymbol{v}(\boldsymbol{q}) = e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}(i/\hbar)[\boldsymbol{r},H]e^{i\boldsymbol{q}\cdot\boldsymbol{r}} = \partial H(\boldsymbol{q},t)/\partial(\hbar\boldsymbol{q})$.
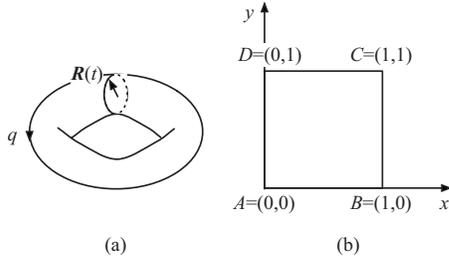
FIG. 1. Brillouin zone as a torus. (a) A torus with its surface parametrized by (q,t). The control parameter $\boldsymbol{R}(t)$ is periodic in t. (b) The equivalence of a torus: a rectangle with periodic boundary conditions: $\overline{AB}=\overline{DC}$ and $\overline{AD}=\overline{BC}$. To make use of Stokes's theorem, we relax the boundary condition and allow the wave functions on parallel sides to have different phases.

$$j = -\sum_n \int_{BZ} \frac{dq}{2\pi} \Omega_{qt}^n, \tag{2.6}$$

where the sum is over filled bands. We have thus derived the result that the adiabatic current induced by a time-dependent perturbation in a band is equal to the q integral of the Berry curvature Ω_{qt}^n (Thouless, 1983).

B. Quantized adiabatic particle transport

Next we consider the particle transport for the nth band over a time cycle given by

$$c_n = -\int_0^T dt \int_{BZ} \frac{dq}{2\pi} \Omega_{qt}^n. \tag{2.7}$$

Since the Hamiltonian $H(q,t)$ is periodic in both t and q, the parameter space of $H(q,t)$ is a torus, schematically shown in Fig. 1(a). By definition (1.12), $2\pi c_n$ is nothing but the Berry phase over the torus.

In Sec. I.C.3 we showed that the Berry phase over a closed manifold, the surface of a sphere S^2 in that case, is quantized in the unit of 2π. Here we prove that it is also true in the case of a torus. Our strategy is to evaluate the surface integral (2.7) using Stokes's theorem, which requires the surface to be simply connected. To do that, we cut the torus open and transform it into a rectangle, as shown in Fig. 1(b). The basis function along the contour of the rectangle is assumed to be single valued. We introduce $x=t/T$ and $y=q/2\pi$. According to Eq. (1.10), the Berry phase in Eq. (2.7) can be written into a contour integral of the Berry vector potential, i.e.,

$$c = \frac{1}{2\pi}\left\{\int_A^B dx\,\mathcal{A}_x(x,0) + \int_B^C dy\,\mathcal{A}_y(1,y)\right.$$
$$\left. + \int_C^D dx\,\mathcal{A}_x(x,1) + \int_D^A dy\,\mathcal{A}_y(0,y)\right\}$$
$$= \frac{1}{2\pi}\left\{\int_0^1 dx[\mathcal{A}_x(x,0) - \mathcal{A}_x(x,1)]\right.$$

$$\left. - \int_0^1 dy[\mathcal{A}_y(0,y) - \mathcal{A}_y(1,y)]\right\}, \tag{2.8}$$

where the band index n is dropped for simplicity. Now consider the integration over x. By definition $\mathcal{A}_x(x,y) = \langle u(x,y)|i\nabla_x|u(x,y)\rangle$. Recall that $|u(x,0)\rangle$ and $|u(x,1)\rangle$ describe physically equivalent states, therefore they can only differ by a phase factor, i.e., $e^{i\theta_x(x)}|u(x,1)\rangle = |u(x,0)\rangle$. We thus have

$$\int_0^1 dx[\mathcal{A}_x(x,0) - \mathcal{A}_x(x,1)] = \theta_x(1) - \theta_x(0). \tag{2.9}$$

Similarly,

$$\int_0^1 dy[\mathcal{A}_y(0,y) - \mathcal{A}_y(1,y)] = \theta_y(1) - \theta_y(0), \tag{2.10}$$

where $e^{i\theta_y(y)}|u(y,1)\rangle = |u(y,0)\rangle$. The total integral is

$$c = \frac{1}{2\pi}[\theta_x(1) - \theta_x(0) + \theta_y(0) - \theta_y(1)]. \tag{2.11}$$

On the other hand, using the phase matching relations at the four corners A, B, C, and D,

$$e^{i\theta_x(0)}|u(0,1)\rangle = |u(0,0)\rangle,$$

$$e^{i\theta_x(1)}|u(1,1)\rangle = |u(1,0)\rangle,$$

$$e^{i\theta_y(0)}|u(1,0)\rangle = |u(0,0)\rangle,$$

$$e^{i\theta_y(1)}|u(1,1)\rangle = |u(0,1)\rangle,$$

we obtain

$$|u(0,0)\rangle = e^{i[\theta_x(1)-\theta_x(0)+\theta_y(0)-\theta_y(1)]}|u(0,0)\rangle. \tag{2.12}$$

The single valuedness of $|u\rangle$ requires that the exponent must be an integer multiple of 2π. Therefore the transported particle number c, given in Eq. (2.11), must be quantized. This integer is called the first Chern number, which characterizes the topological structure of the mapping from the parameter space (q,t) to the Bloch states $|u(q,t)\rangle$. Note that in our proof we made no reference to the original physical system; the quantization of the Chern number is always true as long as the Hamiltonian is periodic in both parameters.

For a particular case in which the entire periodic potential is sliding, an intuitive picture of the quantized particle transport is the following. If the periodic potential slides its position without changing its shape, we expect that the electrons simply follow the potential. If the potential shifts one spatial period in the time cycle, the particle transport should be equal to the number of filled Bloch bands (double if the spin degeneracy is counted). This follows from the fact that there is on average one state per unit cell in each filled band.

1. Conditions for nonzero particle transport in cyclic motion

We have shown that the adiabatic particle transport over a time period takes the form of the Chern number

and it is quantized. However, the exact quantization does not guarantee that the electrons will be transported at the end of the cycle because zero is also an integer. According to the discussion in Sec. I.C.3, the Chern number counts the net number of monopoles enclosed by the surface. Therefore the number of transported electrons can be related to the number of monopoles, which are degeneracy points in the parameter space.

To formulate this problem, we let the Hamiltonian depend on time through a set of control parameters $\boldsymbol{R}(t)$, i.e.,

$$H(q,t) = H(q,\boldsymbol{R}(t)), \quad \boldsymbol{R}(t+T) = \boldsymbol{R}(t). \tag{2.13}$$

The particle transport is now given by, in terms of $\boldsymbol{R}$,

$$c_n = \frac{1}{2\pi} \oint dR_\alpha \int_{\mathrm{BZ}} dq \, \Omega^n_{qR_\alpha}. \tag{2.14}$$

If $\boldsymbol{R}(t)$ is a smooth function of t, as it is usually the case for physical quantities, then $\boldsymbol{R}$ must have at least two components, say R_1 and R_2. Otherwise, the trajectory of $\boldsymbol{R}(t)$ cannot trace out a circle on the torus [see Fig. 1(a)]. To find the monopoles inside the torus, we now relax the constraint that R_1 and R_2 can only move on the surface and extend their domains inside the torus such that the parameter space of (q, R_1, R_2) becomes a toroid. Thus, the criterion for c_n to be nonzero is that a degeneracy point must occur somewhere inside the torus as one varies q, R_1, and R_2. In the context of quasi-one-dimensional ferroelectrics, Onoda *et al.* (2004b) discussed the situation where $\boldsymbol{R}$ has three components and showed how the topological structure in the $\boldsymbol{R}$ space affects the particle transport.

2. Many-body interactions and disorder

Above we considered only band insulators of noninteracting electrons. However, in real materials both many-body interactions and disorder are ubiquitous. Niu and Thouless (1984) studied this problem and showed that in the general case the quantization of particle transport is still valid as long as the system remains an insulator during the whole process.

Consider a time-dependent Hamiltonian of an N-particle system

$$H(t) = \sum_i^N \left[\frac{\hat{p}_i^2}{2m} + U(x_i, t) \right] + \sum_{i>j}^N V(x_i - x_j), \tag{2.15}$$

where the one-particle potential $U(x_i, t)$ varies slowly in time and repeats itself in period T. Note that we have not assumed any specific periodicity of the potential $U(x_i, t)$. The trick is to use the so-called twisted boundary condition by requiring that the many-body wave function satisfies

$$\Phi(x_1, \dots, x_i + L, \dots, x_N) = e^{i\kappa L} \Phi(x_1, \dots, x_i, \dots, x_N), \tag{2.16}$$

where L is the size of the system. This is equivalent to solving a κ-dependent Hamiltonian

$$H(\kappa, t) = \exp\left(i\kappa \sum x_i\right) H(t) \exp\left(-i\kappa \sum x_i\right) \tag{2.17}$$

with the strict periodic boundary condition

$$\tilde{\Phi}(\kappa; x_1, \dots, x_i + L, \dots, x_N) = \tilde{\Phi}_\kappa(\kappa; x_1, \dots, x_i, \dots, x_N). \tag{2.18}$$

The Hamiltonian $H(\kappa, t)$ together with the boundary condition (2.18) describes a one-dimensional system placed on a ring of length L and threaded by a magnetic flux of $(\hbar/e)\kappa L$ (Kohn, 1964). We note that the above transformation (2.17) with the boundary condition (2.18) is similar to that of the one-particle case, given by Eqs. (1.24) and (1.25).

One can verify that the current operator is given by $\partial H(\kappa, t)/\partial(\hbar\kappa)$. For each κ, we can repeat the same steps in Sec. II.A by identifying $|u_n\rangle$ in Eq. (2.2) as the many-body ground-state $|\tilde{\Phi}_0\rangle$ and $|u_{n'}\rangle$ as the excited state. We have

$$j(\kappa) = \frac{\partial \varepsilon(\kappa)}{\hbar \partial \kappa} - i\left[\left\langle \frac{\partial \tilde{\Phi}_0}{\partial \kappa} \middle| \frac{\partial \tilde{\Phi}_0}{\partial t} \right\rangle - \left\langle \frac{\partial \tilde{\Phi}_0}{\partial t} \middle| \frac{\partial \tilde{\Phi}_0}{\partial \kappa} \right\rangle \right]$$
$$= \frac{\partial \varepsilon(\kappa)}{\hbar \partial \kappa} - \tilde{\Omega}_{\kappa t}. \tag{2.19}$$

So far the derivation is formal and we still cannot see why the particle transport should be quantized. The key step is achieved by realizing that if the Fermi energy lies in a gap, then the current $j(\kappa)$ should be insensitive to the boundary condition specified by κ (Thouless, 1981; Niu and Thouless, 1984). Consequently, we can take the thermodynamic limit and average $j(\kappa)$ over different boundary conditions. Note that κ and $\kappa + 2\pi/L$ describe the same boundary condition in Eq. (2.16). Therefore the parameter space for κ and t is a torus T^2:$\{0 < \kappa < 2\pi/L, 0 < t < T\}$. The particle transport is given by

$$c = -\frac{1}{2\pi} \int_0^T dt \int_0^{2\pi/L} d\kappa \, \tilde{\Omega}_{\kappa t}, \tag{2.20}$$

which, according to the previous discussion, is quantized.

We emphasize that the quantization of the particle transport only depends on two conditions:

(1) The ground state is separated from the excited states in the bulk by a finite energy gap.

(2) The ground state is nondegenerate.

The exact quantization of the Chern number in the presence of many-body interactions and disorder is remarkable. Usually, small perturbations to the Hamiltonian result in small changes of physical quantities. However, the fact that the Chern number must be an integer means that it can only be changed in a discontinuous way and does not change at all if the perturbation is small. This is a general outcome of the topological invariance.

Later we show that the same quantity also appears in the quantum Hall effect. Equation (2.19), the induced current, also provides a many-body formulation for adiabatic transport.

3. Adiabatic pumping

The phenomenon of adiabatic transport is sometimes called adiabatic pumping because it can generate a dc current I via periodic variations of some parameters of the system, i.e.,

$$I = ec\nu, \tag{2.21}$$

where c is the Chern number and ν is the frequency of the variation. Niu (1990) suggested that the exact quantization of the adiabatic transport can be used as a standard for charge current and proposed an experimental realization in nanodevices, which could serve as a charge pump. Later a similar device was realized in the experimental study of acoustoelectric current induced by a surface acoustic wave in a one-dimensional channel in a GaAs-Al$_x$Ga$_{1-x}$ heterostructure (Talyanskii *et al.*, 1997). The same idea has led to the proposal of a quantum spin pump in an antiferromagnetic chain (Shindou, 2005).

Recently much effort has focused on adiabatic pumping in mesoscopic systems (Brouwer, 1998; Zhou *et al.*, 1999; Avron *et al.*, 2001, 2004; Sharma and Chamon, 2001; Mucciolo *et al.*, 2002; Zheng *et al.*, 2003). Experimentally both charge and spin pumping have been observed in a quantum dot system (Switkes *et al.*, 1999; Watson *et al.*, 2003). Instead of the wave function, the central quantity in a mesoscopic system is the scattering matrix. Brouwer (1998) showed that the pumped charge over a time period is given by

$$Q(m) = \frac{e}{\pi} \int_A dX_1 dX_2 \sum_\beta \sum_{\alpha \in m} \Im \frac{\partial S_{\alpha\beta}^*}{\partial X_1} \frac{\partial S_{\alpha\beta}}{\partial X_2}, \tag{2.22}$$

where m labels the contact, X_1 and X_2 are two external parameters whose trace encloses the area A in the parameter space, α and β label the conducting channels, and $S_{\alpha\beta}$ is the scattering matrix. Although the physical descriptions of these open systems are dramatically different from the closed ones, the concepts of gauge field and geometric phase can still be applied. The integrand in Eq. (2.22) can be thought as the Berry curvature $\Omega_{X_1 X_2} = -2\Im \langle \partial_{X_1} u | \partial_{X_2} u \rangle$ if we identify the inner product of the state vector with the matrix product. Zhou *et al.* (2003) showed the pumped charge (spin) is essentially the Abelian (non-Abelian) geometric phase associated with scattering matrix $S_{\alpha\beta}$.

C. Electric polarization of crystalline solids

Electric polarization is one of the fundamental quantities in condensed-matter physics, essential to any proper description of dielectric phenomena of matter. Despite its importance, the theory of polarization in crystals had been plagued by the lack of a proper microscopic understanding. The main difficulty lies in the fact

that in crystals the charge distribution is periodic in space, for which the electric dipole operator is not well defined. This difficulty is most exemplified in covalent solids, where the electron charges are continuously distributed between atoms. In this case, simple integration over charge density would give arbitrary values depending on the choice of the unit cell (Martin, 1972, 1974). It has prompted the question whether the electric polarization can be defined as a bulk property. These problems are eventually solved by the modern theory of polarization (King-Smith and Vanderbilt, 1993; Resta, 1994), where it is shown that only the change in polarization has physical meaning and it can be quantified using the Berry phase of the electronic wave function. The resulting Berry phase formula has been very successful in first-principles studies of dielectric and ferroelectric materials. Resta and Vanderbilt (2007) reviewed recent progress in this field.

Here we discuss the theory of polarization based on the concept of adiabatic transport. Their relation is revealed by elementary arguments from macroscopic electrostatics (Ortiz and Martin, 1994). We begin with

$$\nabla \cdot \boldsymbol{P}(\boldsymbol{r}) = -\rho(\boldsymbol{r}), \tag{2.23}$$

where $\boldsymbol{P}(\boldsymbol{r})$ is the polarization density and $\rho(\boldsymbol{r})$ is the charge density. Coupled with the continuity equation

$$\frac{\partial \rho(\boldsymbol{r})}{\partial t} + \nabla \cdot \boldsymbol{j} = 0, \tag{2.24}$$

Eq. (2.23) leads to

$$\nabla \cdot \left(\frac{\partial \boldsymbol{P}}{\partial t} - \boldsymbol{j} \right) = 0. \tag{2.25}$$

Therefore up to a divergence-free part,[7] the change in the polarization density is given by

$$\Delta P_\alpha = \int_0^T dt\, j_\alpha. \tag{2.26}$$

Equation (2.26) can be interpreted in the following way: The polarization difference between two states is given by the integrated bulk current as the system adiabatically evolves from the initial state at $t=0$ to the final state at $t=T$. This description implies a time-dependent Hamiltonian $H(t)$, and the electric polarization can be regarded as "unquantized" adiabatic particle transport. The above interpretation is also consistent with experiments, as it is always the change in the polarization that has been measured (Resta and Vanderbilt, 2007).

Obviously, the time t in the Hamiltonian can be replaced by any scalar that describes the adiabatic process. For example, if the process corresponds to a deformation of the crystal, then it makes sense to use the param-

---

[7]The divergence-free part of the current is usually given by the magnetization current. In a uniform system, such current vanishes identically in the bulk. Hirst (1997) gave an in-depth discussion on the separation between polarization and magnetization current.

eter that characterizes the atomic displacement within a unit cell. For general purpose, we assume the adiabatic transformation is parametrized by a scalar $\lambda(t)$ with $\lambda(0)=0$ and $\lambda(T)=1$. It follows from Eqs. (2.6) and (2.26) that

$$\Delta P_\alpha = e \sum_n \int_0^1 d\lambda \int_{BZ} \frac{d\boldsymbol{q}}{(2\pi)^d} \Omega^n_{q_\alpha \lambda}, \qquad (2.27)$$

where d is the dimensionality of the system. This is the Berry phase formula obtained by King-Smith and Vanderbilt (1993).

In numerical calculations, a two-point version of Eq. (2.27) that involves only the initial and final states of the system is commonly used to reduce the computational load. This version is obtained under the periodic gauge [see Eq. (1.28)].[8] The Berry curvature $\Omega_{q_\alpha \lambda}$ is written as $\partial_{q_\alpha} \mathcal{A}_\lambda - \partial_\lambda \mathcal{A}_{q_\alpha}$. Under the periodic gauge, A_λ is periodic in q_α, and integration of $\partial_{q_\alpha} \mathcal{A}_\lambda$ over q_α vanishes. Hence

$$\Delta P_\alpha = e \sum \int_{BZ} \frac{d\boldsymbol{q}}{(2\pi)^d} \mathcal{A}^n_{q_\alpha} \Big|_{\lambda=0}^1. \qquad (2.28)$$

In view of Eq. (2.28), both the adiabatic transport and the electric polarization can be regarded as the manifestation of Zak's phase, given in Eq. (1.29).

A price must be paid, however, to use the two-point formula, namely, the polarization in Eq. (2.28) is determined up to an uncertainty quantum. Since the integral (2.28) does not track the history of λ, there is no information on how many cycles λ has gone through. According to our discussion on quantized particle transport in Sec. II.B, for each cycle an integer number of electrons are transported across the sample, hence the polarization is changed by multiple of the quantum

$$\frac{e\boldsymbol{a}}{\mathcal{V}_0}, \qquad (2.29)$$

where $\boldsymbol{a}$ is the Bravais lattice vector and $\mathcal{V}_0$ is the volume of the unit cell.

Because of this uncertainty quantum, the polarization may be regarded as a multivalued quantity with each value differed by the quantum. With this in mind, consider Zak's phase in a one-dimensional system with inversion symmetry. Now we know that Zak's phase is just $2\pi/e$ times the polarization density P. Under spatial inversion, P transforms to $-P$. On the other hand, inversion symmetry requires that P and $-P$ describes the same state, which is only possible if P and $-P$ differ by multiple of the polarization quantum ea. Therefore P is either 0 or $ea/2$ (modulo ea). Any other value of P will break the inversion symmetry. Consequently, Zak's phase can only take the value 0 or π (modulo 2π).

King-Smith and Vanderbilt (1993) further showed that, based on Eq. (2.28), the polarization per unit cell can be defined as the dipole moment of the Wannier charge density,

$$\boldsymbol{P} = -e \sum_n \int d\boldsymbol{r}\, \boldsymbol{r} |W_n(\boldsymbol{r})|^2, \qquad (2.30)$$

where $W_n(\boldsymbol{r})$ is the Wannier function of the nth band,

$$W_n(\boldsymbol{r} - R) = \sqrt{N}\mathcal{V}_0 \int_{BZ} \frac{d\boldsymbol{q}}{(2\pi)^3} e^{i\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{R})} u_{nk}(\boldsymbol{r}). \qquad (2.31)$$

In this definition, one effectively maps a band insulator into a periodic array of localized distributions with truly quantized charges. This resembles an ideal ionic crystal where the polarization can be understood in the classical picture of localized charges. The quantum uncertainty found in Eq. (2.29) is reflected by the fact that the Wannier center position is defined only up to a lattice vector.

Before concluding, we point out that the polarization defined above is clearly a bulk quantity as it is given by the Berry phase of the ground-state wave function. A many-body formulation was developed by Ortiz and Martin (1994) based on the work of Niu and Thouless (1984).

Recent development in this field falls into two categories. On the computational side, calculating polarization in finite electric fields has been addressed, which has a strong influence on density functional theory in extended systems (Nunes and Vanderbilt, 1994; Nunes and Gonze, 2001; Souza *et al.*, 2002). On the theory side, Resta (1998) proposed a quantum-mechanical position operator for extended systems. It was shown that the expectation value of such an operator can be used to characterize the phase transition between the metallic and insulating states (Resta and Sorella, 1999; Souza *et al.*, 2000) and is closely related to the phenomenon of electron localization (Kohn, 1964).

1. The Rice-Mele model

So far our discussion of the adiabatic transport and electric polarization has been rather abstract. We now consider a concrete example: a one-dimensional dimerized lattice model described by the following Hamiltonian:

$$H = \sum_j \left(\frac{t}{2} + (-1)^j \frac{\delta}{2} \right) (c_j^\dagger c_{j+1} + \text{H.c.}) + \Delta(-1)^j c_j^\dagger c_j, \qquad (2.32)$$

where t is the uniform hopping amplitude, δ is the dimerization order, and Δ is a staggered sublattice potential. It is the prototype Hamiltonian for a class of one-dimensional ferroelectrics. At half filling, the system is a metal for $\Delta = \delta = 0$, and an insulator otherwise. Rice and Mele (1982) considered this model in the study of solitons in polyenes. It was later used to study ferroelectricity (Vanderbilt and King-Smith, 1993; Onoda *et al.*,

---

[8]A more general phase choice is given by the path-independent gauge $|u_n(\boldsymbol{q},\lambda)\rangle = e^{i[\theta(\boldsymbol{q})+\boldsymbol{G}\cdot\boldsymbol{r}]}|u_n(\boldsymbol{q}+\boldsymbol{G},\lambda)\rangle$, where $\theta(\boldsymbol{q})$ is an arbitrary phase (Ortiz and Martin, 1994).
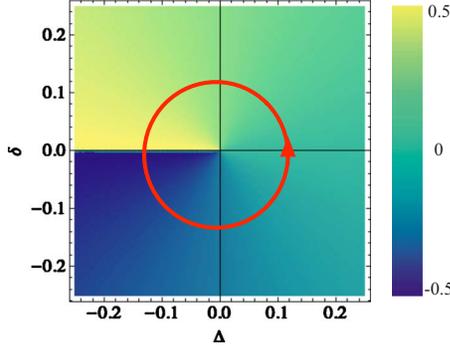
FIG. 2. (Color online) Polarization as a function of Δ and δ in the Rice-Mele model. The unit is ea with a the lattice constant. The line of discontinuity can be chosen anywhere depending on the particular phase choice of the eigenstate.

2004b). If $\Delta = 0$, it reduces to the celebrated Su-Shrieffer-Heeger model (Su *et al.*, 1979).

Assuming periodic boundary conditions, the Bloch representation of the above Hamiltonian is given by $H(q) = \boldsymbol{h}(q) \cdot \boldsymbol{\sigma}$, where

$$\boldsymbol{h} = \left(t \cos \frac{qa}{2}, -\delta \sin \frac{qa}{2}, \Delta \right). \tag{2.33}$$

This is the two-level model discussed in Sec. I.C.3. Its energy spectrum consists of two bands with eigenenergies $\varepsilon_\pm = \pm(\Delta^2 + \delta^2 \sin^2 qa/2 + t^2 \cos^2 qa/2)^{1/2}$. The degeneracy point occurs at

$$\Delta = 0, \quad \delta = 0, \quad q = \pi/a. \tag{2.34}$$

The polarization is calculated using the two-point formula (2.28) with the Berry connection given by

$$\mathcal{A}_q = \partial_q \phi \mathcal{A}_\phi + \partial_q \theta \mathcal{A}_\theta = \sin^2 \frac{\theta}{2} \partial_q \phi, \tag{2.35}$$

where θ and ϕ are the spherical angles of $\boldsymbol{h}$.

Consider the case of $\Delta = 0$. In the parameter space of $\boldsymbol{h}$, it lies in the x-y plane with $\theta = \pi/2$. As q varies from 0 to $2\pi/a$, ϕ changes from 0 to π if $\delta < 0$ and 0 to $-\pi$ if $\delta > 0$. Therefore the polarization difference between $P(\delta)$ and $P(-\delta)$ is $ea/2$. This is consistent with the observation that the state with $P(-\delta)$ can be obtained by shifting the state with $P(\delta)$ by half of the unit cell length a.

Figure 2 shows the calculated polarization for arbitrary Δ and δ. If the system adiabatically evolves along a loop enclosing the degeneracy point $(0,0)$ in the (Δ, δ) space, then the polarization will be changed by ea, which means that if we allow (Δ, δ) to change in time along this loop, for example, $\Delta(t) = \Delta_0 \sin(t)$ and $\delta(t) = \delta_0 \cos(t)$, a quantized charge of e is pumped out of the system after one cycle. On the other hand, if the loop does not contain the degeneracy point, then the pumped charge is zero.

III. ELECTRON DYNAMICS IN AN ELECTRIC FIELD

The dynamics of Bloch electrons under the perturbation of an electric field is one of the oldest problems in solid-state physics. It is usually understood that while the electric field can drive electron motion in momentum space, it does not appear in the electron velocity; the latter is simply given by [see, for example, Chap. 12 of Ashcroft and Mermin (1976)]

$$v_n(\boldsymbol{q}) = \frac{\partial \varepsilon_n(\boldsymbol{q})}{\hbar \partial \boldsymbol{q}}. \tag{3.1}$$

Through recent progress on the semiclassical dynamics of Bloch electrons it has been made increasingly clear that this description is incomplete. In the presence of an electric field, an electron can acquire an anomalous velocity proportional to the Berry curvature of the band (Chang and Niu, 1995, 1996; Sundaram and Niu, 1999). This anomalous velocity is responsible for a number of transport phenomena, in particular various Hall effects, which we study in this section.

A. Anomalous velocity

Consider a crystal under the perturbation of a weak electric field $\boldsymbol{E}$, which enters into the Hamiltonian through the coupling to the electrostatic potential $\phi(\boldsymbol{r})$. A uniform $\boldsymbol{E}$ means that $\phi(\boldsymbol{r})$ varies linearly in space and breaks the translational symmetry of the crystal so that Bloch's theorem cannot be applied. To avoid this difficulty, one can let the electric field enter through a uniform vector potential $\boldsymbol{A}(t)$ that changes in time. Using the Peierls substitution, the Hamiltonian is written as

$$H(t) = \frac{[\hat{\boldsymbol{p}} + e\boldsymbol{A}(t)]^2}{2m} + V(\boldsymbol{r}). \tag{3.2}$$

This is the time-dependent problem studied in last section. Transforming to the $\boldsymbol{q}$-space representation, we have

$$H(\boldsymbol{q}, t) = H\left(\boldsymbol{q} + \frac{e}{\hbar} \boldsymbol{A}(t) \right). \tag{3.3}$$

Introduce the gauge-invariant crystal momentum

$$\boldsymbol{k} = \boldsymbol{q} + \frac{e}{\hbar} \boldsymbol{A}(t). \tag{3.4}$$

The parameter-dependent Hamiltonian can be simply written as $H(\boldsymbol{k}(\boldsymbol{q}, t))$. Hence the eigenstates of the time-dependent Hamiltonian can be labeled by a single parameter $\boldsymbol{k}$. Moreover, because $\boldsymbol{A}(t)$ preserves the translational symmetry, $\boldsymbol{q}$ is still a good quantum number and is a constant of motion $\dot{\boldsymbol{q}} = 0$. It then follows from Eq. (3.4) that $\boldsymbol{k}$ satisfies the following equation of motion:

$$\dot{\boldsymbol{k}} = -\frac{e}{\hbar} \boldsymbol{E}. \tag{3.5}$$

Using $\partial/\partial q_\alpha = \partial/\partial k_\alpha$ and $\partial/\partial t = -(e/\hbar)E_\alpha \partial/\partial k_\alpha$, the general formula (2.5) for the velocity in a given state $\boldsymbol{k}$ becomes

$$\boldsymbol{v}_n(\boldsymbol{k}) = \frac{\partial \varepsilon_n(\boldsymbol{k})}{\hbar \partial \boldsymbol{k}} - \frac{e}{\hbar} \boldsymbol{E} \times \boldsymbol{\Omega}_n(\boldsymbol{k}), \qquad (3.6)$$

where $\boldsymbol{\Omega}_n(\boldsymbol{k})$ is the Berry curvature of the nth band:

$$\boldsymbol{\Omega}_n(\boldsymbol{k}) = i\langle \boldsymbol{\nabla}_{\boldsymbol{k}} u_n(\boldsymbol{k})| \times |\boldsymbol{\nabla}_{\boldsymbol{k}} u_n(\boldsymbol{k})\rangle. \qquad (3.7)$$

We can see that, in addition to the usual band dispersion contribution, an extra term previously known as an anomalous velocity also contributes to $\boldsymbol{v}_n(\boldsymbol{k})$. This velocity is always transverse to the electric field, which will give rise to a Hall current. Historically the anomalous velocity was obtained by Karplus and Luttinger (1954), Kohn and Luttinger (1957), and Adams and Blount (1959); its relation to the Berry phase was realized much later. In Sec. V we rederive this term using a wave-packet approach.

B. Berry curvature: Symmetry considerations

The velocity formula (3.6) reveals that, in addition to the band energy, the Berry curvature of the Bloch bands is also required for a complete description of the electron dynamics. However, the conventional formula [Eq. (3.1)] has much success in describing various electronic properties in the past. It is thus important to know under what conditions the Berry curvature term cannot be neglected.

The general form of the Berry curvature $\boldsymbol{\Omega}_n(\boldsymbol{k})$ can be obtained via symmetry analysis. The velocity formula (3.6) should be invariant under time-reversal and spatial inversion operations if the unperturbed system has these symmetries. Under time reversal, $\boldsymbol{v}_n$ and $\boldsymbol{k}$ change sign while $\boldsymbol{E}$ is fixed. Under spatial inversion, $\boldsymbol{v}_n$, $\boldsymbol{k}$, and $\boldsymbol{E}$ change sign. If the system has time-reversal symmetry, the symmetry condition on Eq. (3.6) requires that

$$\boldsymbol{\Omega}_n(-\boldsymbol{k}) = -\boldsymbol{\Omega}_n(\boldsymbol{k}). \qquad (3.8)$$

If the system has spatial inversion symmetry, then

$$\boldsymbol{\Omega}_n(-\boldsymbol{k}) = \boldsymbol{\Omega}_n(\boldsymbol{k}). \qquad (3.9)$$

Therefore, for crystals with simultaneous time-reversal and spatial inversion symmetry the Berry curvature vanishes identically throughout the Brillouin zone. In this case Eq. (3.6) reduces to the simple expression (3.1). However, in systems with broken either time-reversal or inversion symmetries, their proper description requires the use of the full velocity formula (3.6).

There are many important physical systems where both symmetries are not simultaneously present. For example, in the presence of ferromagnetic or antiferromagnetic ordering the crystal breaks the time-reversal symmetry. Figure 3 shows the Berry curvature on the Fermi surface of fcc Fe. As shown the Berry curvature is negligible in most areas in the momentum space and displays sharp and pronounced peaks in regions where the Fermi lines [intersection of the Fermi surface with
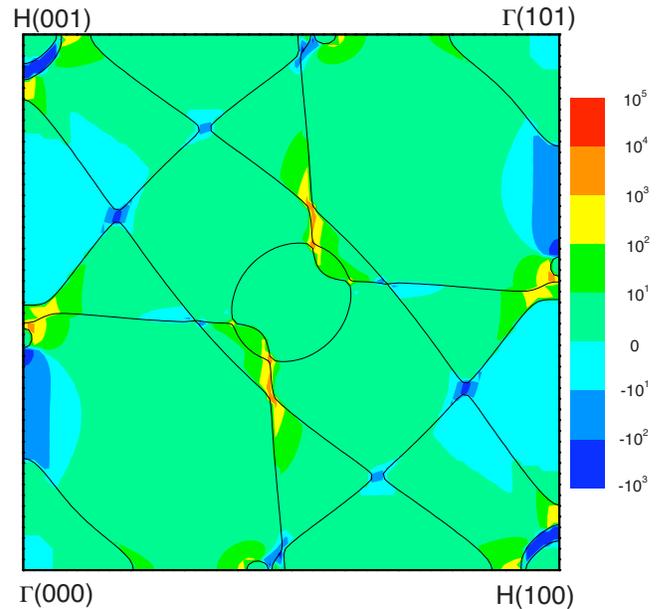

FIG. 3. (Color online) Fermi surface in (010) plane (solid lines) and the integrated Berry curvature $-\Omega_z(\boldsymbol{k})$ in atomic units (color map) of fcc Fe. From Yao et al., 2004.

(010) plane] have avoided crossings due to spin-orbit coupling. Such a structure has been identified in other materials as well (Fang et al., 2003). Another example is provided by single-layered graphene sheet with staggered sublattice potential, which breaks inversion symmetry (Zhou et al., 2007). Figure 4 shows the energy band and Berry curvature of this system. The Berry curvature at valley K_1 and K_2 have opposite signs due to time-reversal symmetry. We note that as the gap approaches zero, the Berry phase acquired by an electron during one circle around the valley becomes exactly $\pm\pi$. This Berry phase of π has been observed in intrinsic
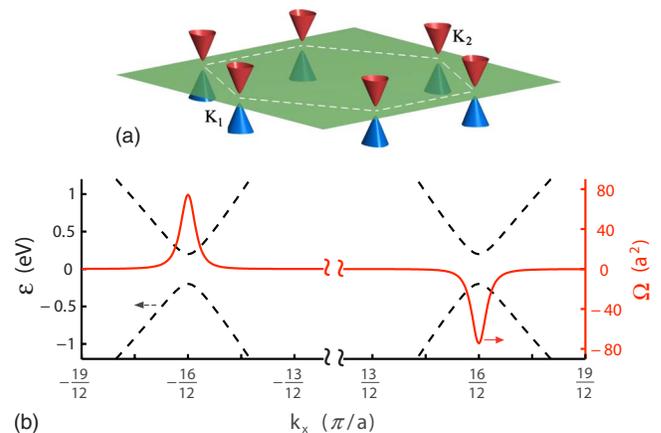

FIG. 4. (Color online) Energy bands (top panel) and Berry curvature of the conduction band (bottom panel) of a graphene sheet with broken inversion symmetry. The first Brillouin zone is outlined by the dashed lines, and two inequivalent valleys are labeled as K_1 and K_2. Details are presented in Xiao, Yao, and Niu (2007).

graphene sheet (Novoselov *et al.*, 2005; Zhang *et al.*, 2005).

C. The quantum Hall effect

The quantum Hall effect was discovered by Klitzing *et al.* (1980). They found that in a strong magnetic field the Hall conductivity of a two-dimensional (2D) electron gas is *exactly* quantized in the units of e^2/h. The exact quantization was subsequently explained by Laughlin (1981) based on gauge invariance and was later related to a topological invariance of the energy bands (Thouless *et al.*, 1982; Avron *et al.*, 1983; Niu *et al.*, 1985). Since then it has blossomed into an important research field in condensed-matter physics. In this section we focus only on the quantization aspect of the quantum Hall effect using the formulation developed so far.

Consider a two-dimensional band insulator. It follows from Eq. (3.6) that the Hall conductivity of the system is given by

$$\sigma_{xy} = \frac{e^2}{\hbar} \int_{\text{BZ}} \frac{d^2k}{(2\pi)^2} \Omega_{k_x k_y}, \qquad (3.10)$$

where the integration is over the entire Brillouin. Once again we encounter the situation where the Berry curvature is integrated over a closed manifold. Here σ_{xy} is the Chern number in the units of e^2/h, i.e.,

$$\sigma_{xy} = n \frac{e^2}{h}. \qquad (3.11)$$

Therefore the Hall conductivity is quantized for a two-dimensional band insulator of noninteracting electrons.

Historically the quantization of the Hall conductivity in a crystal was first shown by Thouless *et al.* (1982) for magnetic Bloch bands (see also Sec. VIII). It was shown that, due to the magnetic translational symmetry, the phase of the wave function in the magnetic Brillouin zone carries a vortex and leads to a *nonzero* quantized Hall conductivity (Kohmoto, 1985). However, it is clear from the above derivation that for the quantum Hall effect to occur the only condition is that the Chern number of the bands must be nonzero. It is possible that in some materials the Chern number can be nonzero even in the absence of an external magnetic field. Haldane (1988) constructed a tight-binding model on a honeycomb lattice which displays the quantum Hall effect with zero net flux per unit cell. Another model is proposed for semiconductor quantum well where the spin-orbit interaction plays the role of the magnetic field (Qi *et al.*, 2006; Liu *et al.*, 2008) and leads to a quantized Hall conductance. The possibility of realizing the quantum Hall effect without a magnetic field is attractive in device design.

Niu *et al.* (1985) further showed that the quantized Hall conductivity in two-dimensions is robust against many-body interactions and disorder [see also Avron and Seiler (1985)]. Their derivation involves the same technique discussed in Sec. II.B.2. A two-dimensional many-body system is placed on a torus by assuming pe-
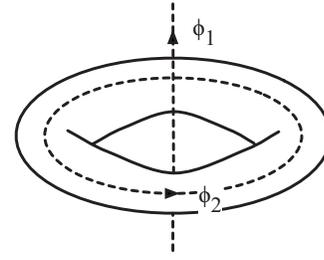

FIG. 5. Magnetic flux going through the holes of the torus.

riodic boundary conditions in both directions. One can then thread the torus with magnetic flux through its holes (Fig. 5) and make the Hamiltonian $H(\phi_1, \phi_2)$ depend on the flux ϕ_1 and ϕ_2. The Hall conductivity is calculated using the Kubo formula

$$\sigma_H = ie^2\hbar \sum_{n>0} \frac{\langle \Phi_0|v_1|\Phi_n\rangle\langle\Phi_n|v_2|\Phi_0\rangle - (1 \leftrightarrow 2)}{(\varepsilon_0 - \varepsilon_n)^2}, \quad (3.12)$$

where Φ_n is the many-body wave function with $|\Phi_0\rangle$ the ground state. In the presence of flux, the velocity operator is given by $v_i = \partial H(\kappa_1, \kappa_2)/\partial(\hbar\kappa_i)$ with $\kappa_i = (e/\hbar)\phi_i/L_i$ and L_i the dimensions of the system. We recognize that Eq. (3.12) is the summation formula (1.13) for the Berry curvature $\Omega_{\kappa_1\kappa_2}$ of the state $|\Phi_0\rangle$. The existence of a bulk energy gap guarantees that the Hall conductivity remains unchanged after thermodynamic averaging, which is given by

$$\sigma_H = \frac{e^2}{\hbar} \int_0^{2\pi/L_1} d\kappa_1 \int_0^{2\pi/L_2} d\kappa_2 \Omega_{\kappa_1\kappa_2}. \qquad (3.13)$$

Note that the Hamiltonian $H(\kappa_1, \kappa_2)$ is periodic in κ_i with period $2\pi/L_i$ because the system returns to its original state after the flux is changed by a flux quantum h/e (and κ_i changed by $2\pi/L_i$). Therefore the Hall conductivity is quantized even in the presence of many-body interaction and disorder. Due to the high precision of the measurement and the robustness of the quantization, the quantum Hall resistance is now used as the primary standard of resistance.

The geometric and topological ideas developed in the study of the quantum Hall effect has a far-reaching impact on modern condensed-matter physics. The robustness of the Hall conductivity suggests that it can be used as a topological invariance to classify many-body phases of electronic states with a bulk energy gap (Avron *et al.*, 1983): states with different topological orders (Hall conductivities in the quantum Hall effect) cannot be adiabatically transformed into each other; if that happens, a phase transition must occur. The Hall conductivity has important applications in strongly correlated electron systems, such as the fractional quantum Hall effect (Wen and Niu, 1990), and most recently the topological quantum computing [for a review, see Nayak *et al.* (2008)].

D. The anomalous Hall effect

Next we discuss the anomalous Hall effect, which refers to the appearance of a large spontaneous Hall current in a ferromagnet in response to an electric field alone [for early works in this field see Chien and Westgate (1980)]. Despite its century-long history and importance in sample characterization, the microscopic mechanism of the anomalous Hall effect has been a controversial subject and it comes to light only recently [for a recent review see Nagaosa *et al.* (2010)]. In the past, three mechanisms have been identified: the intrinsic contribution (Karplus and Luttinger, 1954; Luttinger, 1958), the extrinsic contributions from the skew (Smit, 1958), and side-jump scattering (Berger, 1970). The latter two describe the asymmetric scattering amplitudes for spin-up and spin-down electrons. It was later realized that the scattering-independent intrinsic contribution comes from the Berry phase supported anomalous velocity. This will be our primary interest here.

The intrinsic contribution to the anomalous Hall effect can be regarded as an "unquantized" version of the quantum Hall effect. The Hall conductivity is given by

$$\sigma_{xy} = \frac{e^2}{\hbar} \int \frac{d\boldsymbol{k}}{(2\pi)^d} f(\varepsilon_{\boldsymbol{k}}) \Omega_{k_x k_y}, \qquad (3.14)$$

where $f(\varepsilon_{\boldsymbol{k}})$ is the Fermi-Dirac distribution function. However, unlike the quantum Hall effect, the anomalous Hall effect does not require a nonzero Chern number of the band; for a band with zero Chern number, the local Berry curvature can be nonzero and give rise to a nonzero anomalous Hall conductivity.

Consider a generic Hamiltonian with spin-orbit (SO) split bands (Onoda, Sugimoto, and Nagaosa, 2006),

$$H = \frac{\hbar^2 k^2}{2m} + \lambda (\boldsymbol{k} \times \boldsymbol{\sigma}) \cdot \boldsymbol{e}_z - \Delta \sigma_z, \qquad (3.15)$$

where 2Δ is the SO split gap in the energy spectrum $\varepsilon_{\pm} = \hbar^2 k^2 / 2m \pm \sqrt{\lambda^2 k^2 + \Delta^2}$ and λ gives a linear dispersion in the absence of Δ. This model also describes spin-polarized two-dimensional electron gas with Rashba SO coupling, with λ the SO coupling strength and Δ the exchange field (Culcer *et al.*, 2003). Obviously the Δ term breaks time-reversal symmetry and the system is ferromagnetic. However, the Δ term alone will not lead to a Hall current as it only breaks the time-reversal symmetry in the spin space. The SO interaction is needed to couple the spin and orbital part together. The Berry curvature is given by, using Eq. (1.19),

$$\Omega_{\pm} = \mp \frac{\lambda^2 \Delta}{2(\lambda^2 k^2 + \Delta^2)^{3/2}}. \qquad (3.16)$$

The Berry curvatures of the two energy bands have opposite sign and are highly concentrated around the gap. (In fact, the Berry curvature has the same form of the Berry curvature in one valley of the graphene, shown in Fig. 4.) One can verify that the integration of the Berry
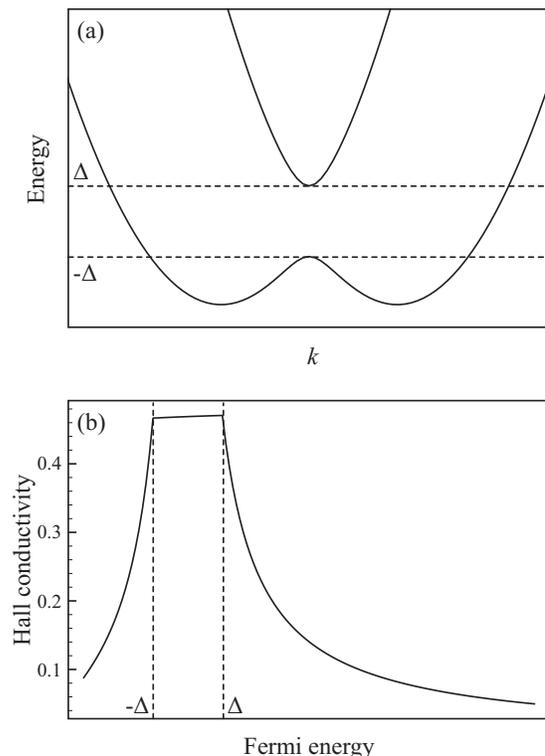

FIG. 6. Anomalous Hall effect in a simple two-band model. (a) Energy dispersion of spin-split bands. (b) The Hall conductivity $-\sigma_{xy}$ in the units of e^2/h as a function of Fermi energy.

curvature of a full band, $2\pi \int_0^{\infty} q \, dq \, \Omega_{\pm}$, is $\pm\pi$ for the upper and lower bands, respectively.[9]

Figure 6 shows the band dispersion, and the intrinsic Hall conductivity [Eq. (3.14)] as the Fermi energy sweeps across the SO split gap. As shown when the Fermi energy ε_F is in the gap region, the Hall conductivity reaches its maximum value (about $-e^2/2h$). If $\varepsilon_F < -\Delta$, the states with energies just below $-\Delta$, which contribute most to the Hall conductivity, are empty. If $\varepsilon_F > \Delta$, contributions from upper and lower bands cancel each other, and the Hall conductivity decreases quickly as ε_F moves away from the band gap. It is only when $-\Delta < \varepsilon_F < \Delta$, the Hall conductivity is resonantly enhanced (Onoda, Sugimoto, and Nagaosa, 2006).

1. Intrinsic versus extrinsic contributions

The above discussion does not take into account the fact that, unlike insulators, in metallic systems electron scattering can be important in transport phenomena. Two contributions to the anomalous Hall effect arises due to scattering: (i) the skew scattering that refers to the asymmetric scattering amplitude with respect to the scattering angle between the incoming and outgoing electron waves (Smit, 1958) and (ii) the side jump which is a sudden shift of the electron coordinates during scattering (Berger, 1970). In a more careful analysis, a sys-

---

[9]Since the integration is performed in an infinite momentum space, the result is not quantized in the unit of 2π.

tematic study of the anomalous Hall effect based on the semiclassical Boltzmann transport theory has been carried out (Sinitsyn, 2008). The basic idea is to solve the following Boltzmann equation:

$$\frac{\partial g_{\boldsymbol{k}}}{\partial t} - e\boldsymbol{E} \cdot \frac{\partial \varepsilon}{\hbar \partial \boldsymbol{k}} \frac{\partial f}{\partial \varepsilon} = \sum_{\boldsymbol{k}'} \omega_{\boldsymbol{k}\boldsymbol{k}'} \left[g_{\boldsymbol{k}} - g_{\boldsymbol{k}'} - \frac{\partial f}{\partial \varepsilon} e\boldsymbol{E} \cdot \delta\boldsymbol{r}_{\boldsymbol{k}\boldsymbol{k}'} \right],$$

(3.17)

where g is the nonequilibrium distribution function, $\omega_{\boldsymbol{k}\boldsymbol{k}'}$ represents the asymmetric skew scattering, and $\delta\boldsymbol{r}_{\boldsymbol{k}\boldsymbol{k}'}$ describes the side-jump of the scattered electrons. The Hall conductivity is the sum of different contributions

$$\sigma_H = \sigma_H^{\text{in}} + \sigma_H^{\text{sk}} + \sigma_H^{\text{sj}},$$

(3.18)

where σ_H^{in} is the intrinsic contribution given by Eq. (3.14), σ_H^{sk} is the skew scattering contribution, which is proportional to the relaxation time τ, and σ_H^{sj} is the side jump contribution, which is independent of τ. Note that, in addition to Berger's original proposal, σ_H^{sj} also includes two other contributions: the intrinsic skew-scattering and anomalous distribution function (Sinitsyn, 2008).

An important question is how to identify the dominant contribution to the anomalous Hall effect (AHE) among these mechanisms. If the sample is clean and the temperature is low, the relaxation time τ can be extremely large, and the skew scattering is expected to dominate. On the other hand, in dirty samples and at high temperatures σ_H^{sk} becomes small compared to both σ_H^{in} and σ_H^{sj}. Because the Berry phase contribution σ_H^{in} is independent of scattering, it can be readily evaluated using first-principles methods or effective Hamiltonians. Excellent agreement with experiments has been demonstrated in ferromagnetic transition metals and semiconductors (Jungwirth et al., 2002; Fang et al., 2003; Yao et al., 2004, 2007; Xiao, Yao, et al., 2006), which leaves little room for the side-jump contribution.

In addition, a number of experimental results also gave favorable evidence for the dominance of the intrinsic contribution (Lee et al., 2004b; Mathieu et al., 2004; Sales et al., 2006; Zeng et al., 2006; Chun et al., 2007). In particular, Tian et al. (2009) recently measured the anomalous Hall conductivity in Fe thin films. By varying the film thickness and the temperature, they are able to control various scattering process such as the impurity scattering and the phonon scattering. Figure 7 shows their measured σ_{ah} as a function of $\sigma_{xx}(T)^2$. One can see that although σ_{ah} in different thin films and at different temperatures shows a large variation at finite σ_{xx}, they converge to a single point as σ_{xx} approaches zero, where the impurity-scattering-induced contribution should be washed out by the phonon scattering and only the intrinsic contribution survives. It turns out that this converged value is very close to the bulk σ_{ah} of Fe, which confirms the dominance of the intrinsic contribution in Fe.

In addition to the semiclassical approach (Sinitsyn et al., 2005; Sinitsyn, 2008), there are a number of works based on a full quantum-mechanical approach (Nozières
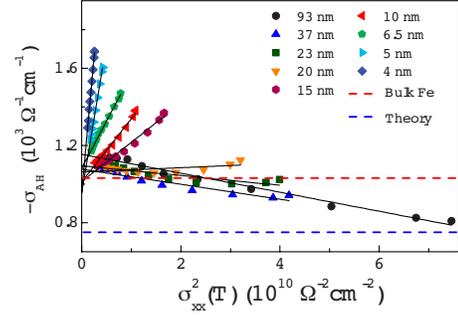

FIG. 7. (Color online) σ_{ah} vs $\sigma_{xx}(T)^2$ in Fe thin films with different film thickness over the temperature range of 5–300 K. From Tian et al., 2009.

and Lewiner, 1973; Onoda and Nagaosa, 2002; Dugaev et al., 2005; Inoue et al., 2006; Onoda, Sugimoto, and Nagaosa, 2006; Kato et al., 2007; Sinitsyn et al., 2007; Onoda et al., 2008). In both approaches, the Berry phase supported intrinsic contribution to the anomalous Hall effect has been firmly established.

2. Anomalous Hall conductivity as a Fermi surface property

An interesting aspect of the intrinsic contribution to the anomalous Hall effect is that the Hall conductivity [Eq. (3.14)] is given as an integration over all occupied states below the Fermi energy. It seems to be against the spirit of the Landau Fermi-liquid theory, which states that the transport property of an electron system is determined by quasiparticles at the Fermi energy. This issue was first raised by Haldane (2004), and he showed that the Hall conductivity can be written, in the units of $e^2/2\pi h$, as the Berry phase of quasiparticles on the Fermi surface, up to a multiple of 2π. Therefore the intrinsic Hall conductivity is also a Fermi surface property. This observation suggests that the Berry phase on the Fermi surface should be added as a topological ingredient to the Landau Fermi-liquid theory.

For simplicity, consider a two-dimensional system. We assume there is only one partially filled band. If we write the Berry curvature in terms of the Berry vector potential and integrate Eq. (3.14) by part, one finds

$$\sigma_{xy}^{2D} = \frac{e^2}{\hbar} \int \frac{d^2\boldsymbol{k}}{2\pi} \left(\frac{\partial f}{\partial k_y} \mathcal{A}_{k_x} - \frac{\partial f}{\partial k_x} \mathcal{A}_{k_y} \right).$$

(3.19)

Note that the Fermi distribution function f is a step function at the Fermi energy. If we assume the Fermi surface is a closed loop in the Brillouin zone, then

$$\sigma_{xy}^{2D} = \frac{e^2}{2\pi h} \oint d\boldsymbol{k} \cdot \mathcal{A}_{\boldsymbol{k}}.$$

(3.20)

The integral is nothing but the Berry phase along the Fermi circle in the Brillouin zone. The three-dimensional case is more complicated; Haldane (2004) showed that the same conclusion can be reached.

Wang et al. (2007) implemented Haldane's idea in first-principles calculations of the anomalous Hall conductivity. From a computational point of view, the ad-

vantage lies in that the integral over the Fermi sea is converted to a more efficient integral on the Fermi surface. On the theory side, Shindou and Balents (2006, 2008) derived an effective Boltzmann equation for quasiparticles on the Fermi surface using the Keldysh formalism, where the Berry phase of the Fermi surface is defined in terms of the quasiparticle Green's functions, which nicely fits into the Landau Fermi-liquid theory.

E. The valley Hall effect

A necessary condition for the charge Hall effect to manifest is the broken time-reversal symmetry of the system. In this section we discuss another type of Hall effect which relies on inversion symmetry breaking and survives in time-reversal invariant systems.

We use graphene as our prototype system. The band structure of intrinsic graphene has two degenerate and inequivalent Dirac points at the corners of the Brillouin zone, where the conduction and valance bands touch each other, forming a valley structure. Because of their large separation in momentum space, the intervalley scattering is strongly suppressed (Morozov $et\ al.$, 2006; Morpurgo and Guinea, 2006; Gorbachev $et\ al.$, 2007), which makes the valley index a good quantum number. Interesting valley-dependent phenomena, which concerns about the detection and generation of valley polarization, are currently being explored (Akhmerov and Beenakker, 2007; Rycerz $et\ al.$, 2007; Xiao, Yao, and Niu, 2007; Yao $et\ al.$, 2008).

The system we are interested in is graphene with broken inversion symmetry. Zhou $et\ al.$ (2007) recently reported the observation of a band-gap opening in epitaxial graphene, attributed to the inversion symmetry breaking by the substrate potential. In addition, in biased graphene bilayer inversion symmetry can be explicitly broken by the applied interlayer voltage (McCann and Fal'ko, 2006; Ohta $et\ al.$, 2006; Min $et\ al.$, 2007). It is this broken inversion symmetry that allows a valley Hall effect. Introducing the valley index $\tau_z = \pm 1$ which labels the two valleys, we can write the valley Hall effect as

$$\boldsymbol{j}^v = \sigma_H^v \hat{z} \times \boldsymbol{E}, \tag{3.21}$$

where σ_H^v is the valley Hall conductivity, and the valley current $\boldsymbol{j}^v = \langle \tau_z \boldsymbol{v} \rangle$ is defined as the average of the valley index τ_z times the velocity operator $\boldsymbol{v}$. Under time reversal, both the valley current and electric field are invariant (τ_z changes sign because the two valleys switch when the crystal momentum changes sign). Under spatial inversion, the valley current is still invariant but the electric field changes sign. Therefore, the valley Hall conductivity can be nonzero when the inversion symmetry is broken even if time-reversal symmetry remains.

In the tight-binding approximation, the Hamiltonian of a single graphene sheet can be modeled with a nearest-neighbor hopping energy t and a site energy difference Δ between sublattices. For relatively low doping, we can resort to the low-energy description near the Dirac points. The Hamiltonian is given by (Semenoff, 1984)

$$H = \frac{\sqrt{3}}{2} at(q_x \tau_z \sigma_x + q_y \sigma_y) + \frac{\Delta}{2} \sigma_z, \tag{3.22}$$

where $\boldsymbol{\sigma}$ is the Pauli matrix accounting for the sublattice index and $\boldsymbol{q}$ is measured from the valley center $\boldsymbol{K}_{1,2} = (\pm 4\pi/3a)\hat{\boldsymbol{x}}$ with a the lattice constant. The Berry curvature of the conduction band is given by

$$\Omega(\boldsymbol{q}) = \tau_z \frac{3a^2 t^2 \Delta}{2(\Delta^2 + 3q^2 a^2 t^2)^{3/2}}. \tag{3.23}$$

Note that the Berry curvatures in two valleys have opposite sign as required by time-reversal symmetry. The energy spectrum and the Berry curvature are shown in Fig. 4. Once the structure of the Berry curvature is revealed, the valley Hall effect becomes transparent: upon the application of an electric field, electrons in different valleys will flow to opposite directions transverse to the electric field, giving rise to a net valley Hall current in the bulk.

We remark that as Δ goes to zero, the Berry curvature vanishes everywhere except at the Dirac points where it diverges. Meanwhile, the Berry phase around the Dirac points becomes exactly $\pm \pi$ (see also Sec. VII.C).

As shown the valley transport in systems with broken inversion symmetry is a very general phenomenon. It provides a new and standard pathway to potential applications of valleytronics or valley-based electronic applications in a broad class of semiconductors (Gunawan $et\ al.$, 2006; Xiao, Yao, and Niu, 2007; Yao $et\ al.$, 2008).

IV. WAVE PACKET: CONSTRUCTION AND PROPERTIES

Our discussion so far has focused on crystals under time-dependent perturbations, and we have shown that the Berry phase will manifest itself as an anomalous term in the electron velocity. However, in general situations the electron dynamics can be also driven by perturbations that vary in space. In this case, the most useful description is provided by the semiclassical theory of Bloch electron dynamics, which describes the motion of a narrow wave packet obtained by superposing the Bloch states of a band [see, for example, Chap. 12 of Ashcroft and Mermin (1976)]. The current and next sections are devoted to the study of the wave-packet dynamics, where the Berry curvature naturally appears in the equations of motion.

In this section we discuss the construction and the general properties of the wave packet. Two quantities emerge in the wave-packet approach, i.e., the orbital magnetic moment of the wave packet and the dipole moment of a physical observable. For their applications, we consider the problems of orbital magnetization and anomalous thermoelectric transport in ferromagnets.

A. Construction of the wave packet and its orbital moment

We assume the perturbations are sufficiently weak that transitions between different bands can be neglected; i.e., the electron dynamics is confined within a single band. Under this assumption, we construct a wave packet using the Bloch functions $|\psi_n(\boldsymbol{q})\rangle$ from the nth band,

$$|W_0\rangle = \int d\boldsymbol{q} w(\boldsymbol{q},t)|\psi_n(\boldsymbol{q})\rangle. \tag{4.1}$$

There are two requirements on the envelope function $w(\boldsymbol{q},t)$. First, $w(\boldsymbol{q},t)$ must have a sharp distribution in the Brillouin zone such that it makes sense to speak of the wave vector $\boldsymbol{q}_c$ of the wave packet given by

$$\boldsymbol{q}_c = \int d\boldsymbol{q}\boldsymbol{q}|w(\boldsymbol{q},t)|^2. \tag{4.2}$$

To first order, $|w(\boldsymbol{q},t)|^2$ can be approximated by $\delta(\boldsymbol{q} - \boldsymbol{q}_c)$ and one has

$$\int d\boldsymbol{q} f(\boldsymbol{q})|w(\boldsymbol{q},t)|^2 \approx f(\boldsymbol{q}_c), \tag{4.3}$$

where $f(\boldsymbol{q})$ is an arbitrary function of $\boldsymbol{q}$. Equation (4.3) is useful in evaluating various quantities related to the wave packet. Second, the wave packet has to be narrowly localized around its center of mass, denoted by $\boldsymbol{r}_c$, in the real space, i.e.,

$$\boldsymbol{r}_c = \langle W_0|\boldsymbol{r}|W_0\rangle. \tag{4.4}$$

Using Eq. (4.3) we obtain

$$\boldsymbol{r}_c = -\frac{\partial}{\partial \boldsymbol{q}_c} \arg w(\boldsymbol{q}_c,t) + \mathcal{A}_{\boldsymbol{q}}^n(\boldsymbol{q}_c), \tag{4.5}$$

where $\mathcal{A}_{\boldsymbol{q}}^n = i\langle u_n(\boldsymbol{q})|\boldsymbol{\nabla}_{\boldsymbol{q}}|u_n(\boldsymbol{q})\rangle$ is the Berry connection of the nth band defined using $|u_n(\boldsymbol{q})\rangle = e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}|\psi_n(\boldsymbol{q})\rangle$. A more rigorous consideration of the wave-packet construction is given by Hagedorn (1980).

In principle, more dynamical variables, such as the width of the wave packet in both the real space and momentum space, can be added to allow a more elaborate description of the time evolution of the wave packet. However, in short period the dynamics is dominated by the motion of the wave-packet center, and Eqs. (4.2) and (4.5) are sufficient requirements.

When more than one band come close to each other or even become degenerate, the single-band approximation breaks down and the wave packet must be constructed using Bloch functions from multiple bands. Culcer et al. (2005) and Shindou and Imura (2005) developed the multiband formalism for electron dynamics, which will be presented in Sec. IX. For now, we focus on the single-band formulation and drop the band index n for simplicity.

The wave packet, unlike a classical point particle, has a finite spread in real space. In fact, since it is constructed using an incomplete basis of the Bloch functions, the size of the wave packet has a nonzero lower bound (Marzari and Vanderbilt, 1997). Therefore, a wave packet may possess a self-rotation around its center of mass, which will in turn give rise to an orbital magnetic moment. By calculating the angular momentum of a wave packet directly, one finds (Chang and Niu, 1996)

$$\boldsymbol{m}(\boldsymbol{q}) = -\frac{e}{2}\langle W_0|(\boldsymbol{r} - \boldsymbol{r}_c) \times \boldsymbol{j}|W_0\rangle$$

$$= -i\frac{e}{2\hbar}\langle\boldsymbol{\nabla}_{\boldsymbol{q}}u| \times [H(\boldsymbol{q}) - \varepsilon(\boldsymbol{q})]|\boldsymbol{\nabla}_{\boldsymbol{q}}u\rangle, \tag{4.6}$$

where $H(\boldsymbol{q}) = e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}He^{i\boldsymbol{q}\cdot\boldsymbol{r}}$ is the $\boldsymbol{q}$-dependent Hamiltonian. Equation (4.3) is used to obtain this result. This shows that the wave packet of a Bloch electron generally rotates around its mass center and carries an orbital magnetic moment in addition to its spin moment.

We emphasize that the orbital moment is an intrinsic property of the band. Its final expression [Eq. (4.6)] does not depend on the actual shape and size of the wave packet and only depends on the Bloch functions. Under symmetry operations, the orbital moment transforms exactly like the Berry curvature. Therefore unless the system has both time-reversal and inversion symmetry, $\boldsymbol{m}(\boldsymbol{q})$ is in general nonzero. Information of the orbital moment can be obtained by measuring magnetic circular dichroism (MCD) spectrum of a crystal (Souza and Vanderbilt, 2008; Yao et al., 2008). In the single-band case, MCD directly measures the magnetic moment.

This orbital moment behaves exactly like the electron spin. For example, upon the application of a magnetic field, the orbital moment will couple to the field through a Zeeman-like term $-\boldsymbol{m}(\boldsymbol{q})\cdot\boldsymbol{B}$. If one can construct a wave packet using only the positive energy states (the electron branch) of the Dirac Hamiltonian, its orbital moment in the nonrelativistic limit is exactly the Bohr magneton (Sec. IX). For Bloch electrons, the orbital moment can be related to the electron g factor (Yafet, 1963). Consider a specific example. For the graphene model with broken-inversion symmetry, discussed in Sec. III.E, the orbital moment of the conduction band is given by (Xiao, Yao, and Niu, 2007)

$$m(\tau_z,\boldsymbol{q}) = \tau_z\frac{3ea^2\Delta t^2}{4\hbar(\Delta^2 + 3q^2a^2t^2)}. \tag{4.7}$$

So orbital moments in different valleys have opposite signs, as required by time-reversal symmetry. Interestingly, the orbital moment at exactly the band bottom takes the following form:

$$m(\tau_z) = \tau_z\mu_B^*, \quad \mu_B^* = \frac{e\hbar}{2m^*}, \tag{4.8}$$

where m^* is the effective mass at the band bottom. The close analogy with the Bohr magneton for the electron spin is transparent. In realistic situations, the moment can be 30 times larger than the spin moment and can be used as an effective way to detect and generate the valley polarization (Xiao, Yao, and Niu, 2007; Yao et al., 2008).
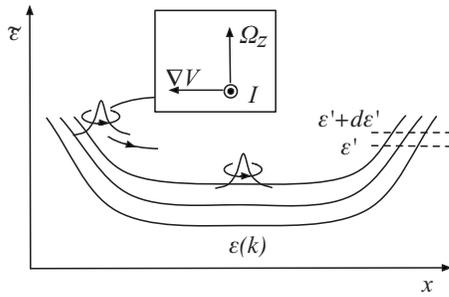
FIG. 8. Electron energy $\tilde{\varepsilon}$ in a slowly varying confining potential $V(r)$. In addition to the self-rotation, wave packets near the boundary will also move along the boundary due to the potential V. Level spacings between different bulk q states are exaggerated; they are continuous in the semiclassical limit. The inset shows directions of the Berry curvature, the effective force, and the current carried by a wave packet on the left boundary.

B. Orbital magnetization

A closely related quantity to the orbital magnetic moment is the orbital magnetization in a crystal. Although this phenomenon has been known for a long time, our understanding of orbital magnetization in crystals has remained in a primitive stage. In fact, there was no proper way to calculate this quantity until recently when the Berry phase theory of orbital magnetization is developed (Thonhauser *et al.*, 2005; Xiao *et al.*, 2005; Shi *et al.*, 2007). Here we provide a pictorial derivation of the orbital magnetization based on the wave-packet approach. This derivation gives an intuitive picture of different contributions to the total orbital magnetization.

The main difficulty of calculating the orbital magnetization is exactly the same as when calculating the electric polarization: the magnetic dipole $er \times p$ is not defined in a periodic system. For a wave packet this is not a problem because it is localized in space. As shown in the previous section, each wave packet carries an orbital moment. Thus, it is tempting to conclude that the orbital magnetization is simply the thermodynamic average of the orbital moment. As it turns out, this is only part of the contribution. There is another contribution due to the center-of-mass motion of the wave packet.

For simplicity, consider a finite system of electrons in a two-dimensional lattice with a confining potential $V(r)$. We further assume that the potential $V(r)$ varies slowly at atomic length scale such that the wave-packet description of the electron is still valid on the boundary. In the bulk where $V(r)$ vanishes, the electron energy is just the bulk band energy; near the boundary, it will be tilted up due to the increase in $V(r)$. Thus to a good approximation, we can write the electron energy as

$$\tilde{\varepsilon}(r,q) = \varepsilon(q) + V(r). \tag{4.9}$$

The energy spectrum in real space is shown in Fig. 8.

Before proceeding further, we need to generalize the velocity formula (3.6), which is derived in the presence of an electric field. In our derivation the electric field

enters through a time-dependent vector potential $A(t)$ so that we can avoid the technical difficulty of calculating the matrix element of the position operator. However, the electric field may be also given by the gradient of the electrostatic potential. In both cases, the velocity formula should stay the same because it should be gauge invariant. Therefore, in general a scalar potential $V(r)$ will induce a transverse velocity of the following form:

$$\frac{1}{\hbar} \boldsymbol{\nabla} V(r) \times \boldsymbol{\Omega}(q). \tag{4.10}$$

This generalization will be justified in Sec. VI.

Now consider a wave packet in the boundary region of the finite system (Fig. 8). It will feel a force $\boldsymbol{\nabla} V(r)$ due to the presence of the confining potential. Consequently, according to Eq. (4.10) the electron acquires a transverse velocity, whose direction is parallel with the boundary (Fig. 8). This transverse velocity will lead to a boundary current (of the dimension current density × width in two dimensions) given by

$$I = \frac{e}{\hbar} \int dx \int \frac{dq}{(2\pi)^2} \frac{dV}{dx} f[\varepsilon(q) + V]\Omega_z(q), \tag{4.11}$$

where x is in the direction perpendicular to the boundary, and the integration range is taken from deep into the bulk to outside the sample. Recall that for a current I flowing in a closed circuit enclosing a sufficiently small area A, the circuit carries a magnetic moment given by IA. Therefore the magnetization (magnetic moment per unit area) has the magnitude of the current I. Integrating Eq. (4.11) by part, we obtain

$$M_f = \frac{1}{e} \int d\varepsilon f(\varepsilon)\sigma_{xy}(\varepsilon), \tag{4.12}$$

where $\sigma_{xy}(\varepsilon)$ is the zero-temperature Hall conductivity for a system with Fermi energy ε:

$$\sigma_{xy}(\varepsilon) = \frac{e^2}{\hbar} \int \frac{dq}{(2\pi)^d}\Theta(\varepsilon - \varepsilon(q))\Omega_z(q). \tag{4.13}$$

Since the boundary current corresponds to the global movement of the wave-packet center, we call this contribution the "free current" contribution, whereas the orbital moment are due to "localized" current. Thus the total magnetization is

$$M_z = \int \frac{dq}{(2\pi)^d} f(q)m_z(q) + \frac{1}{e} \int d\varepsilon f(\varepsilon)\sigma_{xy}(\varepsilon). \tag{4.14}$$

The orbital magnetization has two different contributions: one is from the self-rotation of the wave packet, and the other is due to the center-of-mass motion. Gat and Avron obtained an equivalent result for the special case of the Hofstadter model (Gat and Avron, 2003a, 2003b).

The above derivation relies on the existence of a confining potential, which seems to contradict the fact that the orbital magnetization is a bulk property. This is a wrong assertion as the final expression [Eq. (4.14)] is given in terms of the bulk Bloch functions and does not
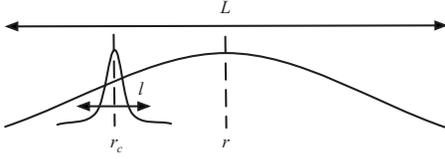
FIG. 9. Sampling function and a wave packet at r_c. The width L of the sampling function is sufficiently small so that it can be treated as a δ function at the macroscopic level and is sufficiently large so that it contains a large number of wave packets of width l inside its range. Equation (4.15) is indeed a microscopic average over the distance L around the point r. See Sec. 6.6 in Jackson (1998) for an analogy in macroscopic electromagnetism.

depend on the boundary condition. Here the boundary is merely a tool to expose the "free current" contribution because in a uniform system, the magnetization current always vanishes in the bulk. Finally, in more rigorous approaches (Xiao *et al.*, 2005; Shi *et al.*, 2007) the boundary is not needed and the derivation is based on a pure bulk picture. It is similar to the quantum Hall effect, which can be understood in terms of either the bulk states (Thouless *et al.*, 1982) or the edge states (Halperin, 1982).

C. Dipole moment

The finite size of the wave packet not only allows an orbital magnetic moment but also leads to the concept of the dipole moment associated with an operator.

The dipole moment appears naturally when we consider the thermodynamic average of a physical quantity, with its operator denoted by $\hat{\mathcal{O}}$. In the wave-packet approach, the operator is given by

$$\mathcal{O}(r) = \int \frac{dr_c dq_c}{(2\pi)^3} g(r_c, q_c) \langle W | \hat{\mathcal{O}} \delta(r - \hat{r}) | W \rangle, \qquad (4.15)$$

where $g(r, q)$ is the distribution function, $\langle W | \cdots | W \rangle$ denotes the expectation in the wave-packet state, and $\delta(r - \hat{r})$ plays the role as a sampling function, as shown in Fig. 9. An intuitive way to view Eq. (4.15) is to think of the wave packets as small molecules, then Eq. (4.15) is the quantum-mechanical version of the familiar coarse graining process which averages over the length scale larger than the size of the wave packet. A multipole expansion can be carried out. But for most purposes the dipole term is enough. Expand the δ function to first order of $\hat{r} - r_c$:

$$\delta(r - \hat{r}) = \delta(r - r_c) - (\hat{r} - r_c) \cdot \nabla \delta(r - r_c). \qquad (4.16)$$
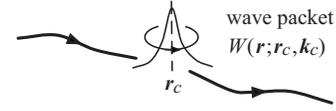
Inserting the function into Eq. (4.15) yields

wave packet $W(r; r_c, k_c)$

FIG. 10. The wave-packet description of a charge carrier whose center is (r_c, q_c). A wave packet generally possesses two kinds of motion: the center-of-mass motion and the self-rotation around its center. From Xiao, Yao, *et al.*, 2006.

$$\mathcal{O}(r) = \int \frac{dq}{(2\pi)^3} g(r, q) \langle W | \hat{\mathcal{O}} | W \rangle |_{r_c = r}$$
$$- \nabla \cdot \int \frac{dq}{(2\pi)^3} g(r, q) \langle W | \hat{\mathcal{O}}(\hat{r} - r_c) | W \rangle |_{r_c = r}. \qquad (4.17)$$

The first term is obtained if the wave packet is treated as a point particle. The second term is due to the finite size of the wave packet. We can see that the bracket in the second integral has the form of a dipole of the operator $\mathcal{O}$ defined by

$$P_{\hat{\mathcal{O}}} = \langle W | \hat{\mathcal{O}}(\hat{r} - r_c) | W \rangle. \qquad (4.18)$$

The dipole moment of an observable is a general consequence of the wave-packet approach and must be included in the semiclassical theory. Its contribution appears only when the system is inhomogeneous.

In particular, we find the following:

(1) If $\hat{\mathcal{O}} = e$, then $P_e = 0$. This is consistent with the fact that the charge center coincides with the mass center of the electron.

(2) If $\hat{\mathcal{O}} = \hat{v}$, one finds the expression for the local current,

$$j^L = \int \frac{dq}{(2\pi)^3} g(r, q) \dot{r} + \nabla \times \int \frac{dq}{(2\pi)^3} g(r, q) m(q). \qquad (4.19)$$

We explain the meaning of local later. Interestingly, this is the second time we encounter the quantity $m(q)$ but in an entirely different context. The physical meaning of the second term becomes transparent if we make reference to the self-rotation of the wave packet. The self-rotation can be thought as localized circuit. Therefore if the distribution is not uniform, the localized circuit will contribute to the local current j^L (see Fig. 10).

(3) If $\hat{\mathcal{O}}$ is the spin operator $\hat{s}$, then Eq. (4.18) gives the spin dipole

$$P_s = \langle u | s \left(i \frac{\partial}{\partial q} - \mathcal{A}_q \right) | u \rangle. \qquad (4.20)$$

The spin dipole shows that in general the spin center and the mass center do not coincide, which is usually

due to the spin-orbit interaction. The time derivative of the spin dipole contributes to the total spin current (Culcer *et al.*, 2004).

D. Anomalous thermoelectric transport

In applying the above concepts, we consider the problem of anomalous thermoelectric transport in ferromagnets, which refers to the appearance of a Hall current driven by statistical forces, such as the gradient of temperature and chemical potential (Chien and Westgate, 1980). Similar to the anomalous Hall effect, there are also intrinsic and extrinsic contributions, and we focus on the former.

A question immediately arises when one tries to formulate this problem. Recall that in the presence of an electric field the electron acquires an anomalous velocity proportional to the Berry curvature, which gives rise to a Hall current. In this case, the driving force is of mechanical nature: it exists on the microscopic level and can be described by a perturbation to the Hamiltonian for the carriers. On the other hand, transport can be also driven by the statistical force. However, the statistical force manifests on the macroscopic level and makes sense only through the statistical distribution of the carriers. Since there is no force acting directly on individual particles, the obvious cause for the Berry phase assisted transport is eliminated. This conclusion would introduce a number of basic contradictions to the standard transport theory. First, a chemical potential gradient would be distinct from the electric force, violating the basis for the Einstein relation. Second, a temperature gradient would not induce an intrinsic charge Hall current, violating the Mott relation. Finally, it is also unclear whether the Onsager relation is satisfied or not.

It turns out the correct description of anomalous thermoelectric transport in ferromagnets requires knowledge of both the magnetic moment and orbital magnetization. First, as shown in Eq. (4.19), the local current is given by

$$j^L = \int \frac{d\boldsymbol{q}}{(2\pi)^d} g(\boldsymbol{r}, \boldsymbol{q}) \dot{\boldsymbol{r}} + \boldsymbol{\nabla} \times \int \frac{d\boldsymbol{q}}{(2\pi)^d} f(\boldsymbol{r}, \boldsymbol{q}) \boldsymbol{m}(\boldsymbol{q}),$$
(4.21)

where in the second term we have replaced the distribution function $g(\boldsymbol{r}, \boldsymbol{q})$ with the local Fermi-Dirac function $f(\boldsymbol{r}, \boldsymbol{q})$, which is sufficient for a first order calculation. Second, in ferromagnetic systems, it is important to discount the contribution from the magnetization current. It was argued that the magnetization current cannot be measured by conventional transport experiments (Cooper *et al.*, 1997). Therefore the transport current is given by

$$j = j^L - \boldsymbol{\nabla} \times M(\boldsymbol{r}).$$
(4.22)

Using Eq. (4.14), one finds

$$j = \int \frac{d\boldsymbol{q}}{(2\pi)^d} g(\boldsymbol{r}, \boldsymbol{q}) \dot{\boldsymbol{r}} - \frac{1}{e} \boldsymbol{\nabla} \times \int d\varepsilon f(\varepsilon) \boldsymbol{\sigma}^{\mathrm{AH}}(\varepsilon).$$
(4.23)

Equation (4.23) is the most general expression for the transport current. We notice that the contribution from the orbital magnetic moment $\boldsymbol{m}(\boldsymbol{q})$ cancels out. This agrees with the intuitive picture developed in Sec. IV.B, i.e., the orbital moment is due to the self-rotation of the wave packet, therefore it is localized and cannot contribute to transport (see Fig. 10).

In the presence of a statistical force, there are two ways for a Hall current to occur. The asymmetric scattering will have an effect on the distribution $g(\boldsymbol{r}, \boldsymbol{q})$, which is obtained from the Boltzmann equation (Berger, 1972). This results in a transverse current in the first term of Eq. (4.23). In addition, there is an intrinsic contribution comes from the orbital magnetization, which is the second term of Eq. (4.23). Note that the spatial dependence enters through $T(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$ in the distribution function. It is straightforward to verify that for the intrinsic contribution to the anomalous thermoelectric transport both the Einstein relation and Mott relation still hold (Xiao, Yao, *et al.*, 2006; Onoda *et al.*, 2008). Hence, the measurement of this type of transport, such as the anomalous Nernst effect, can give further insight into the intrinsic mechanism of the anomalous Hall effect. Much experimental efforts have been put along this line. The intrinsic contribution has been verified in $CuCr_2Se_{4-x}Br_x$ (Lee *et al.*, 2004a, 2004b), $La_{1-x}Sr_xCoO_3$ (Miyasato *et al.*, 2007), $Nd_2Mo_2O_7$ and $Sm_2Mo_2O_7$ (Hanasaki *et al.*, 2008), and $Ga_{1-x}Mn_x$ (Pu *et al.*, 2008).

Equation (4.23) is not limited to transport driven by statistical forces. As shown later, at the microscopic level the mechanical force generally has two effects: it can drive the electron motion directly and appears in the expression for $\dot{\boldsymbol{r}}$; it can also make the electron energy and the Berry curvature spatially dependent, hence also manifest in the second term in Eq. (4.23). The latter provides another route for the Berry phase to enter the transport problems in inhomogeneous situations, which can be caused by a nonuniform distribution function or a spatially dependent perturbation, or both. One example is the electrochemical potential $-e\phi(\boldsymbol{r}) + \mu(\boldsymbol{r})/e$, which can induce an anomalous velocity term in the equation of motion through $-e\phi(\boldsymbol{r})$ and also affect the distribution function through $\mu(\boldsymbol{r})$.

V. ELECTRON DYNAMICS IN ELECTROMAGNETIC FIELDS

In the previous section we discussed the construction and general properties of a wave packet. Now we are set to study its dynamics under external perturbations. The most common perturbations to a crystal are the electromagnetic fields. The study of the electron dynamics under such perturbations dates back to Bloch, Peierls, Jones, and Zener in the early 1930s and is continued by Slater (1949), Luttinger (1951), Adams (1952), Karplus and Luttinger (1954), Kohn and Luttinger (1957), Adams and Blount (1959), Blount (1962a), Brown (1968), Zak (1977), Rammal and Bellissard (1990), and Wilkinson and Kay (1996). In this section we present the semiclas-

sical theory based on the wave-packet approach (Chang and Niu, 1995, 1996).

A. Equations of motion

In the presence of electromagnetic fields, the Hamiltonian is given by

$$H = \frac{[\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r})]^2}{2m} + V(\boldsymbol{r}) - e\phi(\boldsymbol{r}), \tag{5.1}$$

where $V(\boldsymbol{r})$ is the periodic lattice potential and $\boldsymbol{A}(\boldsymbol{r})$ and $\phi(\boldsymbol{r})$ are the electromagnetic potentials. If the length scale of the perturbations is much larger than the spatial spread of the wave packet, the approximate Hamiltonian that the wave packet "feels" may be obtained by linearizing the perturbations about the wave-packet center $\boldsymbol{r}_c$ as

$$H \approx H_c + \Delta H, \tag{5.2}$$

$$H_c = \frac{[\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r}_c)]^2}{2m} + V(\boldsymbol{r}) - e\phi(\boldsymbol{r}_c), \tag{5.3}$$

$$\Delta H = \frac{e}{2m}\{\boldsymbol{A}(\boldsymbol{r}) - \boldsymbol{A}(\boldsymbol{r}_c), \boldsymbol{p}\} - e\boldsymbol{E} \cdot (\boldsymbol{r} - \boldsymbol{r}_c), \tag{5.4}$$

where $\{,\}$ is the anticommutator. Naturally, we can then construct the wave packet using the eigenstates of the local Hamiltonian H_c. The effect of a uniform $\boldsymbol{A}(\boldsymbol{r}_c)$ is to add a phase to the eigenstates of the unperturbed Hamiltonian. Therefore the wave packet can be written as

$$|W(\boldsymbol{k}_c, \boldsymbol{r}_c)\rangle = e^{-(ie/\hbar)\boldsymbol{A}(\boldsymbol{r}_c)\cdot\boldsymbol{r}}|W_0(\boldsymbol{k}_c, \boldsymbol{r}_c)\rangle, \tag{5.5}$$

where $|W_0\rangle$ is the wave packet constructed using the unperturbed Bloch functions.

The wave-packet dynamics can be obtained from the time-dependent variational principles (Kramer and Saraceno, 1981). The basic recipe is to first obtain the Lagrangian from the following equation:

$$L = \langle W|i\hbar\frac{\partial}{\partial t} - H|W\rangle, \tag{5.6}$$

then obtain the equations of motion using the Euler equations. Using Eq. (4.5) we find that $\langle W|i\hbar\partial/\partial t|W\rangle = e\dot{\boldsymbol{A}}\cdot\boldsymbol{R}_c - \hbar(\partial/\partial t)\text{arg } w(\boldsymbol{k}_c, t)$. For the wave-packet energy, we have $\langle W|\Delta H|W\rangle = -\boldsymbol{m}(\boldsymbol{k})\cdot\boldsymbol{B}$. This is expected as we already showed that the wave packet carries an orbital magnetic moment $\boldsymbol{m}(\boldsymbol{k})$ that will couple to the magnetic field. Using Eq. (4.5), we find that the Lagrangian is given by, up to some unimportant total time-derivative terms (dropping the subscript c on $\boldsymbol{r}_c$ and $\boldsymbol{k}_c$),

$$L = \hbar\boldsymbol{k}\cdot\dot{\boldsymbol{r}} - \varepsilon_M(\boldsymbol{k}) + e\phi(\boldsymbol{r}) - e\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r}, t) + \hbar\dot{\boldsymbol{k}}\cdot\mathcal{A}_n(\boldsymbol{k}), \tag{5.7}$$

where $\varepsilon_M(\boldsymbol{k}) = \varepsilon_0(\boldsymbol{k}) - \boldsymbol{B}\cdot\boldsymbol{m}(\boldsymbol{k})$ with $\varepsilon_0(\boldsymbol{k})$ the unperturbed band energy. The equations of motion are given by

$$\dot{\boldsymbol{r}} = \frac{\partial\varepsilon_M(\boldsymbol{k})}{\hbar\partial\boldsymbol{k}} - \dot{\boldsymbol{k}}\times\boldsymbol{\Omega}(\boldsymbol{k}), \tag{5.8a}$$

$$\hbar\dot{\boldsymbol{k}} = -e\boldsymbol{E} - e\dot{\boldsymbol{r}}\times\boldsymbol{B}. \tag{5.8b}$$

Compared to the conventional equations of motion for Bloch electrons (Ashcroft and Mermin, 1976), there are two differences: (1) the electron energy is modified by the orbital magnetic moment and (2) the electron velocity gains an extra velocity term proportional to the Berry curvature. As shown, in the case of only an electric field, Eq. (5.8a) reduces to the anomalous velocity formula (3.6) we derived before.

B. Modified density of states

The Berry curvature not only modifies the electron dynamics but also has a profound effect on the electron density of states in the phase space (Xiao *et al.*, 2005).[10]

Recall that in solid-state physics the expectation value of an observable, in the Bloch representation, is given by

$$\sum_{n\boldsymbol{k}} f_{n\boldsymbol{k}}\langle\psi_{n\boldsymbol{k}}|\hat{O}|\psi_{n\boldsymbol{k}}\rangle, \tag{5.9}$$

where $f_{n\boldsymbol{k}}$ is the distribution function. In the semiclassical limit, the sum is converted to an integral in $\boldsymbol{k}$ space,

$$\sum_{\boldsymbol{k}} \rightarrow \frac{1}{V}\int\frac{d\boldsymbol{k}}{(2\pi)^d}, \tag{5.10}$$

where V is the volume and $(2\pi)^d$ is the density of states, i.e., number of states per unit $\boldsymbol{k}$ volume. From a classical point of view, the constant density of states is guaranteed by the Liouville theorem, which states that the volume element is a conserved quantity during the time evolution of the system.[11] However, as shown below, this is no longer the case for the Berry phase modified dynamics.

The time evolution of a volume element $\Delta V = \Delta\boldsymbol{r}\Delta\boldsymbol{k}$ is given by

$$\frac{1}{\Delta V}\frac{\partial\Delta V}{\partial t} = \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} + \boldsymbol{\nabla}_{\boldsymbol{k}}\cdot\dot{\boldsymbol{k}}. \tag{5.11}$$

Insert the equations of motion (5.8) into Eq. (5.11). After some algebra, we find

---

[10]The phase space is spanned by $\boldsymbol{k}$ and $\boldsymbol{r}$. Although the Lagrangian depends on both $(\boldsymbol{k}, \boldsymbol{r})$ and their velocities, the dependence on the latter is only to linear order. This means that the momenta, defined as the derivative of the Lagrangian with respect to these velocities, are functions of $(\boldsymbol{k}, \boldsymbol{r})$ only and are independent of their velocities. Therefore, $(\boldsymbol{k}, \boldsymbol{r})$ also span the phase space of the Hamiltonian dynamics.

[11]The actual value of this constant volume for a quantum state, however, can be determined only from the quantization conditions in quantum mechanics.

$$\Delta V = \frac{V_0}{1 + (e/\hbar)\boldsymbol{B}\cdot\boldsymbol{\Omega}}. \tag{5.12}$$

The fact that the Berry curvature is generally $\boldsymbol{k}$ dependent and the magnetic field is $\boldsymbol{r}$ dependent implies that the phase-space volume ΔV changes during time evolution of the state variables $(\boldsymbol{r}, \boldsymbol{k})$.

Although the phase-space volume is no longer conserved, it is a local function of the state variables and has nothing to do with the history of time evolution. We can thus introduce a modified density of states

$$D(\boldsymbol{r}, \boldsymbol{k}) = \frac{1}{(2\pi)^d}\left(1 + \frac{e}{\hbar}\boldsymbol{B}\cdot\boldsymbol{\Omega}\right) \tag{5.13}$$

such that the number of states in the volume element $D_n(\boldsymbol{r}, \boldsymbol{k})\Delta V$ remains constant in time. Therefore, the correct semiclassical limit of the sum in Eq. (5.9) is

$$O(\boldsymbol{R}) = \int d\boldsymbol{k} D(\boldsymbol{r}, \boldsymbol{k})\langle O\,\delta(\hat{\boldsymbol{r}} - \boldsymbol{R})\rangle_W, \tag{5.14}$$

where $\langle\cdots\rangle_W$ is the expectation value in a wave packet, which could includes the dipole contribution due to the finite size of the wave packet [see Eq. (4.18)]. In a uniform system it is given by

$$O = \int d\boldsymbol{k} D(\boldsymbol{k})f(\boldsymbol{k})O(\boldsymbol{k}). \tag{5.15}$$

We emphasize that although the density of states is no longer a constant, the dynamics itself is still Hamiltonian. The modification comes from the fact that the dynamical variables $\boldsymbol{r}$ and $\boldsymbol{k}$ are no longer canonical variables, and the density of states can be regarded as the phase-space measure (Bliokh, 2006b; Duval *et al.*, 2006a, 2006b; Xiao, Shi, and Niu, 2006). The phase-space measure $d\boldsymbol{k}d\boldsymbol{r}$ is true only when $\boldsymbol{k}$ and $\boldsymbol{r}$ form a canonical set. However, the phase-space variables obtained from the wave packet are generally not canonical as testified by their equations of motion. A more profound reason for this modification has its quantum-mechanical origin in noncommutative quantum mechanics, discussed in Sec. VII.

In the following we discuss two direct applications of the modified density of states in metals and in insulators.

1. Fermi volume

We show that the Fermi volume can be changed linearly by a magnetic field when the Berry curvature is nonzero. Assume zero temperature, the electron density is given by

$$n_e = \int \frac{d\boldsymbol{k}}{(2\pi)^d}\left(1 + \frac{e}{\hbar}\boldsymbol{B}\cdot\boldsymbol{\Omega}\right)\Theta(\varepsilon_F - \varepsilon). \tag{5.16}$$

We work in the canonical ensemble by requiring the electron number fixed, therefore, to first order of $\boldsymbol{B}$, the Fermi volume must be changed by

$$\delta V_F = -\int \frac{d\boldsymbol{k}}{(2\pi)^d}\frac{e}{\hbar}\boldsymbol{B}\cdot\boldsymbol{\Omega}. \tag{5.17}$$

It is particularly interesting to look at insulators, where the integration is limited to the Brillouin zone. Then the electron must populate a higher band if $\int_{\text{BZ}}d\boldsymbol{k}\boldsymbol{B}\cdot\boldsymbol{\Omega}$ is negative. When this quantity is positive, holes must appear at the top of the valance bands. Discontinuous behavior of physical properties in a magnetic field is therefore expected for band insulators with a nonzero integral of the Berry curvatures (Chern numbers).

2. Streda formula

In the context of the quantum Hall effect, Streda (1982) derived a formula relating the Hall conductivity to the field derivative of the electron density at a fixed chemical potential

$$\sigma_{xy} = -e\left(\frac{\partial n_e}{\partial B_z}\right)_\mu. \tag{5.18}$$

There is a simple justification for this relation given by a thermodynamic argument by considering the following adiabatic process in two dimensions. A time-dependent magnetic flux generates an electric field with an emf around the boundary of some region, and the Hall current leads to a net flow of electrons across the boundary and thus a change in electron density inside. Note that this argument is valid only for insulators because in metals the adiabaticity would break down. Using Eq. (5.16) for an insulator, we obtain, in two dimensions,

$$\sigma_{xy} = -\frac{e^2}{\hbar}\int_{\text{BZ}}\frac{d\boldsymbol{k}}{(2\pi)^2}\Omega_{xy}. \tag{5.19}$$

This is what Thouless *et al.* (1982) obtained using the Kubo formula. The fact that the quantum Hall conductivity can be derived using the modified density of states further confirms the necessity to introduce this concept.

C. Orbital magnetization: Revisit

We have discussed the orbital magnetization using a rather pictorial derivation in Sec. IV.B. Here we derive the formula again using the field-dependent density of states [Eq. (5.13)].

The equilibrium magnetization density can be obtained from the grand canonical potential, which, within first order in the magnetic field, may be written as

$$F = -\frac{1}{\beta}\sum_{\boldsymbol{k}}\ln(1 + e^{-\beta(\varepsilon_M - \mu)})$$

$$= -\frac{1}{\beta}\int \frac{d\boldsymbol{k}}{(2\pi)^d}\left(1 + \frac{e}{\hbar}\boldsymbol{B}\cdot\boldsymbol{\Omega}\right)\ln(1 + e^{-\beta(\varepsilon_M - \mu)}),$$

$$\tag{5.20}$$

where the electron energy $\varepsilon_M = \varepsilon(\boldsymbol{k}) - \boldsymbol{m}(\boldsymbol{k})\cdot\boldsymbol{B}$ includes a correction due to the orbital magnetic moment $\boldsymbol{m}(\boldsymbol{k})$. The magnetization is then the field derivative at fixed temperature and chemical potential, $\boldsymbol{M} = -(\partial F/\partial \boldsymbol{B})_{\mu,T}$, with the result

$$M(r) = \int \frac{dk}{(2\pi)^d} f(k) m(k)$$
$$+ \frac{1}{\beta} \int \frac{dk}{(2\pi)^d} \frac{e}{\hbar} \Omega(k) \ln(1 + e^{-\beta(\varepsilon - \mu)}). \qquad (5.21)$$

Integration by parts of the second term will give us the exact formula obtained in Eq. (4.14). We have thus derived a general expression for the equilibrium orbital magnetization density, valid at zero magnetic field but at arbitrary temperatures. From this derivation we can clearly see that the orbital magnetization is indeed a bulk property. The center-of-mass contribution identified before comes from the Berry phase correction to the electron density of states.

Following the discussions on band insulators in our first example in Sec. V.B.1, there will be a discontinuity of the orbital magnetization if the integral of the Berry curvature over the Brillouin zone or the anomalous Hall conductivity is nonzero and quantized. Depending on the direction of the field, the chemical potential μ_0 in Eq. (5.21) should be taken at the top of the valence bands or the bottom of the conduction bands. The size of the discontinuity is given by the quantized anomalous Hall conductivity times E_g/e, where E_g is the energy gap.

A similar formula for insulators with zero Chern number has been obtained by Thonhauser *et al.* (2005) and Ceresoli *et al.* (2006) using the Wannier function approach and by Gat and Avron (2003a, 2003b) for the special case of the Hofstadter model. Recently Shi *et al.* (2007) provided a full quantum-mechanical derivation of the formula, and showed that it is valid in the presence of electron-electron interaction, provided the one-electron energies and wave functions are calculated self-consistently within the framework of the exact current and spin-density functional theory (Vignale and Rasolt, 1988).

The appearance of the Hall conductivity is not a coincidence. Consider an insulator. The free energy is given by

$$dF = -M dB - n d\mu - S dT. \qquad (5.22)$$

Using the Maxwell relation, we have

$$\sigma_H = -e \left(\frac{\partial n}{\partial B} \right)_{\mu,T} = -e \left(\frac{\partial M}{\partial \mu} \right)_{B,T}. \qquad (5.23)$$

On the other hand, the zero-temperature formula of the magnetization for an insulator is given by

$$M = \int_{BZ} \frac{dk}{(2\pi)^3} \left\{ m(k) + \frac{e}{\hbar} (\mu - \varepsilon) \Omega \right\}. \qquad (5.24)$$

Inserting this formula into Eq. (5.23) gives us once again the quantized Hall conductivity.

D. Magnetotransport

The equations of motion (5.8) and the density of states [Eq. (5.13)] gives us a complete description of the elec-

tron dynamics in the presence of electromagnetic fields. In this section we apply these results to the problem of magnetotransport. For simplicity, we set $e = \hbar = 1$ and introduce the shorthand $[dk] = dk/(2\pi)^d$.

1. Cyclotron period

Semiclassical motion of a Bloch electron in a uniform magnetic field is important to understand various magnetoeffects in solids. In this case, the equations of motion reduce to

$$\bar{D}(k) \dot{r} = v + (v \cdot \Omega) B, \qquad (5.25a)$$

$$\bar{D}(k) \dot{k} = -v \times B, \qquad (5.25b)$$

where $\bar{D}(k) = D(k)/(2\pi)^d = 1 + (e/\hbar) B \cdot \Omega$.

We assume the field is along the z axis. From Eq. (5.25b) we can see that motion in k space is confined in the x-y plane and is completely determined once the energy ε and the z component of the wave vector k_z is given. We now calculate the period of the cyclotron motion. The time for the wave vector to move from k_1 to k_2 is

$$t_2 - t_1 = \int_{t_1}^{t_2} dt = \int_{k_1}^{k_2} \frac{dk}{|\dot{k}|}. \qquad (5.26)$$

From the equations of motion (5.25) we have

$$|\dot{k}| = \frac{B|v_\perp|}{\bar{D}(k)} = \frac{B|(\partial \varepsilon / \partial k)_\perp|}{\hbar \bar{D}(k)}. \qquad (5.27)$$

On the other hand, the quantity $(\partial \varepsilon / \partial k)_\perp$ can be written as $\Delta \varepsilon / \Delta k$, where Δk denotes the vector in the plane connecting points on neighboring orbits of energy ε and $\varepsilon + \Delta \varepsilon$. Then

$$t_2 - t_1 = \frac{\hbar}{B} \int_{k_1}^{k_2} \frac{\bar{D}(k) \Delta k \, dk}{\Delta \varepsilon}. \qquad (5.28)$$

Introducing the 2D electron density for given ε and k_z,

$$n_2(\varepsilon, k_z) = \int \int_{k_z, \varepsilon(k) < \varepsilon} \frac{\bar{D}(k) dk_x dk_y}{(2\pi)^2}, \qquad (5.29)$$

the period of a cyclotron motion can be written as

$$T = (2\pi)^2 \frac{\hbar}{B} \frac{\partial n_2(\varepsilon, k_z)}{\partial \varepsilon}. \qquad (5.30)$$

We thus recovered the usual expression for the cyclotron period, with the 2D electron density [Eq. (5.29)] defined with the modified density of states.

In addition, we note that there is a velocity term proportional to B in Eq. (5.25), which seems to suggest there will be a current along the field direction. We show that after averaging over the distribution function, this current is actually zero. The current along B is given by

$$\boldsymbol{j}_B = - e\boldsymbol{B} \int [d\boldsymbol{k}] f\boldsymbol{v} \cdot \boldsymbol{\Omega} = -\frac{e}{\hbar}\boldsymbol{B} \int [d\boldsymbol{k}] \boldsymbol{\nabla}_{\boldsymbol{k}} F \cdot \boldsymbol{\Omega}$$

$$= -\frac{e}{\hbar}\boldsymbol{B}\left(\int [d\boldsymbol{k}] \boldsymbol{\nabla}_{\boldsymbol{k}}(F\boldsymbol{\Omega}) - \int [d\boldsymbol{k}] F \boldsymbol{\nabla}_{\boldsymbol{k}} \cdot \boldsymbol{\Omega} \right), \quad (5.31)$$

where $F(\varepsilon) = -\int_\varepsilon^\infty f(\varepsilon') d\varepsilon'$ and $f(\varepsilon) = \partial F/\partial \varepsilon$. The first term vanishes[12] and if there is no magnetic monopole in $\boldsymbol{k}$ space, the second term also vanishes. In the above calculation we did not consider the change in the Fermi surface. Since it always comes in the form $(\partial f/\partial \mu)\delta\mu = -(\partial f/\partial \varepsilon)\delta\mu$, we can use the same technique to prove that the corresponding current also vanishes.

2. The high-field limit

We now consider the magnetotransport at the so-called high-field limit, i.e., $\omega_c \tau \gg 1$, where $\omega_c = 2\pi/T$ is the cyclotron frequency and τ is the relaxation time. We consider a configuration where the electric and magnetic fields are perpendicular to each other, i.e., $\boldsymbol{E} = E\hat{\boldsymbol{x}}$, $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, and $\boldsymbol{E} \cdot \boldsymbol{B} = 0$.

In the high-field limit, $\omega_c \tau \gg 1$, the electron can finish several turns between two successive collisions. We can then assume all orbits are closed. According to the theorem of adiabatic drifting (Niu and Sundaram, 2001), an originally closed orbit remains closed for weak perturbations so that

$$0 = \langle \dot{\boldsymbol{k}} \rangle = \boldsymbol{E} + \langle \dot{\boldsymbol{r}} \rangle \times \boldsymbol{B} \qquad (5.32)$$

or

$$\langle \dot{\boldsymbol{r}} \rangle_\perp = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2}. \qquad (5.33)$$

The Hall current is simply the sum over $\langle \dot{\boldsymbol{r}} \rangle_\perp$ of occupied states,

$$\boldsymbol{j}_H = -e\frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} \int [d\boldsymbol{k}] f(\boldsymbol{k})(1 + \boldsymbol{B} \cdot \boldsymbol{\Omega})$$

$$= -e\frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} \int [d\boldsymbol{k}] f(\boldsymbol{k}) \bar{D}(\boldsymbol{k}). \qquad (5.34)$$

Therefore in the high-field limit we reach the following conclusion: the total current in crossed electric and magnetic fields is the Hall current as if calculated from free electron model

$$\boldsymbol{j} = -e\frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} n, \qquad (5.35)$$

and it has no dependence on the relaxation time τ. This result ensures that even in the presence of anomalous

---

[12]For any periodic function $F(\boldsymbol{k})$ with the periodicity of a reciprocal Bravais lattice, the following identity holds for integrals taken over a Brillouin zone, $\int_{BZ} d\boldsymbol{k} \boldsymbol{\nabla}_{\boldsymbol{k}} F(\boldsymbol{k}) = 0$. To see this, consider $I(\boldsymbol{k}') = \int d\boldsymbol{k} F(\boldsymbol{k} + \boldsymbol{k}')$. Because $F(\boldsymbol{k})$ is periodic in $\boldsymbol{k}$, $I(\boldsymbol{k}')$ should not depend on $\boldsymbol{k}'$. Therefore, $\boldsymbol{\nabla}_{\boldsymbol{k}'} I(\boldsymbol{k}') = \int d\boldsymbol{k}, \boldsymbol{\nabla}_{\boldsymbol{k}'} F(\boldsymbol{k} + \boldsymbol{k}') = \int d\boldsymbol{k} \boldsymbol{\nabla}_{\boldsymbol{k}} F(\boldsymbol{k} + \boldsymbol{k}') = 0$. Setting $\boldsymbol{k}' = 0$ gives the desired expression. This is also true if $F(\boldsymbol{k})$ is a vector function.

Hall effect, the high-field Hall current gives the "real" electron density.

We now consider the holelike band. The Hall current is obtained by subtracting the contribution of holes from that of the filled band, which is given by $-e\boldsymbol{E} \times \int [d\boldsymbol{k}]\boldsymbol{\Omega}$. Therefore

$$\boldsymbol{j}^{\text{hole}} = e\frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} \int [d\boldsymbol{k}] \bar{D}(\boldsymbol{k})[1 - f(\boldsymbol{k})]$$

$$- e\boldsymbol{E} \times \int [d\boldsymbol{k}]\boldsymbol{\Omega}. \qquad (5.36)$$

As a result for the holelike band there is an additional term in the current expression proportional to the Chern number (the second integral) of the band.

3. The low-field limit

Next we consider the magnetotransport at the low-field limit, i.e., $\omega_c \tau \ll 1$. In particular, we show that the Berry phase induces a linear magnetoresistance. By solving the Boltzmann equation, one finds that the diagonal element of the conductivity is given by

$$\sigma_{xx} = -e^2 \int [d\boldsymbol{k}] \tau \frac{\partial f_0}{\partial \varepsilon} \frac{v_x^2}{\bar{D}(\boldsymbol{k})}. \qquad (5.37)$$

This is just the zeroth-order expansion based on $\omega_c \tau$. There are four places in this expression depending on B. (1) There is an explicit B dependence in $\bar{D}(\boldsymbol{k})$. (2) The electron velocity v_x is modified by the orbital magnetic moment:

$$v_x = \frac{1}{\hbar}\frac{\partial(\varepsilon_0 - m_z B)}{\partial k_x} = v_x^{(0)} - \frac{1}{\hbar}\frac{\partial m_z}{\partial k_x}B. \qquad (5.38)$$

(3) There is also a modification to the Fermi energy, given by Eq. (5.17). (4) The relaxation time τ can also depend on B. In the presence of the Berry curvature, the collision term in the Boltzmann equation is given by

$$\frac{\partial f}{\partial t}\bigg|_{\text{coll}} = -\int [d\boldsymbol{k}'] \bar{D}(\boldsymbol{k}') W_{\boldsymbol{k}\boldsymbol{k}'}[f(\boldsymbol{k}) - f(\boldsymbol{k}')], \qquad (5.39)$$

where $W_{\boldsymbol{k}\boldsymbol{k}'}$ is the transition probability from $\boldsymbol{k}'$ to $\boldsymbol{k}$ state. In the relaxation-time approximation we make the assumption that a characteristic relaxation time exists so that

$$\frac{f - f_0}{\tau} = \bar{D}(\boldsymbol{k}) \int [d\boldsymbol{k}'] \frac{\bar{D}(\boldsymbol{k}')}{\bar{D}(\boldsymbol{k})} W_{\boldsymbol{k}\boldsymbol{k}'}[f(\boldsymbol{k}) - f(\boldsymbol{k}')]. \quad (5.40)$$

If we assume $\boldsymbol{\Omega}(\boldsymbol{k})$ is smooth and $W_{\boldsymbol{k}\boldsymbol{k}'}$ is localized, the relaxation time is given by

$$\tau = \frac{\tau_0}{\bar{D}(\boldsymbol{k})} \approx \tau_0\left(1 - \frac{e}{\hbar}\boldsymbol{B} \cdot \boldsymbol{\Omega}\right). \qquad (5.41)$$

More generally, we can always expand the relaxation time to first order of $(e/\hbar)\boldsymbol{B} \cdot \boldsymbol{\Omega}$,

$$\tau = \tau_0 + \tau_1 \frac{e}{\hbar} \boldsymbol{B} \cdot \boldsymbol{\Omega}, \tag{5.42}$$

where τ_1 should be regarded as a fitting parameter within this theory.

Expanding Eq. (5.37) to first order of B and taking the spherical band approximation, we obtain

$$\sigma_{xx}^{(1)} = e^2 \tau_0 B \left[\int [d\boldsymbol{k}] \frac{\partial f_0}{\partial \varepsilon} \left(\frac{2e\Omega_z}{\hbar} v_x^2 + \frac{2}{\hbar} \frac{\partial m_z}{\partial k_x} v_x \right) \right.$$
$$\left. - \frac{e}{\hbar} \langle (\boldsymbol{M})_{xx}^{-1} \rangle_{k_F} \int [d\boldsymbol{k}] f \Omega_z \right], \tag{5.43}$$

where $\boldsymbol{M}$ is the effect mass tensor. The zero-field conductivity takes the usual form

$$\sigma_{xx}^{(0)} = -e^2 \tau_0 \int [d\boldsymbol{k}] \frac{\partial f_0}{\partial \varepsilon} v_x^2. \tag{5.44}$$

The ratio $-\sigma_{xx}^{(1)}/\sigma_{xx}^{(0)}$ will then give us the linear magnetoresistance.

VI. ELECTRON DYNAMICS UNDER GENERAL PERTURBATIONS

In this section we present the general theory of electron dynamics in slowly perturbed crystals (Sundaram and Niu, 1999; Panati *et al.*, 2003; Shindou and Imura, 2005). As expected, the Berry curvature enters into the equations of motion and modifies the density of states. The difference is that one needs to introduce the Berry curvature defined in the extended parameter space $(\boldsymbol{r}, \boldsymbol{q}, t)$. Two physical applications are considered: electron dynamics in deformed crystals and adiabatic current induced by inhomogeneity.

A. Equations of motion

We consider a slowly perturbed crystal whose Hamiltonian can be expressed in the following form:

$$H[\boldsymbol{r}, \boldsymbol{p}; \beta_1(\boldsymbol{r}, t), \dots, \beta_g(\boldsymbol{r}, t)], \tag{6.1}$$

where $\{\beta_i(\boldsymbol{r}, t)\}$ are the modulation functions characterizing the perturbations. They may represent either gauge potentials of electromagnetic fields, atomic displacements, charge or spin density waves, helical magnetic structures, or compositional gradients. Following the same procedure as used in the previous section, we expand the Hamiltonian around the wave-packet center and obtain

$$H = H_c + \Delta H, \tag{6.2}$$

$$H_c = H[\boldsymbol{r}, \boldsymbol{p}; \{\beta_i(\boldsymbol{r}_c, t)\}], \tag{6.3}$$

$$\Delta H = \sum_i \boldsymbol{\nabla}_{r_c} \beta_i(\boldsymbol{r}_c, t) \cdot \left\{ (\boldsymbol{r} - \boldsymbol{r}_c), \frac{\partial H}{\partial \beta_i} \right\}. \tag{6.4}$$

Since the local Hamiltonian H_c maintains periodicity of the unperturbed crystal, its eigenstates take the Bloch form

$$H_c(\boldsymbol{r}_c, t) |\psi_{\boldsymbol{q}}(\boldsymbol{r}_c, t)\rangle = \varepsilon_c(\boldsymbol{r}_c, \boldsymbol{q}, t) |\psi_{\boldsymbol{q}}(\boldsymbol{r}_c, t)\rangle, \tag{6.5}$$

where $\boldsymbol{q}$ is the Bloch wave vector and $\varepsilon_c(\boldsymbol{r}_c, \boldsymbol{q}, t)$ is the band energy. Here we have dropped the band index n for simplicity.

Following the discussion in Sec. I.D, we switch to the Bloch Hamiltonian $H_c(\boldsymbol{q}, \boldsymbol{r}_c, t) = e^{-i\boldsymbol{q} \cdot \boldsymbol{r}} H_c(\boldsymbol{r}_c, t) e^{i\boldsymbol{q} \cdot \boldsymbol{r}}$, whose eigenstate is the periodic part of the Bloch function, $|u(\boldsymbol{q}, \boldsymbol{r}_c, t)\rangle = e^{-i\boldsymbol{q} \cdot \boldsymbol{r}} |\psi(\boldsymbol{q}, \boldsymbol{r}_c, t)\rangle$. The Berry vector potentials can be defined for each of the coordinates of the parameter space $(\boldsymbol{q}, \boldsymbol{r}_c, t)$; for example,

$$\mathcal{A}_t = \langle u | i \partial_t | u \rangle. \tag{6.6}$$

After constructing the wave packet using the local Bloch functions $|\psi_{\boldsymbol{q}}(\boldsymbol{r}_c, t)\rangle$, one can apply the time-dependent variational principle to find the Lagrangian governing the dynamics of the wave packet:

$$L = -\varepsilon + \boldsymbol{q}_c \cdot \dot{\boldsymbol{r}}_c + \dot{\boldsymbol{q}}_c \cdot \mathcal{A}_{\boldsymbol{q}} + \dot{\boldsymbol{r}}_c \cdot \mathcal{A}_{\boldsymbol{r}} + \mathcal{A}_t. \tag{6.7}$$

Note that the wave-packet energy $\varepsilon = \varepsilon_c + \Delta\varepsilon$ has a correction $\Delta\varepsilon$ from ΔH,

$$\Delta\varepsilon = \langle W | \Delta H | W \rangle = -\Im \left\langle \frac{\partial u}{\partial \boldsymbol{r}_c} \right| \cdot (\varepsilon_c - H_c) \left| \frac{\partial u}{\partial \boldsymbol{q}} \right\rangle. \tag{6.8}$$

From the Lagrangian (6.7) we obtain the following equations:

$$\dot{\boldsymbol{r}}_c = \frac{\partial \varepsilon}{\partial \boldsymbol{q}_c} - (\overleftrightarrow{\Omega}_{qr} \cdot \dot{\boldsymbol{r}}_c + \overleftrightarrow{\Omega}_{qq} \cdot \dot{\boldsymbol{q}}_c) - \boldsymbol{\Omega}_{qt}, \tag{6.9a}$$

$$\dot{\boldsymbol{q}}_c = -\frac{\partial \varepsilon}{\partial \boldsymbol{r}_c} + (\overleftrightarrow{\Omega}_{rr} \cdot \dot{\boldsymbol{r}}_c + \overleftrightarrow{\Omega}_{rq} \cdot \dot{\boldsymbol{q}}_c) + \boldsymbol{\Omega}_{rt}, \tag{6.9b}$$

where Ω's are the Berry curvatures. For example,

$$(\overleftrightarrow{\Omega}_{qr})_{\alpha\beta} = \partial_{q_\alpha} \mathcal{A}_{r_\beta} - \partial_{r_\beta} \mathcal{A}_{k_\alpha}. \tag{6.10}$$

In the following we also drop the subscript c on $\boldsymbol{r}_c$ and $\boldsymbol{q}_c$.

The form of the equations of motion is quite symmetrical with respect to $\boldsymbol{r}$ and $\boldsymbol{q}$, and there are Berry curvatures between every pair of phase-space variables plus time. The term $\boldsymbol{\Omega}_{qt}$ was identified as the adiabatic velocity vector in Sec. II. In fact, if the perturbation is uniform in space (has the same period as the unperturbed crystal) and only varies in time, all the spatial derivatives vanish; we obtain

$$\dot{\boldsymbol{r}} = \frac{\partial \varepsilon}{\partial \boldsymbol{q}} - \boldsymbol{\Omega}_{qt}, \quad \dot{\boldsymbol{q}} = 0. \tag{6.11}$$

The first equation is the velocity formula (2.5) obtained in Sec. II. The term $\overleftrightarrow{\Omega}_{qq}$ was identified as the Hall conductivity tensor. In the presence of electromagnetic perturbations, we have

$$H = H_0[\boldsymbol{q} + e\boldsymbol{A}(\boldsymbol{r})] - e\phi(\boldsymbol{r}, t). \tag{6.12}$$

Hence the local basis can be written as $|u(\boldsymbol{r}_c, \boldsymbol{q})\rangle = |u(\boldsymbol{k})\rangle$, where $\boldsymbol{k} = \boldsymbol{q} + e\boldsymbol{A}(\boldsymbol{r})$. One can verify that by using the chain rule $\partial_{q_\alpha} = \partial_{k_\alpha}$ and $\partial_{r_\alpha} = (\partial_{r_\alpha} A_\beta) \partial_{k_\beta}$, $\Delta\varepsilon$ given in Eq.

(6.8) becomes $-\boldsymbol{m}(\boldsymbol{k})\cdot\boldsymbol{B}$, and the equations of motion (6.9) reduce to Eq. (5.8). The physics of quantum adiabatic transport and the quantum and anomalous Hall effect can be described from a unified point of view. The Berry curvature $\overset{\leftrightarrow}{\Omega}_{rt}$ plays a role like the electric force. The antisymmetric tensor $\overset{\leftrightarrow}{\Omega}_{rr}$ is realized in terms of the magnetic field in the Lorenz force and is also seen in the singular form (δ-function-like distribution) of dislocations in a deformed crystal (Bird and Preston, 1988). Finally, the Berry curvature between $\boldsymbol{r}$ and $\boldsymbol{q}$ can be realized in deformed crystals as a quantity proportional to the strain and the electronic mass renormalization in the crystal (Sundaram and Niu, 1999).

B. Modified density of states

The electron density of states is also modified by the Berry curvature. Consider the time-independent case. To better appreciate the origin of this modification, we introduce the phase-space coordinates $\boldsymbol{\xi}=(\boldsymbol{r},\boldsymbol{q})$. The equations of motion can be written as

$$\Gamma_{\alpha\beta}\dot{\xi}_\beta = \nabla_{\xi_\alpha}\varepsilon, \tag{6.13}$$

where $\overset{\leftrightarrow}{\Gamma}=\overset{\leftrightarrow}{\Omega}-\overset{\leftrightarrow}{J}$ is an antisymmetric matrix with

$$\overset{\leftrightarrow}{\Omega} = \begin{pmatrix} \overset{\leftrightarrow}{\Omega}_{rr} & \overset{\leftrightarrow}{\Omega}_{rq} \\ \overset{\leftrightarrow}{\Omega}_{qr} & \overset{\leftrightarrow}{\Omega}_{qq} \end{pmatrix}, \quad \overset{\leftrightarrow}{J} = \begin{pmatrix} 0 & \overset{\leftrightarrow}{I} \\ -\overset{\leftrightarrow}{I} & 0 \end{pmatrix}. \tag{6.14}$$

According to standard theory of Hamiltonian dynamics (Arnold, 1978), the density of states, which is proportional to the phase-space measure, is given by

$$D(\boldsymbol{r},\boldsymbol{q}) = \frac{1}{(2\pi)^d}\sqrt{\det(\overset{\leftrightarrow}{\Omega}-\overset{\leftrightarrow}{J})}. \tag{6.15}$$

One can show that in the time-dependent case $D(\boldsymbol{r},\boldsymbol{q})$ has the same form.

Consider the following situations. (i) If the perturbation is electromagnetic field, by the variable substitution $\boldsymbol{k}=\boldsymbol{q}+e\boldsymbol{A}(\boldsymbol{r})$, Eq. (6.15) reduces to Eq. (5.13). (ii) In many situations we are aiming at a first-order calculation in the spatial gradient. In this case, the density of states is given by

$$D = \frac{1}{(2\pi)^d}(1 + \mathrm{Tr}\,\overset{\leftrightarrow}{\Omega}_{qr}). \tag{6.16}$$

Note that if the Berry curvature vanishes, Eq. (6.13) becomes the canonical equations of motion for Hamiltonian dynamics, and $\boldsymbol{r}$ and $\boldsymbol{q}$ are called canonical variables. The density of states is a constant in this case. The presence of the Berry curvature renders the variables noncanonical and, as a consequence, modifies the density of states. The noncanonical variables are a common feature of Berry phase participated dynamics (Littlejohn and Flynn, 1991). The modified density of states also arises naturally from a nonequilibrium approach (Olson and Ao, 2007).

To demonstrate the modified density of states, we again consider the Rice-Mele model discussed in Sec.
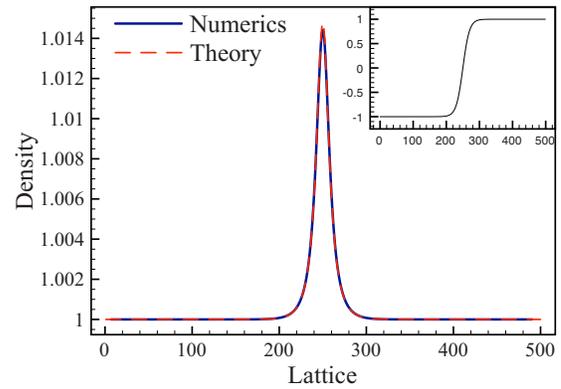

FIG. 11. (Color online) Electron density of the Rice-Mele model with a spatial varying dimerization parameter. The parameters used are $\Delta=0.5$, $t=2$, and $\delta=\tanh(0.02x)$. Inset: The profile of $\delta(x)$. From Xiao, Clougherty, and Niu, 2007.

II.C.1. Now we introduce the spatial dependence by letting the dimerization parameter $\delta(x)$ vary in space. Using Eq. (1.19) we find

$$\Omega_{qx} = \frac{\Delta t \sin^2(q/2)\partial_x\delta}{4(\Delta^2 + t^2\cos^2 q/2 + \delta^2\sin^2 q/2)^{3/2}}. \tag{6.17}$$

At half filling, the system is an insulator and its electron density is given by

$$n_e = \int_{-\pi}^{\pi}\frac{dq}{2\pi}\Omega_{qx}. \tag{6.18}$$

We let $\delta(x)$ have a kink in its profile. Such a domain wall is known to carry fractional charge (Su *et al.*, 1979; Rice and Mele, 1982). Figure 11 shows the calculated electron density using Eq. (6.18) together with numerical result obtained by direct diagonalization of the tight-binding Hamiltonian. These two results are virtually indistinguishable in the plot, which confirms the Berry phase modification to the density of states.

C. Deformed crystal

In this section we present a general theory of electron dynamics in crystals with deformation (Sundaram and Niu, 1999), which could be caused by external pressure, defects in the lattice, or interfacial strain.

First we set up the basic notations for this problem. Consider a deformation described by the atomic displacement $\{\boldsymbol{u}_l\}$. We denote the deformed crystal potential as $V(\boldsymbol{r};\{\boldsymbol{R}_l+\boldsymbol{u}_l\})$, where $\boldsymbol{R}_l$ is the atomic position with l labeling the atomic site. Introducing a smooth displacement field $\boldsymbol{u}(\boldsymbol{r})$ such that $\boldsymbol{u}(\boldsymbol{R}_l+\boldsymbol{u}_l)=\boldsymbol{u}_l$, the Hamiltonian can be written as

$$H = \frac{p^2}{2m} + V[\boldsymbol{r}-\boldsymbol{u}(\boldsymbol{r})] + s_{\alpha\beta}(\boldsymbol{r})\mathcal{V}_{\alpha\beta}[\boldsymbol{r}-\boldsymbol{u}(\boldsymbol{r})], \tag{6.19}$$

where $s_{\alpha\beta}=\partial u_\alpha/\partial r_\beta$ is the unsymmetrized strain and $\mathcal{V}_{\alpha\beta}[\boldsymbol{r}-\boldsymbol{u}(\boldsymbol{r})]=\Sigma_l[\boldsymbol{R}_l+\boldsymbol{u}(\boldsymbol{r})-\boldsymbol{r}]_\beta(\partial V/\partial R_{l\alpha})$ is a gradient expansion of the crystal potential. The last term, being

proportional to the strain, can be treated perturbatively. The local Hamiltonian is given by

$$H_c = \frac{p^2}{2m} + V[\mathbf{r} - \mathbf{u}(\mathbf{r}_c)], \qquad (6.20)$$

with its eigenstates $|\psi_q(\mathbf{r} - \mathbf{u}(\mathbf{r}_c))\rangle$.

To write down the equations of motion, two pieces of information are needed. One is the gradient correction to the electron energy, given by Eq. (6.8). Sundaram and Niu (1999) found that

$$\Delta\varepsilon = s_{\alpha\beta}D_{\alpha\beta}(\mathbf{q}), \qquad (6.21)$$

where

$$D_{\alpha\beta} = m[v_\alpha v_\beta - \langle \hat{v}_\alpha \hat{v}_\beta \rangle] + \langle \mathcal{V}_{\alpha\beta} \rangle, \qquad (6.22)$$

with $\langle \cdots \rangle$ the expectation value of the enclosed operators in the Bloch state and $\hat{v}_\alpha$ is the velocity operator. Note that in the free electron limit ($V \to 0$) this quantity vanishes. This is expected since a wave packet should not feel the effect of a deformation of the lattice in the absence of electron-phonon coupling. The other piece is the Berry curvature, which is derived from the Berry vector potentials. For deformed crystals, in addition to $\mathcal{A}_q$, there are two other vector potentials

$$\mathcal{A}_r = f_\alpha \frac{\partial u_\alpha}{\partial \mathbf{r}}, \quad \mathcal{A}_t = f_\alpha \frac{\partial u_\alpha}{\partial t}, \qquad (6.23)$$

with

$$\mathbf{f}(\mathbf{q}) = \frac{m}{\hbar}\frac{\partial\varepsilon}{\partial\mathbf{q}} - \hbar\mathbf{q}. \qquad (6.24)$$

This then leads to the following Berry curvatures:

$$\Omega_{q_\alpha r_\beta} = -\frac{\partial u_\gamma}{\partial r_\beta}\frac{\partial f_\gamma}{\partial q_\alpha}, \quad \Omega_{q_\alpha t} = \frac{\partial u_\gamma}{\partial t}\frac{\partial f_\gamma}{\partial q_\alpha},$$

$$\Omega_{r_\alpha r_\beta} = \Omega_{x_\alpha t} = 0. \qquad (6.25)$$

With the above information we plug in the electron energy as well as the Berry curvatures into Eq. (6.9) to obtain the equations of motion.

We first consider a one-dimensional insulator with lattice constant a. Suppose the system is under a uniform strain with a new lattice constant $a + \delta a$, i.e., $\partial_x u = \delta a / a$. Assuming one electron per unit cell, the electron density goes from $1/a$ to

$$\frac{1}{a + \delta a} = \frac{1}{a}\left(1 - \frac{\delta a}{a}\right). \qquad (6.26)$$

On the other hand, we can also directly calculate the change in the electron density using the modified density of states (6.16), which gives

$$\int_0^{2\pi/a} \frac{dq}{2\pi}\Omega_{qx} = -\frac{\delta a}{a^2}. \qquad (6.27)$$

From a physical point of view, this change says an insulator under a uniform strain remains an insulator.

The above formalism is also applicable to dislocation strain fields, which are well defined except in a region of a few atomic spacings around the line of dislocation. Outside this region, the displacement field $\mathbf{u}(\mathbf{r})$ is a smooth but multiple valued function. On account of this multiple valuedness, a wave packet of incident wave vector $\mathbf{q}$ taken around the line of dislocation acquires a Berry phase

$$\gamma = \oint_c d\mathbf{r} \cdot \mathcal{A}_\mathbf{r} = \oint_c d\mathbf{u} \cdot f(\mathbf{k}) \approx \mathbf{b} \cdot \mathbf{f}(\mathbf{k}), \qquad (6.28)$$

where $\mathbf{b} = \oint dr_\alpha \partial \mathbf{u}/\partial r_\alpha$ is known as the Burgers vector. What we have here is a situation similar to the Aharonov-Bohm effect (Aharonov and Bohm, 1959), with the dislocation playing the role of the solenoid, and the Berry curvature $\mathbf{\Omega}_{rr}$ the role of the magnetic field. Bird and Preston (1988) showed that this Berry phase can affect the electron diffraction pattern of a deformed crystal.

The above discussion only touches a few general ideas of the Berry phase effect in deformed crystals. With complete information on the equations of motion, the semiclassical theory provides a powerful tool to investigate the effects of deformation on electron dynamics and equilibrium properties.

D. Polarization induced by inhomogeneity

In Sec. II.C we discussed the Berry phase theory of polarization in crystalline solids, based on the basic idea that the polarization is identical to the integration of the adiabatic current flow in the bulk. There the system is assumed to be periodic and the perturbation depends only on time (or any scalar for that matter). In this case, it is straightforward to obtain the polarization based on the equations of motion (6.11). However, in many physical situations the system is in an inhomogeneous state and the electric polarization strongly depends on the inhomogeneity. Examples include flexoelectricity where a finite polarization is produced by a strain gradient (Tagantsev, 1986, 1991) and multiferroic materials where the magnetic ordering varies in space and induces a polarization (Fiebig *et al.*, 2002; Kimura *et al.*, 2003; Hur *et al.*, 2004; Cheong and Mostovoy, 2007).

Consider an insulating crystal with an order parameter that varies slowly in space. We assume that, at least at the mean-field level, the system can be described by a perfect crystal under the influence of an external field $\mathbf{h}(\mathbf{r})$. If, for example, the order parameter is the magnetization, then $\mathbf{h}(\mathbf{r})$ can be chosen as the exchange field that yields the corresponding spin configuration. Our goal is to calculate the electric polarization to first order in the spatial gradient as the field $\mathbf{h}(\mathbf{r})$ is gradually turned on. The Hamiltonian thus takes the form $H[\mathbf{h}(\mathbf{r}); \lambda]$, where λ is the parameter describing the turning on process. Xiao *et al.* (2009) showed that the first-order contribution to the polarization can be classified into two categories: the perturbative contribution due to

the correction to the wave function and the topological contribution which is from the dynamics of the electrons.

First consider the perturbation contribution, which is basically a correction to the polarization formula obtained by King-Smith and Vanderbilt (1993) for a uniform system. The perturbative contribution is obtained by evaluating the Berry curvature Ω_{qt} in Eq. (2.27) to first order of the gradient. Remember that we always expand the Hamiltonian into the form $H = H_c + \Delta H$ and choose the eigenfunctions of H_c as our expansion basis. Hence to calculate the Berry curvature to first order of the gradient, one needs to know the form of the wave function perturbed by ΔH. This calculation has been discussed in the case of an electric field (Nunes and Gonze, 2001; Souza *et al.*, 2002).

The topological contribution is of different nature. Starting from Eq. (6.9) and making use of the modified density of states (6.16), one finds the adiabatic current induced by inhomogeneity is given by

$$j_\alpha^{(2)} = e \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{(2\pi)^d}(\Omega_{\alpha\beta}^{qq}\Omega_\beta^{r\lambda} + \Omega_{\beta\beta}^{qr}\Omega_\alpha^{q\lambda} - \Omega_{\alpha\beta}^{qr}\Omega_\beta^{q\lambda}).$$

(6.29)

We can see that this current is explicitly proportional to the spatial gradient. Comparing this equation with Eq. (2.6) reveals an elegant structure: the zeroth-order contribution [Eq. (2.6)] is given as an integral of the first Chern form, while the first-order contribution [Eq. (6.29)] is given as an integral of the second Chern form. A similar result has been obtained by Qi *et al.* (2008).

The polarization is obtained by integrating the current. As usually in the case of multiferroics, we can assume the order parameter is periodic in space (but in general incommensurate with the crystal lattice). A two-point formula can be written down[13]

$$P_\alpha^{(2)} = \frac{e}{\mathcal{V}}\int d\mathbf{r}\int_{\mathrm{BZ}}\frac{d\mathbf{q}}{(2\pi)^d}(\mathcal{A}_\alpha^q\nabla_\beta^r\mathcal{A}_\beta^q + \mathcal{A}_\beta^q\nabla_\alpha^q\mathcal{A}_\beta^r$$
$$+ \mathcal{A}_\beta^r\nabla_\beta^q\mathcal{A}_\alpha^q)|_0^1,$$

(6.30)

where $\mathcal{V}$ is the volume of the periodic structure of the order parameter. Again, due to the loss of tracking of λ, there is an uncertain quantum which is the second Chern number. If we assume the order parameter has period l_y in the y direction, the polarization quantum in the x direction is given by

$$\frac{e}{l_y a_z},$$

(6.31)

where a is the lattice constant.

Kunz (1986) discussed the charge pumping in incommensurate potentials and showed that in general the

charge transport is quantized and given in the form of Chern numbers, which is consistent with what we have derived.

The second Chern form demands that the system must be two dimensions or higher, otherwise the second Chern form vanishes. It allows us to determine the general form of the induced polarization. Consider a two-dimensional minimal model with $\mathbf{h}(\mathbf{r})$ having two components. If we write $H[\mathbf{h}(\mathbf{r});\lambda]$ as $H[\lambda\mathbf{h}(\mathbf{r})]$, i.e., λ acts like a switch, the polarization can be written as

$$\mathbf{P}^{(2)} = \frac{e}{\mathcal{V}}\int d\mathbf{r}\chi[(\boldsymbol{\nabla}\cdot\mathbf{h})\mathbf{h} - (\mathbf{h}\cdot\boldsymbol{\nabla})\mathbf{h}].$$

(6.32)

The coefficient χ is given by

$$\chi = \frac{e}{8}\int_{\mathrm{BZ}}\frac{d\mathbf{q}}{(2\pi)^2}\int_0^1\frac{d\lambda}{\lambda}\epsilon_{abcd}\Omega_{ab}\Omega_{cd},$$

(6.33)

where the Berry curvature is defined on the parameter space $(\mathbf{q},\mathbf{h})$ and ϵ_{abcd} is the Levi-Città antisymmetric tensor.

Xiao *et al.* (2009) showed how the two-point formula can be implemented in numerical calculations using a discretized version (Kotiuga, 1989).

1. Magnetic-field-induced polarization

An important application of the above result is the magnetic-field-induced polarization. Essin *et al.* (2009) considered an insulator in the presence of a vector potential $\mathbf{A} = By\hat{\mathbf{z}}$ with its associated magnetic field $\mathbf{B} = h/ea_z l_y \hat{\mathbf{x}}$. The inhomogeneity is introduced through the vector potential in the z direction. Note that magnetic flux over the supercell $a_z \times l_y$ in the x direction is exactly h/e, therefore the system is periodic in the y direction with period l_y. According to our discussion in Sec. V, the effect of a magnetic field can be counted by the Peierls substitution $k_z \to k_z + eBy/\hbar$, hence $\nabla_y = (eB/\hbar)\nabla_{kz}$. Applying Eq. (6.30), one obtains the induced polarization

$$P_x = \frac{\theta e^2}{2\pi h}B,$$

(6.34)

with

$$\theta = \frac{1}{2\pi}\int_{\mathrm{BZ}}d\mathbf{k}\varepsilon_{\alpha\beta\gamma}\mathrm{Tr}\left[\mathcal{A}_\alpha\partial_\beta\mathcal{A}_\gamma - i\frac{2}{3}\mathcal{A}_\alpha\mathcal{A}_\beta\mathcal{A}_\gamma\right],$$

(6.35)

where $(\mathcal{A}_\alpha)_{mn} = i\langle u_m|\nabla_{k_\alpha}|u_n\rangle$ is the non-Abelian Berry connection, discussed in Sec. IX. Recall that the polarization is defined as the response of the total energy to an electric field $\mathbf{P} = \partial\mathcal{E}/\partial\mathbf{E}$, such a magnetic-field-induced polarization implies that there is an electromagnetic coupling of the form

$$\Delta\mathcal{L}_{\mathrm{EM}} = \frac{\theta e^2}{2\pi h}\mathbf{E}\cdot\mathbf{B}.$$

(6.36)

This coupling, labeled "axion electrodynamics," was discussed by Wilczek (1987). When $\theta = \pi$, the corresponding insulator is known as a 3D Z_2 topological insulator (Qi *et al.*, 2008).

---

[13]So far we only considered the Abelian Berry case. The non-Abelian result is obtained by replacing the Chern-Simons form with its non-Abelian form.

E. Spin texture

So far our discussion has focused on the physical effects of the Berry curvature in the momentum space ($\mathbf{\Omega}_{kk}$) or in the mixed space of the momentum coordinates and some other parameters ($\mathbf{\Omega}_{kr}$ and $\mathbf{\Omega}_{kt}$). In this section we discuss the Berry curvatures which originate only from the nontrivial real-space configuration of the system.

One of such systems is magnetic materials with domain walls or spin textures. Consider a ferromagnetic thin films described by the following Hamiltonian:

$$H = \frac{\mathbf{p}^2}{2m} - J\hat{\mathbf{n}}(\mathbf{r},t) \cdot \boldsymbol{\sigma}, \tag{6.37}$$

where the first term is the bare Hamiltonian for a conduction electron and the second term is the s-d coupling between the conduction electron and the local d-electron spin along the direction $\hat{\mathbf{n}}(\mathbf{r},t)$ with J the coupling strength. Note that we have allowed the spin texture to vary in both space and time. The simple momentum dependence of the Hamiltonian dictates that all $\mathbf{k}$-dependent Berry curvatures vanish.

Because of the strong s-d coupling, we adopt the adiabatic approximation which states that the electron spin will follow the local spin direction during its motion. Then the spatial variation in local spin textures gives rise to the Berry curvature field

$$\mathbf{\Omega}_{rr} = \tfrac{1}{2} \sin \theta (\boldsymbol{\nabla}\theta \times \boldsymbol{\nabla}\phi), \tag{6.38}$$

where θ and ϕ are the spherical angles specifying the direction of $\hat{\mathbf{n}}$. According to Eqs. (6.9), this field acts on the electrons as an effective magnetic field. In addition, the time-dependence of the spin texture also gives rises to

$$\mathbf{\Omega}_{rt} = \tfrac{1}{2} \sin \theta (\partial_t\phi \boldsymbol{\nabla}\theta - \partial_t\theta \boldsymbol{\nabla}\phi). \tag{6.39}$$

Similarly, $\mathbf{\Omega}_{rt}$ acts on the electrons as an effective electric field. This is in analogy with a moving magnetic field ($\mathbf{\Omega}_{rr}$) generating an electric field ($\mathbf{\Omega}_{rt}$).

The physical consequences of these two fields are obvious by analogy with the electromagnetic fields. The Berry curvature $\mathbf{\Omega}_{rr}$ will drive a Hall current, just like the ordinary Hall effect (Ye *et al.*, 1999; Bruno *et al.*, 2004). Unlike the anomalous Hall effect discussed in Sec. III.D, this mechanism for a nonvanishing Hall effect does not require the spin-orbit coupling but does need a topologically nontrivial spin texture, for example, a skyrmion. On the other hand, for a moving domain wall in a thin magnetic wire the Berry curvature $\mathbf{\Omega}_{rt}$ will induce an electromotive force, which results in a voltage difference between the two ends. This Berry curvature induced emf has recently been experimentally measured (Yang *et al.*, 2009).

VII. QUANTIZATION OF ELECTRON DYNAMICS

In previous sections we have reviewed several Berry phase effects in solid-state systems. Berry curvature of-

ten appears as a result of restricting (or projecting) the extent of a theory to its subspace. In particular, the Berry curvature plays a crucial role in the semiclassical dynamics of electrons, which is valid under the one-band approximation. In the following, we explain how the semiclassical formulation could be requantized. This is necessary, for example, in studying the quantized Wannier-Stark ladders from the Bloch oscillation, or the quantized Landau levels from the cyclotron orbit (Ashcroft and Mermin, 1976). The requantized theory is valid in the same subspace of the semiclassical theory. It will become clear that, the knowledge of the Bloch energy, the Berry curvature, and the magnetic moment in the semiclassical theory constitute sufficient information for building the requantized theory. In this section we focus on the following methods of quantization: the Bohr-Sommerfeld quantization and the canonical quantization.

A. Bohr-Sommerfeld quantization

A method of quantization is a way to select quantum mechanically allowed states out of a continuum of classical states. This is often formulated using the generalized coordinates q_i and their conjugate momenta p_i. The Bohr-Sommerfeld quantization requires the action integral for each set of the conjugate variables to satisfy

$$\oint_{C_i} p_i dq_i = \left(m_i + \frac{\nu_i}{4}\right)h, \quad i = 1, \dots, d, \tag{7.1}$$

where C_i are closed trajectories in the phase space with dimension $2d$, m_i are integers, and ν_i are the so-called Maslov indices, which are usually integers. Notice that since the choice of conjugate variables may not be unique, the Bohr-Sommerfeld quantization method could give inequivalent quantization rules. This problem can be fixed by the following Einstein-Brillouin-Keller (EBK) quantization rule.

For a completely integrable system, there are d constants of motion. As a result, the trajectories in the phase space are confined to a d-dimensional torus. On such a torus, one can have d closed loops that are topologically independent of each other (Tabor, 1989). That is, one cannot be deformed continuously to the other. Since $\Sigma_{\ell=1}^{d} p_\ell dq_\ell$ is invariant under coordinate transformation, instead of Eq. (7.1), one can use the following EBK quantization condition:

$$\oint_{C_k} \sum_{\ell=1}^{d} p_\ell dq_\ell = \left(m_k + \frac{\nu_k}{4}\right)h, \quad k = 1, \dots, d, \tag{7.2}$$

where C_k are periodic orbits on invariant tori. Such a formula is geometric in nature (i.e., it is coordinate independent). Furthermore, it applies to the invariant tori of systems that may not be completely integrable. Therefore, the EBK formula plays an important role in quantizing chaotic systems (Stone, 2005). In the following, we refer to both types of quantization simply as the Bohr-Sommerfeld quantization.

In the wave-packet formulation of Bloch electrons, both r_c and q_c are treated as generalized coordinates. With the Lagrangian in Eq. (5.7), one can find their conjugate momenta $\partial L/\partial \dot{r}_c$ and $\partial L/\partial \dot{q}_c$, which are equal to $\hbar q_c$ and $\hbar\langle u|i\partial u/\partial q_c\rangle = \hbar \mathcal{A}$, respectively (Sundaram and Niu, 1999). The quantization condition for an orbit with constant energy thus becomes

$$\oint_C q_c \cdot dr_c = 2\pi\left(m + \frac{\nu}{4} - \frac{\Gamma_C}{2\pi}\right), \qquad (7.3)$$

where $\Gamma_C \equiv \oint_C \mathcal{A} \cdot dq_c$ is the Berry phase of an energy contour C [see also Wilkinson (1984b) and Kuratsuji and Iida (1985)]. Since the Berry phase is path dependent, one may need to solve the equation self-consistently to obtain the quantized orbits.

Before applying the Bohr-Sommerfeld quantization in the following sections, we point out two disadvantages of this method. First, the value of the Maslov index is not always apparent. For example, for a one-dimensional particle bounded by two walls, its value would depend on the slopes of the walls (van Houten *et al.*, 1989). In fact, a noninteger value may give a more accurate prediction of the energy levels (Friedrich and Trost, 1996). Second, this method fails if the trajectory in phase space is not closed, or if the dynamic system is chaotic and invariant tori fail to exist. On the other hand, the method of canonical quantization in Sec. VII.D does not have these problems.

B. Wannier-Stark ladder

Consider an electron moving in a one-dimensional periodic lattice with lattice constant a. Under a weak uniform electric field E, according to the semiclassical equations of motion, the quasimomentum of an electron wave packet is simply [see Eq. (5.8)]

$$\hbar q_c(t) = -eEt. \qquad (7.4)$$

It takes the time $T_B = h/eEa$ for the electron to traverse the first Brillouin zone. Therefore, the angular frequency of the periodic motion is $\omega_B = eEa/\hbar$. This is the so-called Bloch oscillation (Ashcroft and Mermin, 1976). Similar to a simple harmonic oscillator, the energy of the oscillatory electron is quantized in multiples of $\hbar\omega_B$. However, unlike the former, the Bloch oscillator has no zero-point energy (that is, the Maslov index is zero). These equally spaced energy levels are called the Wannier-Stark ladders. Since the Brillouin zone is periodic, the electron orbit is closed. According to the Bohr-Sommerfeld quantization, one has

$$\oint_{C_m} r_c \cdot dq_c = -2\pi\left(m - \frac{\Gamma_{C_m}}{2\pi}\right). \qquad (7.5)$$

For a simple one-dimensional lattice with inversion symmetry, if the origin is located at a symmetric point, then the Berry phase Γ_{C_m} can only have two values, 0 or π (Zak, 1989), as discussed in Sec. II.C.

Starting from Eq. (7.5), it is not difficult to find the average position of the electron,

$$\langle r_c\rangle_m = a\left(m - \frac{\Gamma_C}{2\pi}\right), \qquad (7.6)$$

where we have neglected the subscript m in Γ_{C_m} since all of the paths in the same energy band have the same Berry phase here. Such average positions $\langle r_c\rangle_m$ are the average positions of the Wannier function (Vanderbilt and King-Smith, 1993). Due to the Berry phase, they are displaced from the positive ions located at am.

In Sec. II.C the electric polarization is derived using the theory of adiabatic transport. It can also be obtained from the expectation value of the position operator directly. Because of the charge separation mentioned above, the one-dimensional crystal has a polarization $\Delta P = e\Gamma_c/2\pi$ (compared to the state without charge separation), which is the electric dipole per unit cell. This is consistent with the result in Eq. (2.28).

After time average, the quantized energies of the electron wave packet are

$$\langle \mathcal{E}\rangle_m = \langle\varepsilon(q_c)\rangle - eE\langle r_c\rangle_m = \varepsilon_0 - eEa\left(m - \frac{\Gamma_C}{2\pi}\right), \qquad (7.7)$$

which are the energy levels of the Wannier-Stark ladders.

Two short comments are in order: First, beyond the one-band approximation, there exist Zener tunnelings between Bloch bands. Therefore, the quantized levels are not stationary states of the system. They should be understood as resonances with finite lifetimes (Avron, 1982; Glück *et al.*, 1999). Second, the fascinating phenomenon of Bloch oscillation is not commonly observed in laboratory for the following reason: In an usual solid, the electron scattering time is shorter than the period T_B by several orders of magnitude. Therefore, the phase coherence of the electron is destroyed within a small fraction of a period. Nonetheless, with the help of a superlattice that has a much larger lattice constant the period T_B can be reduced by two orders of magnitude, which could make the Bloch oscillation and the accompanying Wannier-Stark ladders detectable (Mendez and Bastard, 1993). Alternatively, the Bloch oscillation and Wannier-Stark ladders can also be realized in an optical lattice (Ben Dahan *et al.*, 1996; Wilkinson and Kay, 1996), in which the atom can be coherent over a long period of time.

C. de Haas–van Alphen oscillation

When a uniform B field is applied to a solid, the electron would execute a cyclotron motion in both r and the k space. From Eq. (5.25b), it is not difficult to see that an orbit C in k space lies on the intersection of a plane perpendicular to the magnetic field and the constant-energy surface (Ashcroft and Mermin, 1976). Without quantization, the size of an orbit is determined by the initial energy of the electron and can be varied continuously. One then applies the Bohr-Sommerfeld quantiza-

tion rule, as Onsager did, to quantize the size of the orbit (Onsager, 1952). That is, only certain orbits satisfying the quantization rule are allowed. Each orbit corresponds to an energy level of the electron (i.e., the Landau level). Such a method remains valid in the presence of the Berry phase.

With the help of the semiclassical equation [see Eq. (5.8)],

$$\hbar \dot{\mathbf{k}}_c = -e\dot{\mathbf{r}}_c \times \mathbf{B}, \tag{7.8}$$

the Bohr-Sommerfeld condition in Eq. (7.3) can be written as (note that $\hbar \mathbf{q}_c = \hbar \mathbf{k}_c - e\mathbf{A}$, and $\nu = 2$)

$$\frac{\mathbf{B}}{2} \cdot \oint_{C_m} \mathbf{r}_c \times d\mathbf{r}_c = \left(m + \frac{1}{2} - \frac{\Gamma_{C_m}}{2\pi}\right)\phi_0, \tag{7.9}$$

where $\phi_0 \equiv h/e$ is the flux quantum. The integral on the left-hand side is simply the magnetic flux enclosed by the real-space orbit (allowing a drift along the $\mathbf{B}$ direction). Therefore, the enclosed flux has to jump in steps of the flux quantum (plus a Berry phase correction).

Similar to the Bohr atom model, in which the electron has to form a standing wave, here the total phase acquired by the electron after one circular motion also has to be integer multiples of 2π. Three types of phases contribute to the total phase: (a) the Aharonov-Bohm phase—an electron circulating a flux quantum picks up a phase of 2π; (b) the phase lag of π at each turning point (there are two of them)—this explains why the Maslov index is two; and (c) the Berry phase intrinsic to the solid. Therefore, Eq. (7.9) simply says that the summation of these three phases should be equal to $2\pi m$.

The orbit in k space can be obtained by rescaling the r-space orbit in Eq. (7.9) with a linear factor of λ_B^2, followed by a rotation of $90°$, where $\lambda_B \equiv \sqrt{\hbar/eB}$ is the magnetic length (Ashcroft and Mermin, 1976). Therefore, one has

$$\frac{\hat{\mathbf{B}}}{2} \cdot \oint_{C_m} \mathbf{k}_c \times d\mathbf{k}_c = 2\pi\left(m + \frac{1}{2} - \frac{\Gamma_{C_m}}{2\pi}\right)\frac{eB}{\hbar}. \tag{7.10}$$

The size of the orbit combined with the knowledge of the electron energy $E(\mathbf{k}_c) = \varepsilon(\mathbf{k}_c) - \mathbf{M} \cdot \mathbf{B}$ help determine the quantized energy levels. For an electron with a quadratic energy dispersion (before applying the magnetic field), these levels are equally spaced. However, with the Berry phase correction, which is usually different for different orbits, the energy levels are no longer uniformly distributed. This is related to the discussion in Sec. V on the relation between the density of states and the Berry curvature (Xiao *et al.*, 2005).

As a demonstration, we apply the quantization rule to graphene and calculate the energies of Landau levels near the Dirac point. Before applying a magnetic field, the energy dispersion near the Dirac point is linear, $\varepsilon(\mathbf{k}) = \hbar v_F k$. It is known that if the energy dispersion

near a degenerate point is linear, then the cyclotron orbit will acquire a Berry phase $\Gamma_C = \pi$, independent of the shape of the orbit (Blount, 1962b). As a result, the $1/2$ on the right-hand side of Eq. (7.10) is canceled by the Berry phase term. According to Eq. (7.10), the area of a cyclotron orbit is thus $\pi k^2 = 2\pi meB/\hbar$, where m is a nonnegative integer, from which one can easily obtain the Landau-level energy $\varepsilon_m = v_F\sqrt{2eB\hbar m}$. The experimental observation of a quantum Hall plateau at zero energy is thus a direct consequence of the Berry phase (Novoselov *et al.*, 2005, 2006; Zhang *et al.*, 2005).

In addition to point degeneracy, other types of degeneracy in momentum space can also be a source of the Berry phase. For example, the effect of the Berry phase generated by a line of band contact on magneto-oscillations is studied by Mikitik and Sharlai (1999, 2004).

The discussion so far is based on the one-band approximation. In reality, the orbit in one band would couple with the orbits in other bands. As a result, the Landau levels are broadened into minibands (Wilkinson, 1984a). A similar situation occurs in a magnetic Bloch band, which is the subject of Sec. VIII.

D. Canonical quantization (Abelian case)

In addition to the Bohr-Sommerfeld quantization, an alternative way to quantize a classical theory is by finding out position and momentum variables that satisfy the following Poisson brackets:

$$\{x_i, p_j\} = \delta_{ij}. \tag{7.11}$$

Afterwards, these classical canonical variables are promoted to operators that satisfy the commutation relation

$$[x_i, p_j] = i\hbar \delta_{ij}, \tag{7.12}$$

that is, all we need to do is to substitute the Poisson bracket $\{x_i, p_j\}$ by the commutator $[x_i, p_j]/i\hbar$. Based on the commutation relation, these variables can be written explicitly using either the differential-operator representation or the matrix representation. Once this is done, one can proceed to obtain the eigenvalues and eigenstates of the Hamiltonian $H(\mathbf{x}, \mathbf{p})$.

Even though one can always have canonical pairs in a Hamiltonian system, as guaranteed by the Darboux theorem (Arnold, 1989), in practice, however, finding them may not be a trivial task. For example, the center-of-mass variables $\mathbf{r}_c$ and $\mathbf{k}_c$ in the semiclassical dynamics in Eq. (5.8) are not canonical variables since their Poisson brackets below are not of the canonical form (Xiao *et al.*, 2005; Duval *et al.*, 2006a),

$$\{r_i, r_j\} = \epsilon_{ijk}\Omega_k/\kappa, \tag{7.13}$$

$$\{k_i, k_j\} = -\epsilon_{ijk}eB_k/\kappa, \tag{7.14}$$

$$\{r_i, k_j\} = (\delta_{ij} + eB_i\Omega_j)/\kappa, \tag{7.15}$$

where $\kappa \equiv 1 + e\boldsymbol{B}(\boldsymbol{r}) \cdot \boldsymbol{\Omega}(\boldsymbol{k})$. In order to carry out the canonical quantization, canonical variables of position and momentum must be found.

The derivation of Eqs. (7.13)–(7.15) is outlined as follows: One first writes the equations of motion in Eq. (5.7) in the form of Eq. (6.13) [$\boldsymbol{\xi} = (\boldsymbol{r}, \boldsymbol{k})$]. In this case, the only nonzero Berry curvatures are $(\overleftrightarrow{\Omega}_{\mathbf{rr}})_{ij} = -\epsilon_{ijk}eB_k$ and $(\overleftrightarrow{\Omega}_{\mathbf{kk}})_{ij} = \epsilon_{ijk}\Omega_k$, whereas $\overleftrightarrow{\Omega}_{\mathbf{rk}}$ and $\overleftrightarrow{\Omega}_{\mathbf{kr}}$ are zero. The Pois-

son bracket of two functions in the phase space f and g is then defined as

$$\{f, g\} = \left(\frac{\partial f}{\partial \boldsymbol{\xi}}\right)^T \overleftrightarrow{\Gamma}^{-1}\left(\frac{\partial g}{\partial \boldsymbol{\xi}}\right), \tag{7.16}$$

where $\overleftrightarrow{\Gamma} = \overleftrightarrow{\Omega} - \overleftrightarrow{J}$ [see Eq. (6.14)]. It is defined in such a way that the equations of motion can be written in the standard form: $\dot{\boldsymbol{\xi}} = \{\boldsymbol{\xi}, \varepsilon\}$. To linear order of magnetic field or Berry curvature, one can show that

$$\overleftrightarrow{\Gamma}^{-1} = \frac{1}{1 + e\boldsymbol{B} \cdot \boldsymbol{\Omega}}\begin{pmatrix} \overleftrightarrow{\Omega}_{kk} & \overleftrightarrow{I} - \overleftrightarrow{\Omega}_{kk}\overleftrightarrow{\Omega}_{rr} + e\boldsymbol{B} \cdot \boldsymbol{\Omega} \\ -\overleftrightarrow{I} + \overleftrightarrow{\Omega}_{rr}\overleftrightarrow{\Omega}_{kk} - e\boldsymbol{B} \cdot \boldsymbol{\Omega} & \overleftrightarrow{\Omega}_{rr} \end{pmatrix}. \tag{7.17}$$

Equations (7.13)–(7.15) thus follow when f and g are identified with the components of $\boldsymbol{\xi}$.

We start with two special cases. The first is a solid with zero Berry curvature that is under the influence of a magnetic field ($\boldsymbol{\Omega} = 0$, $\boldsymbol{B} \neq 0$). In this case, the factor κ in Eq. (7.14) reduces to 1 and the position variables commute with each other. Obviously, if one assumes $\hbar\boldsymbol{k}_c = \boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{x})$ and requires $\boldsymbol{x}$ and $\boldsymbol{p}$ to be canonical conjugate variables, then the quantized version of Eq. (7.14) (with $i\hbar$ inserted) can easily be satisfied. This is the familiar Peierls substitution (Peierls, 1933).

In the second case, consider a system with Berry curvature but not in a magnetic field ($\boldsymbol{\Omega} \neq 0$, $\boldsymbol{B} = 0$). In this case, again we have $\kappa = 1$. Now the roles of $\boldsymbol{r}_c$ and $\boldsymbol{k}_c$ in the commutators are reversed. The momentum variables commute with each other but not the coordinates. One can apply a Peierls-like substitution to the coordinate variables and write $\boldsymbol{r}_c = \boldsymbol{x} + \mathcal{A}(\boldsymbol{q})$. It is not difficult to see that the commutation relations arising from Eq. (7.13) can indeed be satisfied. After the canonical quantization, $\boldsymbol{x}$ becomes $i\partial/\partial\boldsymbol{q}$ in the quasimomentum representation. Blount (1962b) showed that the position operator $\boldsymbol{r}$ in the one-band approximation acquires a correction, which is our Berry connection $\mathcal{A}$. Therefore, $\boldsymbol{r}_c$ can be identified with the projected position operator $P\boldsymbol{r}P$, where P projects to the energy band of interest.

When both $\boldsymbol{B}$ and $\boldsymbol{\Omega}$ are nonzero, applying both of the Peierls substitutions simultaneously is not enough to produce the correct commutation relations, mainly because of the nontrivial factor κ there. In general, exact canonical variables cannot be found easily. However, since the semiclassical theory itself is valid to linear order of field, we only need to find the canonical variables correct to the same order in practice. The result is (Chang and Niu, 2008)

$$\boldsymbol{r}_c = \boldsymbol{x} + \mathcal{A}(\boldsymbol{\pi}) + \boldsymbol{G}(\boldsymbol{\pi}), \tag{7.18}$$

$$\hbar\boldsymbol{k}_c = \boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{x}) + e\boldsymbol{B} \times \mathcal{A}(\boldsymbol{\pi}),$$

where $\boldsymbol{\pi} = \boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{x})$ and $G_\alpha(\boldsymbol{k}_c) \equiv (e/\hbar)(\mathcal{A} \times \boldsymbol{B}) \cdot \partial\mathcal{A}/\partial k_{c\alpha}$. This is the generalized Peierls substitution for systems with Berry connection $\mathcal{A}$ and vector potential $\boldsymbol{A}$. With these equations, one can verify Eqs. (7.13) and (7.15) to linear orders of $\boldsymbol{B}$ and $\boldsymbol{\Omega}$.

A few comments are in order: First, if a physical observable is a product of several canonical variables, the order of the product may become a problem after the quantization since the variables may not commute with each other. To preserve the Hermitian property of the physical observable, the operator product needs to be symmetrized. Second, the Bloch energy, Berry curvature, and orbital moment of the semiclassical theory contains sufficient information for building a quantum theory that accounts for all physical effects to first order in external fields. We return to this in Sec. IX, where the non-Abelian generalization of the canonical quantization method is addressed.

VIII. MAGNETIC BLOCH BANDS

The semiclassical dynamics in previous sections is valid when the external field is weak, so that the latter can be treated as a perturbation to the Bloch states. Such a premise is no longer valid if the external field is so strong that the structure of the Bloch bands is significantly altered. This happens, for example, in quantum Hall systems where the magnetic field is of the order of Tesla and a Bloch band would break into many subbands. The translational symmetry and the topological property of the subband are very different from those of the usual Bloch band. To distinguish between the two, the former is called the magnetic Bloch band (MBB).

The MBB usually carries nonzero quantum Hall conductance and has a nontrivial topology. Compared to the usual Bloch band, the MBB is a more interesting playground for many physics phenomena. In fact, nontrivial

topology of magnetic Bloch state was first revealed in the MBB (Thouless *et al.*, 1982). In this section, we review some basic facts of the MBB, as well as the semiclassical dynamics of the magnetic Bloch electron when it is subject to *further* electromagnetic perturbation (Chang and Niu, 1995, 1996). Such a formulation provides a clear picture of the hierarchical subbands split by the strong magnetic field [called the Hofstadter spectrum (Hofstadter, 1976), which could also be realized in a Bose-Einstein condensate, e.g., see Umucalılar *et al.* (2008)].

A. Magnetic translational symmetry

In the presence of a strong magnetic field, one needs to treat the magnetic field and the lattice potential on equal footing and solve the following Schrodinger equation:

$$\left\{ \frac{1}{2m}[\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r})]^2 + V_L(\boldsymbol{r}) \right\} \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \qquad (8.1)$$

where V_L is the periodic lattice potential. For convenience of discussion, we assume the magnetic field is uniform along the z axis and the electron is confined to the x-y plane. Because of the vector potential, the Hamiltonian H above no longer has the lattice translation symmetry.

Since the lattice symmetry of the charge density is not broken by an uniform magnetic field, one should be able to define translation operators that differ from the usual ones only by phase factors (Lifshitz and Landau, 1980). First, consider a system translated by a lattice vector $\boldsymbol{a}$,

$$\left\{ \frac{1}{2m}[\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r} + \boldsymbol{a})]^2 + V_L(\boldsymbol{r}) \right\} \psi(\boldsymbol{r} + \boldsymbol{a}) = E\psi(\boldsymbol{r} + \boldsymbol{a}), \qquad (8.2)$$

where $V_L(\boldsymbol{r} + \boldsymbol{a}) = V_L(\boldsymbol{r})$ has been used. One can write

$$\boldsymbol{A}(\boldsymbol{r} + \boldsymbol{a}) = \boldsymbol{A}(\boldsymbol{r}) + \nabla f(\boldsymbol{r}), \qquad (8.3)$$

where $\nabla f(\boldsymbol{r}) = \boldsymbol{A}(\boldsymbol{r} + \boldsymbol{a}) - \boldsymbol{A}(\boldsymbol{r}) \equiv \Delta \boldsymbol{A}(\boldsymbol{a})$. In a uniform magnetic field, one can choose a gauge such that $\boldsymbol{A}$ is perpendicular to the magnetic field $B\hat{z}$ and its components linear in x and y. As a result, $\Delta \boldsymbol{A}$ is independent of $\boldsymbol{r}$ and $f = \Delta \boldsymbol{A} \cdot \boldsymbol{r}$. The extra vector potential ∇f can be removed by a gauge transformation,

$$\left\{ \frac{1}{2m}[\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r})]^2 + V_L(\boldsymbol{r}) \right\} e^{i(e/\hbar)f} \psi(\boldsymbol{r} + \boldsymbol{a})$$

$$= E e^{i(e/\hbar)f} \psi(\boldsymbol{r} + \boldsymbol{a}). \qquad (8.4)$$

We now identify the state above as the magnetic translated state $T_{\boldsymbol{a}}\psi(\boldsymbol{r})$,

$$T_{\boldsymbol{a}}\psi(\boldsymbol{r}) = e^{i(e/\hbar)\Delta \boldsymbol{A} \cdot \boldsymbol{r}} \psi(\boldsymbol{r} + \boldsymbol{a}). \qquad (8.5)$$

The operator $T_{\boldsymbol{a}}$ being defined this way has the desired property that $[H, T_{\boldsymbol{a}}] = 0$.
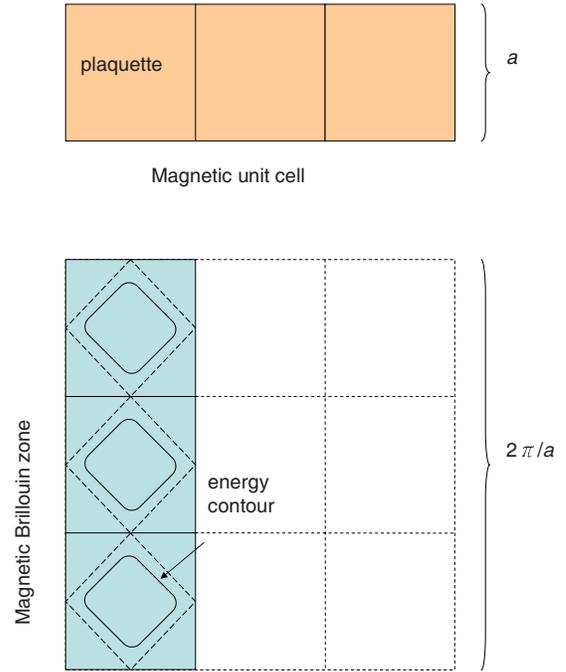
FIG. 12. (Color online) When the magnetic flux per plaquette is $\phi_0/3$, the magnetic unit cell is composed of three plaquettes. The magnetic Brillouin zone is three times smaller than the usual Brillouin zone. Furthermore, the magnetic Bloch states are threefold degenerate.

Unlike usual translation operators, magnetic translations along different directions usually do not commute. For example, let $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ be lattice vectors, then

$$T_{\boldsymbol{a}_2} T_{\boldsymbol{a}_1} = T_{\boldsymbol{a}_1} T_{\boldsymbol{a}_2} \exp\left(i\frac{e}{\hbar} \oint \boldsymbol{A} \cdot d\boldsymbol{r} \right), \qquad (8.6)$$

where $\oint \boldsymbol{A} \cdot d\boldsymbol{r}$ is the magnetic flux going through the unit cell defined by $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. That is, the noncommutativity is a result of the Aharonov-Bohm phase. $T_{\boldsymbol{a}_1}$ commutes with $T_{\boldsymbol{a}_2}$ only if the flux ϕ is an integer multiple of the flux quantum $\phi_0 = e/h$.

If the magnetic flux ϕ enclosed by a plaquette is $(p/q)\phi_0$, where p and q are coprime integers, then $T_{q\boldsymbol{a}_1}$ would commute with $T_{\boldsymbol{a}_2}$ (see Fig. 12). The simultaneous eigenstate of H, $T_{q\boldsymbol{a}_1}$, and $T_{\boldsymbol{a}_2}$ is called a magnetic Bloch state and its energy the magnetic Bloch energy,

$$H\psi_{n\boldsymbol{k}} = E_{n\boldsymbol{k}}\psi_{n\boldsymbol{k}}, \qquad (8.7)$$

$$T_{q\boldsymbol{a}_1}\psi_{n\boldsymbol{k}} = e^{i\boldsymbol{k} \cdot q\boldsymbol{a}_1}\psi_{n\boldsymbol{k}}, \qquad (8.8)$$

$$T_{\boldsymbol{a}_2}\psi_{n\boldsymbol{k}} = e^{i\boldsymbol{k} \cdot \boldsymbol{a}_2}\psi_{n\boldsymbol{k}}. \qquad (8.9)$$

Since the magnetic unit cell is q times larger than the usual unit cell, the magnetic Brillouin zone (MBZ) has to be q times smaller. If $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are defined as the lattice vectors reciprocal to $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, then, in this example, the MBZ is folded back q times along the $\boldsymbol{b}_1$ direction.

In addition, with the help of Eqs. (8.6) and (8.8), one can show that the eigenvalues of the $T_{\mathbf{a}_2}$ operator for the following translated states,

$$T_{\mathbf{a}_1}\psi_{n\mathbf{k}}, T_{2\mathbf{a}_1}\psi_{n\mathbf{k}}, \ldots, T_{(q-1)\mathbf{a}_1}\psi_{n\mathbf{k}}, \tag{8.10}$$

are

$$e^{i(\mathbf{k}+\mathbf{b}_2 p/q)\cdot\mathbf{a}_2}, e^{i(\mathbf{k}+2\mathbf{b}_2 p/q)\cdot\mathbf{a}_2}, \ldots, e^{i[\mathbf{k}+(q-1)\mathbf{b}_2 p/q]\cdot\mathbf{a}_2}, \tag{8.11}$$

respectively. These states are not equivalent, but have the same energy as $\psi_{n\mathbf{k}}$ since $[H, T_{\mathbf{a}_1}] = 0$. Therefore, the MBZ has a q-fold degeneracy along the $\mathbf{b}_2$ direction. Each repetition unit in the MBZ is sometimes called a reduced magnetic Brillouin zone. More discussions on the magnetic translation group can be found in Zak (1964a, 1964b, 1964c).

B. Basics of magnetic Bloch band

In this section we review some basic properties of the magnetic Bloch band. These include the pattern of band splitting due to a quantizing magnetic field, the phase of the magnetic Bloch state, and its connection with the Hall conductance.

The rules of band splitting are simple in two opposite limits, which are characterized by the relative strength between the lattice potential and the magnetic field. When the lattice potential is much stronger than the magnetic field, it is more appropriate to start with the zero-field Bloch band as a reference. It was found that if each plaquette encloses a magnetic flux $(p/q)\phi_0$, then each Bloch band would split to q subbands (Obermair and Wannier, 1976; Schellnhuber and Obermair, 1980; Wannier, 1980; Kohmoto, 1989; Hatsugai and Kohmoto, 1990). We know that if N is the total number of lattice sites on the two-dimensional plane, then the number of allowed states in the Brillouin zone (and in one Bloch band) is N. Since the area of the MBZ (and the number of states within) is smaller by a factor of q, each MBB has N/q states, sharing the number of states of the original Bloch band equally.

On the other hand, if the magnetic field is much stronger than the lattice potential, then one should start from the Landau level as a reference. In this case, if each plaquette has a magnetic flux $\phi = (p/q)\phi_0$, then after turning on the lattice potential each Landau level (LL) will split to p subbands. The state counting is quite different from the previous case: The degeneracy of the original LL is $\Phi/\phi_0 = Np/q$, where $\Phi = N\phi$ is the total magnetic flux through the two-dimensional sample. Therefore, after splitting, each MBB again has only N/q states, the number of states in a MBZ.

Between the two limits, when the magnetic field is neither very strong nor very weak, the band splitting does not follow a simple pattern. When the field is tuned from weak to strong, the subbands will split, merge, and interact with each other in a complicated manner, such that in the end there are only p subbands in the strong-field limit.

According to Laughlin's gauge-invariance argument (Laughlin, 1981), each of the isolated magnetic Bloch bands carries a quantized Hall conductivity (see Secs. II.B and III.C). This is closely related to the nontrivial topological property of the magnetic Bloch state (Kohmoto, 1985; Morandi, 1988). Furthermore, the distribution of Hall conductivities among the split subbands follows a simple rule first discovered by Thouless *et al.* (1982). This rule can be derived with the help of the magnetic translation symmetry (Dana *et al.*, 1985). We show the derivation below following the analysis of Dana *et al.* since it reveals the important role played by the Berry phase in the magnetic Bloch state.

In general, the phases of Bloch states at different $\mathbf{k}$'s are unrelated and can be defined independently.[14] However, the same does not apply to a MBZ. For one thing, the phase has to be nonintegrable in order to account for the Hall conductivity. One way to assign the phase of the MBS $u_{\mathbf{k}}(\mathbf{r})$ is by imposing the parallel-transport condition [see Thouless's article in Prange and Girvin (1987)],

$$\langle u_{k_1 0}|\frac{\partial}{\partial k_1}|u_{k_1 0}\rangle = 0, \tag{8.12}$$

$$\langle u_{k_1 k_2}|\frac{\partial}{\partial k_2}|u_{k_1 k_2}\rangle = 0. \tag{8.13}$$

The first equation defines the phase of the states on the k_1 axis; the second equation defines the phase along a line with fixed k_1 [see Fig. 13(a)]. As a result, the phases of any two states in the MBZ have a definite relation.

The states on opposite sides of the MBZ boundaries represent the same physical state. Therefore, they can only differ by a phase factor. Following Eqs. (8.12) and (8.13), we have

$$u_{k_1 + b_1/q, k_2} = u_{k_1, k_2}, \tag{8.14}$$

$$u_{k_1, k_2 + b_2} = e^{i\delta(k_1)} u_{k_1, k_2}, \tag{8.15}$$

where b_1 and b_2 are the lengths of the primitive vectors reciprocal to $\mathbf{a}_1$ and $\mathbf{a}_2$. That is, the states on the opposite sides of the k_1 boundaries have the same phase. The same cannot also be true for the k_2 boundaries, otherwise the topology will be too trivial to accommodate the quantum Hall conductivity.

Periodicity of the MBZ requires that

$$\delta(k_1 + b_1/q) = \delta(k_1) + 2\pi \times \text{integer}. \tag{8.16}$$

In order for the integral $(1/2\pi)\oint_{\partial\text{MBZ}} d\mathbf{k} \cdot \mathcal{A}(\mathbf{k})$ (which is nonzero only along the upper horizontal boundary) to be the Hall conductivity σ_H (in units of h/e^2), the integer in Eq. (8.16) obviously has to be equal to σ_H.

Following the periodicity condition in Eq. (8.16), it is possible to assign the phase in the form

---

[14]In practice, the phases are usually required to be continuous and differentiable so that the Wannier function can behave well.
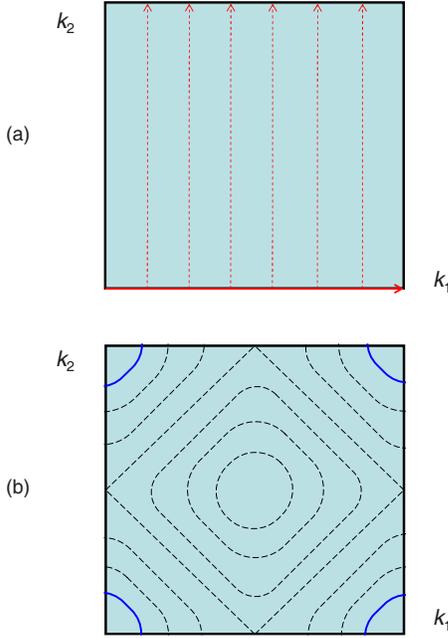
FIG. 13. (Color online) Reduced magnetic Brillouin zone. (a) The phases of the MBS in the reduced MBZ can be assigned using the parallel transport conditions, first along the k_1 axis, then along the paths parallel to the k_2 axis. (b) Hyperorbits in a reduced MBZ. Their sizes are quantized following the Bohr-Sommerfeld quantization condition. The orbit enclosing the largest area is indicated by solid lines.

$$\delta(k_1) = \tilde{\delta}(k_1) + \sigma_H k_1 q a_1, \tag{8.17}$$

where $\tilde{\delta}(k_1 + b_1/q) = \tilde{\delta}(k_1)$. On the other hand, from the discussion at the end of Sec. VIII.A, we know that

$$T_{a_1} u_{k_1 k_2} = e^{i\theta(k_1)} u_{k_1 k_2 + 2\pi p/qa_2}. \tag{8.18}$$

Again from the periodicity of the MBZ, one has

$$\theta(k_1 + b_1/q) = \theta(k_1) + 2\pi m, \quad m \in Z, \tag{8.19}$$

which gives

$$\theta(k_1) = \tilde{\theta}(k_1) + m k_1 q a_1. \tag{8.20}$$

Choosing $\tilde{\delta}(k_1)$ and $\tilde{\theta}(k_1)$ to be 0, one finally gets

$$\begin{aligned} T_{qa_1} u_{k_1 k_2} &= e^{iqm k_1 qa_1} u_{k_1 k_2 + 2\pi p/a_2} \\ &= e^{iqm k_1 qa_1} e^{ip\sigma_H k_1 qa_1} u_{k_1 k_2}. \end{aligned} \tag{8.21}$$

But this state should also be equal to $e^{iqk_1 a_1} u_{k_1 k_2}$. Therefore, the Hall conductivity should satisfy

$$p\sigma_H + qm = 1. \tag{8.22}$$

This equation determines the Hall conductivity (mod q) of a MBB (Dana *et al.*, 1985). In Sec. VIII.D, we will see that the semiclassical analysis can also help us finding out the Hall conductivity of a MBB.

C. Semiclassical picture: Hyperorbits

When a weak magnetic field is applied to a Bloch band, the electron experiences a Lorentz force and executes a cyclotron motion on the surface of the Fermi sea. In the case of the MBB, the magnetic field B_0 changes the band structure itself. On the other hand, the *magnetic* quasimomentum $\hbar k$ is a good quantum number with $\hbar \dot{k} = 0$. Therefore, there is no cyclotron motion of k (even though there is a magnetic field B_0). Similar to the case of the Bloch band, one can construct a wave packet out of the magnetic Bloch states, and study its motion in both r and k spaces when it is subject to an additional weak electromagnetic field E and δB. The semiclassical equations of motion that are valid under the one-band approximation have exactly the same form as Eq. (5.8). One simply needs to reinterpret $\hbar k$, $E_0(k)$, and B in Eq. (5.8) as the magnetic momentum, the magnetic band energy, and the extra magnetic field δB, respectively (Chang and Niu, 1995, 1996). As a result, when δB is not zero, there exists similar circulating motion in the MBB. This type of orbit will be called "hyperorbit."

We first consider the case without the electric field (the case with both E and δB will be considered in Sec. VIII.D). By combining the following two equations of motion [cf. Eq. (5.8)],

$$\hbar \dot{k} = -e \dot{r} \times \delta B, \tag{8.23}$$

$$\hbar \dot{r} = \frac{\partial E}{\partial k} - \hbar \dot{k} \times \Omega, \tag{8.24}$$

one has

$$\hbar \dot{k} = -\frac{1}{\kappa} \frac{\partial E}{\partial k} \times \delta B \frac{e}{\hbar}, \tag{8.25}$$

where $\kappa(k) = 1 + \Omega(k)\delta B e/\hbar$. This determines the k orbit moving along a path with constant $E(k) = E_0(k) - M(k) \cdot \delta B$, which is the magnetic Bloch band energy shifted by the magnetization energy. Similar to the Bloch band case, it is not difficult to see from Eq. (8.23) that the r orbit is simply the k orbit rotated by $\pi/2$ and (linearly) scaled by the factor $\hbar/e\delta B$. These orbits in the MBB and their real-space counterparts are the hyperorbits mentioned above (Chambers, 1965).

The size of a real-space hyperorbit may be very large (if phase coherence can be maintained during the circulation) since it is proportional to the inverse of the residual magnetic field δB. Furthermore, since the split magnetic subband is narrower and flatter than the original Bloch band, the electron group velocity is small. As a result, the frequency of the hyperorbit motion can be very low. Nevertheless, it is possible to detect the hyperorbit using, for example, resonant absorption of ultrasonic wave or conductance oscillation in an electron focusing device.

Similar to the cyclotron orbit, the hyperorbit motion can also be quantized using the Bohr-Sommerfeld quan-

tization rule [see Eq. (7.3)]. One only needs to bear in mind that k is confined to the smaller MBZ and the magnetic field in Eq. (7.3) should be δB. After the quantization, there can only be a finite number of hyperorbits in the MBZ. The area of the largest hyperorbit should be equal to or slightly smaller (assuming $\delta B \ll B_0$ so that the number of hyperorbits is large) than the area of the MBZ $(2\pi/a)^2/q$ [see Fig. 13(b)]. For such an orbit, the Berry phase correction $\Gamma/2\pi$ in Eq. (7.3) is very close to the integer Hall conductivity σ_H of the MBB. Therefore, it is not difficult to see that the number of hyperorbits should be $|1/(q\,\delta\phi)+\sigma_H|$, where $\delta\phi \equiv \delta B a^2/\phi_0$ is the residual flux per plaquette.

Because the MBZ is q-fold degenerate (see Sec. VIII.A), the number of energy levels produced by these hyperorbits are (Chang and Niu, 1995)

$$D = \frac{|1/(q\,\delta\phi) + \sigma_H|}{q}. \tag{8.26}$$

If one further takes the tunneling between degenerate hyperorbits into account (Wilkinson, 1984a), then each energy level will be broadened into an energy band. These are the magnetic energy subbands at a finer energy scale compared to the original MBB.

D. Hall conductivity of hyperorbit

According to Laughlin's argument, each of the isolated subband should have its own integer Hall conductivity. That is, as a result of band splitting, the integer Hall conductivity σ_H of the parent band is split to a distribution of integers σ_r (there are q of them). The sum of these integers should be equal to the original Hall conductivity: $\sigma_H = \Sigma_r \sigma_r$. There is a surprisingly simple way to determine this distribution using the semiclassical formulation: one only needs to study the response of the hyperorbit to an electric field.

After adding a term $-eE$ to Eq. (8.23), one obtains

$$\dot{r} = \frac{\hbar}{e\,\delta B}\dot{k} \times \hat{z} + \frac{E \times \hat{z}}{\delta B}. \tag{8.27}$$

For a closed orbit, this is just a cyclotron motion superimposed with a drift along the $E \times \delta B$ direction. After time average, the former does not contribute to a net transport. Therefore the Hall current density for a filled magnetic band in a clean sample is

$$J_H = -e\int_{MBZ}\frac{d^2k}{(2\pi)^2}\dot{r} = -e\rho\frac{E \times \hat{z}}{\delta B}, \tag{8.28}$$

where ρ is the number of states in the MBZ divided by the sample area. Therefore, the Hall conductivity is $\sigma_r^{close} = e\rho/\delta B$. If the areal electron density of a sample is ρ_0, then after applying a flux $\phi = p/q$ per plaquette the MBZ shrinks by q times and $\rho = \rho_0/q$.

How can one be sure that both the degeneracy in Eq. (8.26) and the Hall conductivity σ_r^{close} are integers? This is closely related to the following question: How does one divide an uniform magnetic field B into the quantiz-

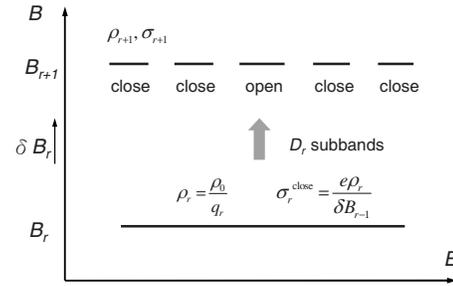FIG. 14. A parent magnetic Bloch band at magnetic field B_r splits to D_r subbands ($D_r = 5$ here) due to a perturbation δB_{r+1}. The subbands near the band edges of the parent band are usually originated from closed hyperorbits. The subband in the middle is from an open hyperorbit.

ing part B_0 and the perturbation δB? The proper way to separate them was first proposed by Azbel (1964). Since then, such a recipe has been used widely in the analysis of the Hofstadter spectrum (Hofstadter, 1976).

One first expands the flux $\phi = p/q(<1)$ as a continued fraction,

$$\frac{p}{q} = \cfrac{1}{f_1 + \cfrac{1}{f_2 + \cfrac{1}{f_3 + \cfrac{1}{\cdots}}}} \equiv [f_1, f_2, f_3, \ldots], \tag{8.29}$$

then the continued fraction is truncated to obtain various orders of approximate magnetic flux. For example, $\phi_1 = [f_1] \equiv p_1/q_1$, $\phi_2 = [f_1, f_2] \equiv p_2/q_2$, $\phi_3 = [f_1, f_2, f_3] \equiv p_3/q_3, \ldots$, etc. What is special about these truncations is that p_r/q_r is the closest approximation to p/q among all fractions with $q \leq q_r$ (Khinchin, 1964).

As a reference, we show two identities that will be used below:

$$q_{r+1} = f_{r+1}q_r + q_{r-1}, \tag{8.30}$$

$$p_{r+1}q_r - p_r q_{r+1} = (-1)^r. \tag{8.31}$$

According to desired accuracy, one chooses a particular ϕ_r to be the quantizing flux, and takes $\delta\phi_r \equiv \phi_{r+1} - \phi_r$ as a perturbation (see Fig. 14). With the help of Eq. (8.31) one has

$$\delta\phi_r = \frac{(-1)^r}{q_r q_{r+1}}. \tag{8.32}$$

As a result, the Hall conductivity for a closed hyperorbit produced by $\delta B_{r-1} \equiv \delta\phi_{r-1}/a^2$ is (recall that $\rho_r = \rho_0/q_r$)

$$\sigma_r^{close} = \frac{e\rho_r}{\delta B_{r-1}} = (-1)^{r-1}q_{r-1}. \tag{8.33}$$

Substituting this value back into Eq. (8.26) for D_r^{close} (the number of subbands split by $\delta\phi_r$), and using Eq. (8.30), one has

$$D_r^{\text{close}} = \frac{|1/(q_r \delta\varphi_r) + \sigma_r^{\text{close}}|}{q_r} = f_{r+1}. \qquad (8.34)$$

This is the number of subbands split from a parent band that is originated from a closed hyperorbit. One can see that the Hall conductivity and the number of splitting subbands are indeed integers.

For lattices with square or triangular symmetry, there is one and only one nesting (open) hyperorbit in the MBZ [for example, see the diamond-shaped energy contour in Fig. 13(b)]. Because of its open trajectory, the above analysis fails for the nesting orbit since the first term in Eq. (8.27) also contributes to the Hall conductivity. However, since the total number of hyperorbits in the parent band can be determined by the quantization rule, we can easily pin down the value of σ_r^{open} with the help of the sum rule $\sigma_H^{\text{parent}} = \Sigma_r \sigma_r$. Furthermore, D_r^{open} can be calculated from Eq. (8.26) once σ_r^{open} is known. Therefore, both the distribution of the σ_r's and the pattern of splitting can be determined entirely within the semiclassical formulation. The computation in principle can be carried out to all orders of r. Interested readers may consult Chang and Niu (1996) and Bohm *et al.* (2003) (Chap. 13) for more details.

IX. NON-ABELIAN FORMULATION

In previous sections we have considered the semiclassical electron dynamics with an Abelian Berry curvature. Such a formalism can be extended to the cases where the energy bands are degenerate or nearly degenerate (e.g., due to spin) (Culcer *et al.*, 2005; Shindou and Imura, 2005) [also see Strinati (1978), Gosselin *et al.* (2006), Gosselin *et al.* (2007), and Dayi (2008)]. Because the degenerate Bloch states have multiple components, the Berry curvature becomes a matrix with non-Abelian gauge structure. We report recent progress on requantizing the semiclassical theory that helps turning the wave-packet energy into an effective quantum Hamiltonian (Chang and Niu, 2008). After citing the dynamics of the Dirac electron as an example, this approach is applied to semiconductor electrons with spin degrees of freedom. Finally, we point out that the effective Hamiltonian is only part of an effective theory, and that the variables in the effective Hamiltonian are often gauge dependent and therefore cannot be physical variables. In order to have a complete effective theory, one also needs to identify the correct physical variables relevant to experiments.

A. Non-Abelian electron wave packet

The wave packet in an energy band with D-fold degeneracy is a superposition of the Bloch states ψ_{nq} (cf. Sec. IV),

$$|W\rangle = \sum_{n=1}^{D} \int d^3 q\, a(\boldsymbol{q},t)\, \eta_n(\boldsymbol{q},t)|\psi_{nq}\rangle, \qquad (9.1)$$

where $\Sigma_n |\eta_n(\boldsymbol{q},t)|^2 = 1$ at each $\boldsymbol{q}$ and $a(\boldsymbol{q},t)$ is a normalized distribution that centers at $\boldsymbol{q}_c(t)$. Furthermore, the wave packet is built to be localized at $\boldsymbol{r}_c$ in the $\boldsymbol{r}$ space. One can first obtain an effective Lagrangian for the wave-packet variables $\boldsymbol{r}_c$, $\boldsymbol{q}_c$, and η_n, then derive their dynamical equations of motion. Without going into details, we only review primary results of such an investigation (Culcer *et al.*, 2005).

Similar to the nondegenerate case, there are three essential quantities in such a formulation. In addition to the Bloch energy $E_0(\boldsymbol{q})$, there are the Berry curvature and the magnetic moment of the wave packet (see Sec. IV). However, because of the spinor degree of freedom, the latter two become vector-valued matrices instead of the usual vectors. The Berry connection becomes

$$\boldsymbol{R}_{mn}(\boldsymbol{q}) \equiv i\left\langle u_{mq} \left| \frac{\partial u_{nq}}{\partial \boldsymbol{q}} \right.\right\rangle. \qquad (9.2)$$

In the rest of this section, (bold-faced) calligraphic fonts are reserved for (vector-valued) matrices. Therefore, the Berry connection in Eq. (9.2) can simply be written as $\boldsymbol{\mathcal{R}}$.

The Berry curvature is defined as

$$\boldsymbol{\mathcal{F}}(\mathbf{q}) = \nabla_{\mathbf{q}} \times \boldsymbol{\mathcal{R}} - i\boldsymbol{\mathcal{R}} \times \boldsymbol{\mathcal{R}}. \qquad (9.3)$$

Recall that the Berry connection and Berry curvature in the Abelian case have the same mathematical structures as the vector potential and the magnetic field in electromagnetism. Here $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{F}}$ also have the same structure as the gauge potential and gauge field in the non-Abelian SU(2) gauge theory (Wilczek and Zee, 1984). Redefining the spinor basis $\{\psi_{nq}\}$ amounts to a gauge transformation. Assuming that the new basis is obtained from the old basis by a gauge transformation U, then $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{F}}$ would change in the following way:

$$\boldsymbol{\mathcal{R}}' = U\boldsymbol{\mathcal{R}}U^\dagger + i\frac{\partial U}{\partial \boldsymbol{\lambda}}U^\dagger,$$

$$\boldsymbol{\mathcal{F}}' = U\boldsymbol{\mathcal{F}}U^\dagger, \qquad (9.4)$$

where $\boldsymbol{\lambda}$ is the parameter of adiabatic change.

The magnetic moment of the wave packet can be found in Eq. (4.6). If the wave packet is narrowly distributed around $\boldsymbol{q}_c$, then it is possible to write it as the spinor average of the following quantity (Culcer *et al.*, 2005):

$$\boldsymbol{M}_{nl}(\boldsymbol{q}_c) = -i\frac{e}{2\hbar}\left\langle \frac{\partial u_n}{\partial \boldsymbol{q}_c} \right| \times [\tilde{H}_0 - E_0(\boldsymbol{q}_c)] \left| \frac{\partial u_l}{\partial \boldsymbol{q}_c} \right\rangle, \quad (9.5)$$

where $\tilde{H}_0 \equiv e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} H_0 e^{i\boldsymbol{q}\cdot\boldsymbol{r}}$. That is, $\boldsymbol{M} = \langle \boldsymbol{\mathcal{M}} \rangle = \boldsymbol{\eta}^\dagger \boldsymbol{\mathcal{M}} \boldsymbol{\eta} = \Sigma_{nl} \eta_n^* \boldsymbol{M}_{nl} \eta_l$. Except for the extension to multiple components, the form of the magnetic moment remains the same as its Abelian counterpart [see Eq. (4.6)].

As a reference, we write down the equations of motion for the non-Abelian wave packet (Culcer *et al.*, 2005),

$$\hbar \dot{\boldsymbol{k}}_c = -e\boldsymbol{E} - e\dot{\boldsymbol{r}}_c \times \boldsymbol{B}, \tag{9.6}$$

$$\hbar \dot{\boldsymbol{r}}_c = \left\langle \left[\frac{\mathcal{D}}{\mathcal{D}\boldsymbol{k}_c}, \mathcal{H} \right] \right\rangle - \hbar \dot{\boldsymbol{k}}_c \times \boldsymbol{F}, \tag{9.7}$$

$$i\hbar \dot{\boldsymbol{\eta}} = (-\boldsymbol{\mathcal{M}} \cdot \boldsymbol{B} - \hbar \dot{\boldsymbol{k}}_c \cdot \boldsymbol{\mathcal{R}}) \boldsymbol{\eta}, \tag{9.8}$$

where $\boldsymbol{k}_c = \boldsymbol{q}_c + (e/\hbar)\boldsymbol{A}(\boldsymbol{r}_c)$, $\boldsymbol{F} = \langle \boldsymbol{\mathcal{F}} \rangle$, and the covariant derivative $\mathcal{D}/\mathcal{D}\boldsymbol{k}_c \equiv \partial/\partial \boldsymbol{k}_c - i\mathcal{R}$. Again the calligraphic fonts represent matrices. A spinor average (represented by the angular bracket) is imposed on the commutator of $\mathcal{D}/\mathcal{D}\boldsymbol{k}_c$ and $\mathcal{H}$. The semiclassical Hamiltonian matrix inside the commutator in Eq. (9.7) is

$$\mathcal{H}(\boldsymbol{r}_c, \boldsymbol{k}_c) = E_0(\boldsymbol{k}_c) - e\phi(\boldsymbol{r}_c) - \boldsymbol{\mathcal{M}}(\boldsymbol{k}_c) \cdot \boldsymbol{B}. \tag{9.9}$$

The spinor-averaged Hamiltonian matrix is nothing but the wave-packet energy $E = \langle \mathcal{H} \rangle$. Like the Abelian case, it has three terms: the Bloch energy, the electrostatic energy, and the magnetization energy.

Compared to the Abelian case in Eq. (5.8), the $\dot{\boldsymbol{k}}_c$ equation also has the electric force and the Lorentz force. The $\dot{\boldsymbol{r}}_c$ equation is slightly more complicated: The derivative in the group velocity $\partial E/\partial \boldsymbol{k}_c$ is replaced by the commutator of the covariant derivative and $\mathcal{H}$. The anomalous velocity in Eq. (9.7) remains essentially the same. One only needs to replace the Abelian Berry curvature with the spinor average of the non-Abelian one.

Equation (9.8) governs the dynamics of the spinor, from which we can derive the equation for the spin vector $\dot{\boldsymbol{J}}$, where $\boldsymbol{J} = \langle \boldsymbol{\mathcal{J}} \rangle$ and $\boldsymbol{\mathcal{J}}$ is the spin matrix,

$$i\hbar \dot{\boldsymbol{J}} = \langle [\boldsymbol{\mathcal{J}}, \mathcal{H} - \hbar \dot{\boldsymbol{k}}_c \cdot \boldsymbol{\mathcal{R}}] \rangle. \tag{9.10}$$

The spin dynamics in Eq. (9.10) is influenced by the Zeeman energy in $\mathcal{H}$, as it should be. However, the significance of the other term that is proportional to the Berry connection is less obvious here. Later we will see that it is in fact the spin-orbit coupling.

B. Spin Hall effect

The anomalous velocity in Eq. (9.7) that is proportional to the Berry curvature $\boldsymbol{F}$ is of physical significance. We have seen earlier that it is the transverse current in the quantum Hall effect and the anomalous Hall effect (Sec. III). The latter requires electron spin with spin-orbit coupling and therefore the carrier dynamics is suitably described by Eqs. (9.6), (9.7), and (9.10).

For the non-Abelian case, the Berry curvature $\boldsymbol{F}$ is often proportional to the spin $\boldsymbol{S}$ (see Secs. IX.D and IX.E). If this is true, then in the presence of an electric field the anomalous velocity is proportional to $\boldsymbol{E} \times \boldsymbol{S}$. That is, the trajectories of spin-up and spin-down electrons are parted toward opposite directions transverse to the electric field. There can be a net transverse current if

the populations of spin-up and spin-down electrons are different, as in the case of a ferromagnet. This then leads to the anomalous Hall effect.

If the populations of different spins are equal, then the net electric Hall current is zero. However, the spin Hall current can still be nonzero. This is the source of the intrinsic spin Hall effect (SHE) in semiconductors predicted by Murakami *et al.* (2003). In the original proposal, a four-band Luttinger model is used to describe the heavy-hole (HH) bands and light-hole (LH) bands. The Berry curvature emerges when one restricts the whole Hilbert space to a particular (HH or LH) subspace. As shown in Sec. IX.E, such a projection of the Hilbert space almost always generates a Berry curvature. Therefore, the SHE should be common in semiconductors or other materials. Indeed, intrinsic SHE has also been theoretically predicted in metals (Guo *et al.*, 2008). The analysis of the SHE from the semiclassical point of view can also be found in Culcer *et al.* (2005).

In addition to the Berry curvature, impurity scattering is another source of the (extrinsic) SHE. This is first predicted by Dyakonov and Perel (1971a, 1971b) [see also Chazalviel (1975)] and the same idea is later revived by Hirsch (1999). Because of the spin-orbit coupling between the electron and the (spinless) impurity, the scattering amplitude is not symmetric with respect to the transverse direction. This is the same skew scattering (or Mott scattering) in AHE (see Sec. III.D.1).

To date most of the experimental evidences for the SHE belongs to the extrinsic case. They are first observed in semiconductors (Kato *et al.*, 2004; Sih *et al.*, 2005; Wunderlich *et al.*, 2005) and later in metals (Valenzuela and Tinkham, 2006; Kimura *et al.*, 2007; Seki *et al.*, 2008). The spin Hall conductivity in metals can be detected at room temperature and can be several orders of magnitude larger than that in semiconductors. Such a large effect could be due to the resonant Kondo scattering from the Fe impurities (Guo *et al.*, 2009). This subject is currently in rapid progress. Complete understanding of the intrinsic or extrinsic SHE is crucial to future devices that could generate a significant amount of spin current.

C. Quantization of electron dynamics

In Sec. VII we have introduced the Bohr-Sommerfeld quantization, which helps predicting quantized energy levels. Such a procedure applies to the Abelian case and is limited to closed orbits in phase space. In this section we report on the method of canonical quantization that applies to more general situations. With both the semiclassical theory and the method of requantization at hand, one can start from a quantum theory of general validity (such as the Dirac theory of electrons) and descend to an effective quantum theory with a smaller range of validity. Such a procedure can be applied iteratively to generate a hierarchy of effective quantum theories.

As mentioned in Sec. VII.D, even though a Hamiltonian system always admits canonical variables, it is not

always easy to find them. In the wave-packet theory, the variables $\boldsymbol{r}_c$ and $\boldsymbol{k}_c$ have clear physical meaning, but they are not canonical variables. The canonical variables $\boldsymbol{r}$ and $\boldsymbol{p}$ accurate to linear order of the fields are related to the center-of-mass variables as follows (Chang and Niu, 2008):

$$\boldsymbol{r}_c = \boldsymbol{r} + \boldsymbol{\mathcal{R}}(\boldsymbol{\pi}) + \boldsymbol{\mathcal{G}}(\boldsymbol{\pi}),$$

$$\hbar \boldsymbol{k}_c = \boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r}) + e\boldsymbol{B} \times \boldsymbol{\mathcal{R}}(\boldsymbol{\pi}),$$

(9.11)

where $\boldsymbol{\pi} = \boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r})$ and $\mathcal{G}_\alpha(\boldsymbol{\pi}) \equiv (e/\hbar)(\boldsymbol{\mathcal{R}} \times \boldsymbol{B}) \cdot \partial \boldsymbol{\mathcal{R}}/\partial \pi_\alpha$. This is a generalization of the Peierls substitution to the non-Abelian case. The last terms in both equations can be neglected on some occasions. For example, they will not change the force and the velocity in Eqs. (9.6) and (9.7).

When expressed in the new variables, the semiclassical Hamiltonian in Eq. (9.9) can be written as

$$\mathcal{H}(\boldsymbol{r},\boldsymbol{p}) = E_0(\boldsymbol{\pi}) - e\phi(\boldsymbol{r}) + e\boldsymbol{E} \cdot \boldsymbol{\mathcal{R}}(\boldsymbol{\pi})$$
$$- \boldsymbol{B} \cdot \left[\boldsymbol{\mathcal{M}}(\boldsymbol{\pi}) - e\boldsymbol{\mathcal{R}} \times \frac{\partial E_0}{\partial \boldsymbol{\pi}} \right],$$

(9.12)

where we have used the Taylor expansion and neglected terms nonlinear in fields. Finally, one promotes the canonical variables to quantum conjugate variables and converts $\mathcal{H}$ to an effective quantum Hamiltonian.

Compared to the semiclassical Hamiltonian in Eq. (9.9), the quantum Hamiltonian has two additional terms from the Taylor expansion. The dipole-energy term $e\boldsymbol{E} \cdot \boldsymbol{\mathcal{R}}$ is originated from the shift between the charge center $\boldsymbol{r}_c$ and the center of the canonical variable $\boldsymbol{r}$. Although the exact form of the Berry connection $\boldsymbol{\mathcal{R}}$ depends on the physical model, we show that for both the Dirac electron (Sec. IX.D) and the semiconductor electron (Sec. IX.E) the dipole term is closely related to the spin-orbit coupling. The correction to the Zeeman energy is sometimes called the Yafet term, which vanishes near a band edge (Yafet, 1963).

Three remarks are in order. First, the form of the Hamiltonian, especially the spin-orbit term and Yafet term, is clearly gauge dependent because of the gauge-dependent Berry connection. Such gauge dependence has prevented one from assigning a clear physical significance to the Yafet term. For that matter, it is also doubtful that the electric dipole, or the spin-orbit energy, can be measured independently. Second, in a neighborhood of a k point, one can always choose to work within a particular gauge. However, if the first Chern number (or its non-Abelian generalization) is not zero, one cannot choose a global gauge in which $\boldsymbol{\mathcal{R}}$ is smooth everywhere in the Brillouin zone. In such a nontrivial topological situation one has to work with patches of the Brillouin zone for a single canonical quantum theory. Third, the semiclassical theory based on the variables $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{M}}$, on the other hand, is gauge independent. Therefore, the effective quantum theory can be smooth globally.

D. Dirac electron

To illustrate the application of the non-Abelian wave-packet theory and its requantization, we use the Dirac electron as an example. The starting quantum Hamiltonian is

$$H = c\boldsymbol{\alpha} \cdot (\boldsymbol{p} + e\boldsymbol{A}) + \beta mc^2 - e\phi(\boldsymbol{r})$$
$$= H_0 + ce\boldsymbol{\alpha} \cdot \boldsymbol{A} - e\phi(\boldsymbol{r}),$$

(9.13)

where $\boldsymbol{\alpha}$ and β are the Dirac matrices (Strange, 1998) and H_0 is the free-particle Hamiltonian. The energy spectrum of H_0 has positive-energy branch and negative-energy branch, each with twofold degeneracy due to the spin. These two branches are separated by a large energy gap mc^2. One can construct a wave packet out of the positive-energy eigenstates and study its dynamics under the influence of an external field. The size of the wave packet can be as small as the Compton wavelength $\lambda_c = \hbar/mc$ (but not smaller), which is two orders of magnitude smaller than the Bohr radius. Therefore, the adiabatic condition on the external electromagnetic field can be easily satisfied: the spatial variation in the potential only needs to be much smoother than λ_c. In this case, even the lattice potential in a solid can be considered as a semiclassical perturbation. Furthermore, because of the large gap between branches, interbranch tunneling happens (and the semiclassical theory fails) only if the field is so strong that electron-positron pair production can no longer be ignored.

Since the wave packet is living on a branch with twofold degeneracy, the Berry connection and curvature are 2×2 matrices (Chang and Niu, 2008),

$$\boldsymbol{\mathcal{R}}(\boldsymbol{q}) = \frac{\lambda_c^2}{2\gamma(\gamma+1)} \boldsymbol{q} \times \boldsymbol{\sigma},$$

(9.14)

$$\boldsymbol{\mathcal{F}}(\boldsymbol{q}) = -\frac{\lambda_c^2}{2\gamma^3} \left(\boldsymbol{\sigma} + \lambda_c^2 \frac{\boldsymbol{q} \cdot \boldsymbol{\sigma}}{\gamma+1} \boldsymbol{q} \right),$$

(9.15)

where $\gamma(q) \equiv \sqrt{1 + (\hbar q/mc)^2}$ is the relativistic dilation factor. To calculate these quantities, we only need the free particle eigenstates of H_0 [see Eqs. (9.2) and (9.3)]. That is, the nontrivial gauge structure exists in the free particle already.

It may come as a surprise that the free wave packet also possesses an intrinsic magnetic moment. Straightforward application of Eq. (9.5) gives (Chuu *et al.*, 2010)

$$\boldsymbol{\mathcal{M}}(\boldsymbol{q}) = -\frac{e\hbar}{2m\gamma^2} \left(\boldsymbol{\sigma} + \lambda_c^2 \frac{\boldsymbol{q} \cdot \boldsymbol{\sigma}}{\gamma+1} \boldsymbol{q} \right).$$

(9.16)

This result agrees with the one calculated from the abstract spin operator $\hat{\boldsymbol{S}}$ in the Dirac theory (Chuu *et al.*, 2010),

$$\boldsymbol{\mathcal{M}}(\boldsymbol{q}) = -g \frac{e\hbar}{2m\gamma(\boldsymbol{q})} \langle W | \hat{\boldsymbol{S}} | W \rangle,$$

(9.17)

in which the g factor is 2. The Zeeman coupling in the wave-packet energy is $-\boldsymbol{\mathcal{M}} \cdot \boldsymbol{B}$. Therefore, this magnetic moment gives the correct magnitude of the Zeeman en-

ergy with the correct g factor. Recall that Eq. (9.5) is originated from Eq. (4.6), which is the magnetic moment due to a circulating charge current. Therefore, the magnetic moment here indeed is a result of the spinning wave packet.

The present approach is a revival of Uhlenbeck and Goudsmit's rotating sphere model of the electron spin but without its problem. The size of the wave packet λ_c constructed from the positive-energy states is two orders of magnitude larger than the classical electron radius e^2/mc^2. Therefore, the wave packet does not have to rotate faster than the speed of light to have the correct magnitude of spin. This semiclassical model for spin is pleasing since it gives a clear and heuristic picture of the electron spin. Also, one does not have to resort to the more complicated Foldy-Wouthuysen approach to extract the spin from the Dirac Hamiltonian (Foldy and Wouthuysen, 1950).

From the equation of motion in Eq. (9.10), one obtains

$$\langle \dot{\boldsymbol{\sigma}} \rangle = \frac{e}{\gamma m} \left[\boldsymbol{B} + \boldsymbol{E} \times \frac{\hbar \boldsymbol{k}_c}{(\gamma + 1)mc^2} \right] \times \langle \boldsymbol{\sigma} \rangle. \tag{9.18}$$

This is the Bargmann-Michel-Telegdi equation for a relativistic electron (Bargmann et al., 1959). More discussions on the equations of motion for $\boldsymbol{r}_c$ and $\boldsymbol{k}_c$ can be found in Chang and Niu (2008).

Finally, substituting the Berry connection and the magnetic moment into Eq. (9.12) and using $E_0(\boldsymbol{\pi}) = \sqrt{c^2\boldsymbol{\pi}^2 + m^2c^4}$, one can obtain the effective quantum Hamiltonian,

$$\mathcal{H}(\boldsymbol{r},\boldsymbol{p}) = \gamma(\boldsymbol{\pi})mc^2 - e\phi(\boldsymbol{r}) + \frac{\mu_B}{\gamma(\gamma+1)} \frac{\boldsymbol{\pi}}{mc^2} \times \boldsymbol{\sigma} \cdot \boldsymbol{E}$$

$$+ \frac{\mu_B}{\gamma} \boldsymbol{\sigma} \cdot \boldsymbol{B}, \tag{9.19}$$

in which all γ's are functions of $\boldsymbol{\pi}$ and $\mu_B = e\hbar/2m$. This is the relativistic Pauli Hamiltonian. At low velocity, $\gamma \simeq 1$, and it reduces to the more familiar form. Notice that the spin-orbit coupling comes from the dipole energy term with the Berry connection, as mentioned earlier [also see Mathur (1991) and Shankar and Mathur (1994)].

E. Semiconductor electron

When studying the transport properties of semiconductors, one often only focuses on the carriers near the fundamental gap at the Γ point. In this case, the band structure far away from this region is not essential. It is then a good approximation to use the $\boldsymbol{k} \cdot \boldsymbol{p}$ expansion and obtain the four-band Luttinger model or the eight-band Kane model (Luttinger, 1951; Kane, 1957; Winkler, 2003) to replace the more detailed band structure (see Fig. 15). In this section, we start from the eight-band Kane model and study the wave-packet dynamics in one of its subspace: the conduction band. It is also possible to investigate the wave-packet dynamics in other subspaces: the
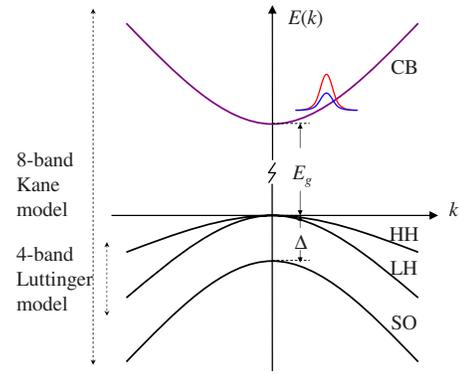

FIG. 15. (Color online) Schematic of the semiconductor band structure near the fundamental gap. The wave packet in the conduction band is formed from a two-component spinor.

HH-LH complex or the spin-orbit split-off band. The result of the latter is not reported in this review. Interested readers can consult Chang and Niu (2008) for more details, including the explicit form of the Kane Hamiltonian that the calculations are based upon.

To calculate the Berry connection in Eq. (9.2), one needs to obtain the eigenstates of the Kane model, which have eight components. Similar to the positive-energy branch of the Dirac electron, the conduction band is twofold degenerate. Detailed calculation shows that, to linear order in k [and up to a SU(2) gauge rotation], the Berry connection is a 2×2 matrix of the form

$$\mathcal{R} = \frac{V^2}{3} \left[\frac{1}{E_g^2} - \frac{1}{(E_g + \Delta)^2} \right] \boldsymbol{\sigma} \times \boldsymbol{k}, \tag{9.20}$$

where E_g is the fundamental gap, Δ is the spin-orbit spit-off gap, and $V = (\hbar/m_0)\langle S|\hat{p}_x|X \rangle$ is a matrix element of the momentum operator.

As a result, the dipole term $e\boldsymbol{E} \cdot \mathcal{R}$ becomes

$$\mathcal{H}_{SO} = e\boldsymbol{E} \cdot \mathcal{R} = \alpha \boldsymbol{E} \cdot \boldsymbol{\sigma} \times \boldsymbol{k}, \tag{9.21}$$

where $\alpha \equiv (eV^2/3)[1/E_g^2 - 1/(E_g+\Delta)^2]$. The coefficient α and the form of the spin-orbit coupling are the same as the Rashba coupling (Rashba, 1960; Bychkov and Rashba, 1984). However, unlike the usual Rashba coupling that requires structural inversion asymmetry to generate an internal field, this term exists in a bulk semiconductor with inversion symmetry but requires an external field $\boldsymbol{E}$.

From the Berry connection, we can calculate the Berry curvature in Eq. (9.3) to the leading order of $\boldsymbol{k}$ as

$$\mathcal{F} = \frac{2V^2}{3} \left[\frac{1}{E_g^2} - \frac{1}{(E_g + \Delta)^2} \right] \boldsymbol{\sigma}. \tag{9.22}$$

In the presence of an electric field, this would generate the transverse velocity in Eq. (9.7),

$$\boldsymbol{v}_T = 2e\alpha \boldsymbol{E} \times \langle \boldsymbol{\sigma} \rangle. \tag{9.23}$$

As a result, spin-up and spin-down electrons move toward opposite directions, which results in a spin-Hall effect (see Sec. IX.B for related discussion).

The wave packet in the conduction band also spontaneously rotates with respect to its own center of mass. To the lowest order of $\boldsymbol{k}$, it has the magnetic moment,

$$\boldsymbol{\mathcal{M}} = \frac{eV^2}{3\hbar}\left(\frac{1}{E_g} - \frac{1}{E_g + \Delta}\right)\boldsymbol{\sigma}. \tag{9.24}$$

With these three basic quantities, $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\mathcal{F}}$, and $\boldsymbol{\mathcal{M}}$, the requantized Hamiltonian in Eq. (9.12) can be established as

$$\mathcal{H}(\boldsymbol{r},\boldsymbol{p}) = E_0(\boldsymbol{\pi}) - e\phi(\boldsymbol{r}) + \alpha\boldsymbol{E}\cdot\boldsymbol{\sigma}\times\boldsymbol{\pi} + \delta g\mu_B \boldsymbol{B}\cdot\frac{\hbar\boldsymbol{\sigma}}{2}, \tag{9.25}$$

where E_0 includes the Zeeman energy from the bare spin and

$$\delta g = -\frac{4}{3}\frac{mV^2}{\hbar^2}\left(\frac{1}{E_g} - \frac{1}{E_g + \Delta}\right). \tag{9.26}$$

In most textbooks on solid-state physics, one can find this correction of the g factor. However, a clear identification with electron's angular momentum is often lacking. In the wave-packet formulation, we see that δg is indeed originated from the electron's spinning motion.

F. Incompleteness of effective Hamiltonian

Once the effective Hamiltonian $\mathcal{H}(\boldsymbol{r},\boldsymbol{p})$ is obtained, one can go on to study its spectra and states, without referring back to the original Hamiltonian. Based on the spectra and states, any physics observables of interest can be calculated. These physics variables may be position, momentum, or other related quantities. Nevertheless, we emphasize that the canonical variables in the effective Hamiltonian may not be physical observables. They may differ, for example, by a Berry connection in the case of the position variable. The effective Hamiltonian itself is not enough for correct prediction if the physical variables have not been identified properly.

This is best illustrated using the Dirac electron as an example. At low velocity, the effective Pauli Hamiltonian is [see Eq. (9.19)]

$$\mathcal{H}(\boldsymbol{r},\boldsymbol{p}) = \frac{\pi^2}{2m} - e\phi(\boldsymbol{r}) + \frac{\mu_B}{2}\frac{\boldsymbol{\pi}}{mc^2}\times\boldsymbol{\sigma}\cdot\boldsymbol{E} + \mu_B\boldsymbol{\sigma}\cdot\boldsymbol{B}, \tag{9.27}$$

which is a starting point of many solid-state calculations. It is considered accurate for most of the low-energy applications in solid state. When one applies an electric field, then according to the Heisenberg equation of motion, the velocity of the electron is

$$\dot{\boldsymbol{r}} = \frac{\boldsymbol{\pi}}{m} + \frac{e\lambda_c^2}{4\hbar}\boldsymbol{\sigma}\times\boldsymbol{E}, \tag{9.28}$$
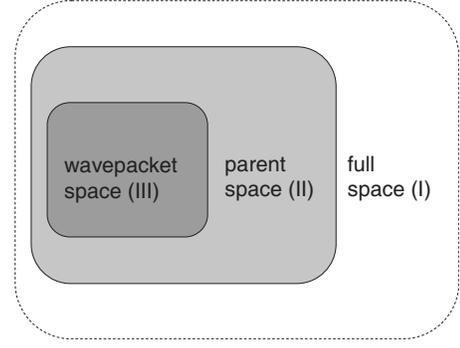
where λ_c is the Compton wavelength.

FIG. 16. The extent of wave-packet space, parent space, and full space.

If one calculates the velocity of a Dirac electron according to Eq. (9.7), then the result is

$$\dot{\boldsymbol{r}}_c = \frac{\hbar\boldsymbol{k}}{m} + \frac{e\lambda_c^2}{2\hbar}\langle\boldsymbol{\sigma}\rangle\times\boldsymbol{E}. \tag{9.29}$$

That is, the transverse velocity is larger by a factor of 2. The source of this discrepancy can be traced back to the difference between the two position variables: $\boldsymbol{r}_c$ and $\boldsymbol{r}$ [see Eq. (9.11)]. One should regard the equation for $\dot{\boldsymbol{r}}_c$ as the correct one since it is based on the Dirac theory [see also Bliokh (2005)].

Such a discrepancy between the same physical variable in different theories can also be understood from the perspective of the Foldy-Wouthuysen transformation. The Pauli Hamiltonian can also be obtained from block diagonalizing the Dirac Hamiltonian using an unitary transformation. Since the basis of states has been rotated, the explicit representations of all observables should be changed as well. For example, $\boldsymbol{r}_c$ in Eq. (9.11) can be obtained by a Foldy-Wouthuysen rotation, followed by a projection to the positive-energy subspace (Foldy and Wouthuysen, 1950).

G. Hierarchy structure of effective theories

Finally, we report on the hierarchical relations for the basic quantities, the Berry curvature $\boldsymbol{\mathcal{F}}$ and the magnetic moment $\boldsymbol{\mathcal{M}}$. We consider theories on three different levels of hierarchy (I, II, and III) with progressively smaller and smaller Hilbert spaces. These spaces will be called the full space, the parent space, and the wave-packet space, respectively (see Fig. 16).

Alternative to Eqs. (9.3) and (9.5), the Berry curvature and the magnetic moment can be written in the following forms (Chang and Niu, 2008):

$$\boldsymbol{F}_{mn} = i\sum_{l\in\text{out}}\boldsymbol{R}_{ml}\times\boldsymbol{R}_{ln}, \tag{9.30}$$

$$\boldsymbol{M}_{mn} = \frac{ie}{2\hbar}\sum_{l\in\text{out}}(E_{0,m} - E_{0,l})\boldsymbol{R}_{ml}\times\boldsymbol{R}_{ln}, \tag{9.31}$$

where $\boldsymbol{R}_{ml}$ is the Berry connection and l sums over the states *outside* of the space of interest. From Eqs. (9.30)

and (9.31), one sees that the Berry curvature and the magnetic moment for theory I are 0 since there is no state outside the full space. With the help of the states in the full space, one can calculate the Berry curvatures and the magnetic moment in theory II and theory III. They are designated as $(\mathcal{F}_p, \mathcal{M}_p)$ and $(\mathcal{F}, \mathcal{M})$, respectively. These two sets of matrices have different ranks since the parent space and the wave-packet space have different dimensions.

If one starts from the parent space, on the other hand, then the Berry curvature and the magnetic moment for theory II is 0 (instead of $\mathcal{F}_p$ and $\mathcal{M}_p$). The Berry curvature and the magnetic moment for theory III are now designated as $\mathcal{F}'$ and $\mathcal{M}'$. They are different from $\mathcal{F}$ and $\mathcal{M}$ since the former are obtained from the summations with more outside states from the full space. It is straightforward to see from Eqs. (9.30) and (9.31) that

$$\mathcal{F} = \mathcal{F}' + P\mathcal{F}_p P, \quad \mathcal{M} = \mathcal{M}' + P\mathcal{M}_p P, \qquad (9.32)$$

where P is a dimension-reduction projection from the parent space to the wave-packet subspace. This means that starting from theory II, instead of theory I, as the parent theory one would have the errors $P\mathcal{F}_p P$ and $P\mathcal{M}_p P$. On the other hand, however, whenever the scope of the parent theory needs to be extended, e.g., from II to I, instead of starting all of the calculations anew, one only needs additional input from $\mathcal{F}_p$ and $\mathcal{M}_p$ and the accuracy can be improved easily.

For example, in the original proposal of Murakami *et al.* (2003, 2004) of the spin Hall effect of holes, the parent space is the HH-LH complex. The heavy hole (or the light hole) acquires a nonzero Berry curvature as a result of the projection from this parent space to the HH band (or the LH band). This Berry curvature corresponds to the $\mathcal{F}'$ above. It gives rise to a spin-dependent transverse velocity $e\boldsymbol{E} \times \mathcal{F}'$ that is crucial to the spin Hall effect.

Instead of the HH-LH complex, if one chooses the eight bands in Fig. 15 as the full space, then the Berry curvatures of the heavy hole and the light hole will get new contributions from $P\mathcal{F}_p P$. The projection from the full space with eight bands to the HH-LH complex of four bands generates a Berry curvature $\mathcal{F}_p = -(2V^2/3E_g^2)\mathcal{J}$ (Chang and Niu, 2008), where $\mathcal{J}$ is the spin-3/2 matrix. Therefore, after further projections, we would get additional anomalous velocities $(eV^2/E_g^2)\boldsymbol{E} \times \boldsymbol{\sigma}$ and $(eV^2/3E_g^2)\boldsymbol{E} \times \boldsymbol{\sigma}$ for HH and LH, respectively.

X. OUTLOOK

In most of the researches mentioned in this review, the Berry phase and semiclassical theory are explored in the single-particle context. The fact that they are so useful and that in some of the materials the many-body effect is crucial naturally motivates one to extend this approach to many-body regime. For example, the Berry phase effect has been explored in the density functional theory with spin degrees of freedom (Niu and Kleinman, 1998; Niu *et al.*, 1999). Also, the Berry phase and rel-

evant quantities are investigated in the context of Fermi-liquid theory (Haldane, 2004). Furthermore, the Berry curvature on the Fermi surface, if strong enough, is predicted to modify a repulsive interaction between electrons to an attractive interaction and causes pairing instability (Shi and Niu, 2006). In addition to the artificial magnetic field generated by the monopole of Berry curvature, a slightly different Berry curvature involving the time component is predicted to generate an artificial electric field, which would affect the normalization factor and the transverse conductivity (Shindou and Balents, 2006). This latter work has henceforth been generalized to multiple-band Fermi liquid with non-Abelian Berry phase (Shindou and Balents, 2008). Recently, the ferrotoroidic moment in multiferroic materials is also found to be a quantum geometric phase (Batista *et al.*, 2008). Research along such a path is exciting and still at its early stage.

There has been a growing amount of research on the Berry phase effect in light-matter interaction. The Berry curvature is responsible for a transverse shift (side jump) of a light beam reflecting off an interface (Onoda *et al.*, 2004a; Sawada and Nagaosa, 2005; Onoda, Murakami, and Nagaosa, 2006). The shift is of the order of the wavelength and is a result of the conservation of angular momentum. The direction of the shift depends on the circular polarization of the incident beam. This "optical Hall effect" can be seen as a rediscovery of the Imbert-Federov effect (Federov, 1955; Imbert, 1972). More detailed study of the optical transport involving spin has also been carried out by Bliokh and others (Bliokh, 2006a; Bliokh and Bliokh, 2006; Duval *et al.*, 2006a). The similarity between the side jump of a light beam and analogous "jump" of an electron scattering off an impurity has been noticed quite early in Berger and Bergmann's review on anomalous Hall effect (Chien and Westgate, 1980). In fact, the side jump of the electromagnetic wave and the electron can be unified using similar dynamical equations. This shows that the equation of motion approach in this review has general validity. Indeed, a similar approach has also been extended to the quasiparticle dynamics in Bose-Einstein condensate (Zhang *et al.*, 2006).

Even though the Berry curvature plays a crucial role in the electronic structure and electron dynamics of crystals, direct measurement of such a quantity is still lacking. There does exist sporadic and indirect evidence of the effect of the Berry phase or the Berry curvature through the measurement of, for example, the quantum Hall conductance, the anomalous Hall effect, or the Hall plateau in graphene. However, this is just a beginning. In this review, one can see that in many circumstances the Berry curvature should be as important as the Bloch energy. Condensed-matter physicists over the years have compiled a large database on the band structures and Fermi surfaces of all kinds of materials. It is about time to add theoretical and experimental results of the Berry curvature that will deepen our understanding of material properties. There is still plenty of room in the quasimomentum space.

ACKNOWLEDGMENTS

APPENDIX: ADIABATIC EVOLUTION

Suppose the Hamiltonian $H[\boldsymbol{R}(t)]$ depends on a set of parameters $\boldsymbol{R}(t)$. Consider an adiabatic process in which $\boldsymbol{R}(t)$ changes slowly in time. The wave function $|\psi(t)\rangle$ must satisfy the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle = H(t)|\psi(t)\rangle. \tag{A1}$$

If we expand the wave function using the instantaneous eigenstates $|n(t)\rangle$ of $H(t)$ as

$$|\psi(t)\rangle = \sum_n \exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt' E_n(t')\right) a_n(t)|n(t)\rangle, \tag{A2}$$

then the coefficients satisfy

$$\dot{a}_n(t) = -\sum_l a_l(t)\langle n(t)|\frac{\partial}{\partial t}|l(t)\rangle$$
$$\times \exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt'[E_l(t') - E_n(t')]\right). \tag{A3}$$

So far we have not made any choice regarding the phase of the instantaneous eigenstates. For our purpose here, it is convenient to impose the condition of parallel transport,

$$\langle n(t)|\frac{\partial}{\partial t}|n(t)\rangle = \dot{\boldsymbol{R}}(t)\langle n(t)|\frac{\partial}{\partial \boldsymbol{R}}|n(t)\rangle = 0. \tag{A4}$$

We denote the wave function chosen this way as $|\tilde{n}\rangle$. We note that this condition is different from the single-valued phase choice we used in the main text: For a closed path in the parameter space, i.e., $\boldsymbol{R}(t_f) = \boldsymbol{R}(t_0)$, there is no guarantee that the phase at the final time t_f is the same as the phase at the beginning t_0. In other words, under the parallel transport condition, even though $|\tilde{n}(t)\rangle$ is uniquely determined as a function of t, it can still be a multivalued function of $\boldsymbol{R}$. The phase difference γ_n in $|\tilde{n}(t_f)\rangle = e^{i\gamma_n}|\tilde{n}(t_0)\rangle$ is precisely the Berry phase of the closed path. In fact, this can be considered as another definition of the Berry phase.

In the limit of $\dot{\boldsymbol{R}} \to 0$, we have, to zeroth order,

$$\dot{a}_n(t) = 0. \tag{A5}$$

Therefore if the system is initially in the nth eigenstate, it will stay in that state afterwards. This is the quantum adiabatic theorem. Now consider the first-order correction. We have $a_n(0) = 1$ and $a_{n'}(0) = 0$ for $n' \neq n$. For the

nth state, we still have $\dot{a}_n = 0$, therefore $a_n = 1$. However, for $n' \neq n$,

$$\frac{\partial}{\partial t}a_{n'} = -\langle \tilde{n}'|\frac{\partial}{\partial t}|\tilde{n}\rangle \exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt'[E_n(t') - E_{n'}(t')]\right). \tag{A6}$$

Then, since the exponential factor oscillates while its coefficient slowly varies in time, we can integrate the above equation by parts, yielding

$$a_{n'} = -\frac{\langle \tilde{n}'|\partial/\partial t|\tilde{n}\rangle}{E_n - E_{n'}}i\hbar$$
$$\times \exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt'[E_n(t') - E_{n'}(t')]\right). \tag{A7}$$

The wave function including the first-order approximation is given by

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt' E_n(t')\right)$$
$$\times \left\{|\tilde{n}\rangle - i\hbar \sum_{n \neq n'} |\tilde{n}'\rangle \frac{\langle \tilde{n}'|\partial/\partial t|\tilde{n}\rangle}{E_n - E_{n'}}\right\}. \tag{A8}$$

When applying the above result, as long as the expressions involving $|\tilde{n}\rangle$ are gauge independent, we can always replace $|\tilde{n}\rangle$ with eigenstates under another phase choice.

Finally, we mentioned that if the phase of the eigenstates is required to be single valued as a function of $\boldsymbol{R}$, Eq. (A8) can still be obtained with an extra *overall* phase $e^{i\gamma_n}$, where

$$\gamma_n = i\int_{t_0}^t dt'\langle n(\boldsymbol{R}(t))|\frac{\partial}{\partial t}|n(\boldsymbol{R}(t))\rangle. \tag{A9}$$

If the path $\boldsymbol{R}(t)$ is closed, γ_n gives the Berry phase.

REFERENCES

Adams, E. N., 1952, Phys. Rev. **85**, 41.
Adams, E. N., and E. I. Blount, 1959, J. Phys. Chem. Solids **10**, 286.
Aharonov, Y., and D. Bohm, 1959, Phys. Rev. **115**, 485.
Akhmerov, A. R., and C. W. J. Beenakker, 2007, Phys. Rev. Lett. **98**, 157003.
Arnold, V. I., 1978, *Mathematical Methods of Classical Mechanics* (Springer, New York).
Arnold, V. I., 1989, *Mathematical Methods of Classical Mechanics*, 2nd ed. (Springer-Verlag, New York).
Ashcroft, N. W., and N. D. Mermin, 1976, *Solid State Physics* (Saunders, Philadelphia).
Avron, J. E., 1982, Ann. Phys. (Paris) **143**, 33.
Avron, J. E., A. Elgart, G. M. Graf, and L. Sadun, 2001, Phys. Rev. Lett. **87**, 236601.
Avron, J. E., E. Elgart, G. M. Graf, and L. Sadun, 2004, J. Stat. Phys. **116**, 425.
Avron, J. E., and R. Seiler, 1985, Phys. Rev. Lett. **54**, 259.
Avron, J. E., R. Seiler, and B. Simon, 1983, Phys. Rev. Lett. **51**, 51.
Azbel, M. Y., 1964, Sov. Phys. JETP **19**, 634.

Bargmann, V., L. Michel, and V. L. Telegdi, 1959, Phys. Rev. Lett. **2**, 435.

Batista, C. D., G. Ortiz, and A. A. Aligia, 2008, Phys. Rev. Lett. **101**, 077203.

Ben Dahan, M., E. Peik, J. Reichel, Y. Castin, and C. Salomon, 1996, Phys. Rev. Lett. **76**, 4508.

Berger, L., 1970, Phys. Rev. B **2**, 4559.

Berger, L., 1972, Phys. Rev. B **5**, 1862.

Berry, M. V., 1984, Proc. R. Soc. London, Ser. A **392**, 45.

Bird, D. M., and A. R. Preston, 1988, Phys. Rev. Lett. **61**, 2863.

Bliokh, K. Y., 2005, Europhys. Lett. **72**, 7.

Bliokh, K. Y., 2006a, Phys. Rev. Lett. **97**, 043901.

Bliokh, K. Y., 2006b, Phys. Lett. A **351**, 123.

Bliokh, K. Y., and Y. P. Bliokh, 2006, Phys. Rev. Lett. **96**, 073903.

Bliokh, K. Y., A. Niv, V. Kleiner, and E. Hasman, 2008, Nat. Photonics **2**, 748.

Blount, E. I., 1962a, Phys. Rev. **126**, 1636.

Blount, E. I., 1962b, *Solid State Physics* (Academic, New York), Vol. 13.

Bohm, A., A. Mostafazadeh, H. Koizumi, Q. Niu, and J. Zwanziger, 2003, *The Geometric Phase in Quantum Systems: Foundations, Mathematical Concepts, and Applications in Molecular and Condensed Matter Physics* (Springer-Verlag, Berlin).

Brouwer, P. W., 1998, Phys. Rev. B **58**, R10135.

Brown, E., 1968, Phys. Rev. **166**, 626.

Bruno, P., V. K. Dugaev, and M. Taillefumier, 2004, Phys. Rev. Lett. **93**, 096806.

Bychkov, Y., and E. Rashba, 1984, JETP Lett. **39**, 78.

Ceresoli, D., T. Thonhauser, D. Vanderbilt, and R. Resta, 2006, Phys. Rev. B **74**, 024408.

Chambers, W., 1965, Phys. Rev. **140**, A135.

Chang, M.-C., and Q. Niu, 1995, Phys. Rev. Lett. **75**, 1348.

Chang, M.-C., and Q. Niu, 1996, Phys. Rev. B **53**, 7010.

Chang, M.-C., and Q. Niu, 2008, J. Phys.: Condens. Matter **20**, 193202.

Chazalviel, J.-N., 1975, Phys. Rev. B **11**, 3918.

Cheong, S.-W., and M. Mostovoy, 2007, Nature Mater. **6**, 13.

Chien, C. L., and C. R. Westgate, 1980, Eds., *The Hall Effect and Its Application* (Plenum, New York).

Chun, S. H., Y. S. Kim, H. K. Choi, I. T. Jeong, W. O. Lee, K. S. Suh, Y. S. Oh, K. H. Kim, Z. G. Khim, J. C. Woo, and Y. D. Park, 2007, Phys. Rev. Lett. **98**, 026601.

Chuu, C., M.-C. Chang, and Q. Niu, 2010, Solid State Commun. **150**, 533.

Cooper, N. R., B. I. Halperin, and I. M. Ruzin, 1997, Phys. Rev. B **55**, 2344.

Culcer, D., A. MacDonald, and Q. Niu, 2003, Phys. Rev. B **68**, 045327.

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

Culcer, D., Y. Yao, and Q. Niu, 2005, Phys. Rev. B **72**, 085110.

Dana, I., Y. Avron, and J. Zak, 1985, J. Phys. C **18**, L679.

Dayi, O. F., 2008, J. Phys. A: Math. Theor. **41**, 315204.

Dirac, P. A. M., 1931, Proc. R. Soc. London, Ser. A **133**, 60.

Dugaev, V. K., P. Bruno, M. Taillefumier, B. Canals, and C. Lacroix, 2005, Phys. Rev. B **71**, 224423.

Duval, C., Z. Horváth, P. A. Horváthy, L. Martina, and P. C. Stichel, 2006a, Mod. Phys. Lett. B **20**, 373.

Duval, C., Z. Horváth, P. A. Horváthy, L. Martina, and P. C. Stichel, 2006b, Phys. Rev. Lett. **96**, 099701.

Dyakonov, M. I., and V. I. Perel, 1971a, JETP Lett. **13**, 467.

Dyakonov, M. I., and V. I. Perel, 1971b, Phys. Lett. **35A**, 459.

Essin, A. M., J. E. Moore, and D. Vanderbilt, 2009, Phys. Rev. Lett. **102**, 146805.

Fang, Z., N. Nagaosa, K. S. Takahashi, A. Asamitsu, R. Mathieu, T. Ogasawara, H. Yamada, M. Kawasaki, Y. Tokura, and K. Terakura, 2003, Science **302**, 92.

Federov, F. I., 1955, Dokl. Akad. Nauk SSSR **105**, 465.

Fiebig, M., R. Lottermoser, D. Frohlich, A. V. Goltsev, and R. V. Pisarev, 2002, Nature (London) **419**, 818.

Fock, V., 1928, Z. Phys. **49**, 323.

Foldy, L. L., and S. A. Wouthuysen, 1950, Phys. Rev. **78**, 29.

Friedrich, H., and J. Trost, 1996, Phys. Rev. Lett. **76**, 4869.

Gat, O., and J. E. Avron, 2003a, New J. Phys. **5**, 44.

Gat, O., and J. E. Avron, 2003b, Phys. Rev. Lett. **91**, 186801.

Glück, M., A. R. Kolovsky, and H. J. Korsch, 1999, Phys. Rev. Lett. **83**, 891.

Goodrich, R. G., D. L. Maslov, A. F. Hebard, J. L. Sarrao, D. Hall, and Z. Fisk, 2002, Phys. Rev. Lett. **89**, 026401.

Gorbachev, R. V., F. V. Tikhonenko, A. S. Mayorov, D. W. Horsell, and A. K. Savchenko, 2007, Phys. Rev. Lett. **98**, 176805.

Gosselin, P., A. Bérard, and H. J. Mohrbach, 2007, Eur. Phys. J. B **58**, 137.

Gosselin, P., F. Ménas, A. Bérard, and H. Mohrbach, 2006, Europhys. Lett. **76**, 651.

Gunawan, O., Y. P. Shkolnikov, K. Vakili, T. Gokmen, E. P. D. Poortere, and M. Shayegan, 2006, Phys. Rev. Lett. **97**, 186404.

Guo, G.-Y., S. Maekawa, and N. Nagaosa, 2009, Phys. Rev. Lett. **102**, 036401.

Guo, G. Y., S. Murakami, T.-W. Chen, and N. Nagaosa, 2008, Phys. Rev. Lett. **100**, 096401.

Hagedorn, G. A., 1980, Commun. Math. Phys. **71**, 77.

Haldane, F. D. M., 1988, Phys. Rev. Lett. **61**, 2015.

Haldane, F. D. M., 2004, Phys. Rev. Lett. **93**, 206602.

Halperin, B. I., 1982, Phys. Rev. B **25**, 2185.

Hanasaki, N., K. Sano, Y. Onose, T. Ohtsuka, S. Iguchi, I. Kézsmárki, S. Miyasaka, S. Onoda, N. Nagaosa, and Y. Tokura, 2008, Phys. Rev. Lett. **100**, 106601.

Hatsugai, Y., and M. Kohmoto, 1990, Phys. Rev. B **42**, 8282.

Hirsch, J. E., 1999, Phys. Rev. Lett. **83**, 1834.

Hirst, L. L., 1997, Rev. Mod. Phys. **69**, 607.

Hofstadter, D. R., 1976, Phys. Rev. B **14**, 2239.

Hur, N., S. Park, P. A. Sharma, J. S. Ahn, S. Guha, and S.-W. Cheong, 2004, Nature (London) **429**, 392.

Imbert, C., 1972, Phys. Rev. D **5**, 787.

Inoue, J.-I., T. Kato, Y. Ishikawa, H. Itoh, G. E. W. Bauer, and L. W. Molenkamp, 2006, Phys. Rev. Lett. **97**, 046604.

Jackson, J. D., 1998, *Classical Electrodynamics*, 3rd ed. (Wiley, New York).

Jungwirth, T., Q. Niu, and A. H. MacDonald, 2002, Phys. Rev. Lett. **88**, 207208.

Kane, E. O., 1957, J. Phys. Chem. Solids **1**, 249.

Karplus, R., and J. M. Luttinger, 1954, Phys. Rev. **95**, 1154.

Kato, T., 1950, J. Phys. Soc. Jpn. **5**, 435.

Kato, T., Y. Ishikawa, H. Itoh, and J. Inoue, 2007, New J. Phys. **9**, 350.

Kato, Y. K., R. C. Myers, A. C. Gossard, and D. D. Awschalom, 2004, Science **306**, 1910.

Khinchin, A. Y., 1964, *Continued Fractions* (The University of Chicago Press, Chicago).

Kimura, T., T. Goto, H. Shintani, K. Ishizaka, T. Arima, and Y. Tokura, 2003, Nature (London) **426**, 55.

Kimura, T., Y. Otani, T. Sato, S. Takahashi, and S. Maekawa, 2007, Phys. Rev. Lett. **98**, 156601.

King-Smith, R. D., and D. Vanderbilt, 1993, Phys. Rev. B **47**, 1651.

Klitzing, K. v., G. Dorda, and M. Pepper, 1980, Phys. Rev. Lett. **45**, 494.

Kohmoto, M., 1985, Ann. Phys. (Paris) **160**, 343.

Kohmoto, M., 1989, Phys. Rev. B **39**, 11943.

Kohn, W., 1959, Phys. Rev. **115**, 809.

Kohn, W., 1964, Phys. Rev. **133**, A171.

Kohn, W., and J. M. Luttinger, 1957, Phys. Rev. **108**, 590.

Kotiuga, P. R., 1989, IEEE Trans. Magn. **25**, 2813.

Kramer, P., and M. Saraceno, 1981, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics* (Springer-Verlag, Berlin), Vol. 140.

Kunz, H., 1986, Phys. Rev. Lett. **57**, 1095.

Kuratsuji, H., and S. Iida, 1985, Prog. Theor. Phys. **74**, 439.

Laughlin, R. B., 1981, Phys. Rev. B **23**, 5632.

Lee, W.-L., S. Watauchi, V. L. Miller, R. J. Cava, and N. P. Ong, 2004a, Phys. Rev. Lett. **93**, 226601.

Lee, W.-L., S. Watauchi, V. L. Miller, R. J. Cava, and N. P. Ong, 2004b, Science **303**, 1647.

Lifshitz, E. M., and L. D. Landau, 1980, *Statistical Physics: Part 2 (Course of Theoretical Physics)* (Butterworth-Heinemann, Washington, D.C.), Vol. 9.

Littlejohn, R. G., and W. G. Flynn, 1991, Phys. Rev. A **44**, 5239.

Liu, C.-X., X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, 2008, Phys. Rev. Lett. **101**, 146802.

Luttinger, J. M., 1951, Phys. Rev. **84**, 814.

Luttinger, J. M., 1958, Phys. Rev. **112**, 739.

Marder, M. P., 2000, *Condensed-Matter Physics* (Wiley, New York).

Martin, R. M., 1972, Phys. Rev. B **5**, 1607.

Martin, R. M., 1974, Phys. Rev. B **9**, 1998.

Marzari, N., and D. Vanderbilt, 1997, Phys. Rev. B **56**, 12847.

Mathieu, R., A. Asamitsu, H. Yamada, K. S. Takahashi, M. Kawasaki, Z. Fang, N. Nagaosa, and Y. Tokura, 2004, Phys. Rev. Lett. **93**, 016602.

Mathur, H., 1991, Phys. Rev. Lett. **67**, 3325.

McCann, E., and V. I. Fal'ko, 2006, Phys. Rev. Lett. **96**, 086805.

Mendez, E., and G. Bastard, 1993, Phys. Today **46** (6), 34.

Messiah, A., 1962, *Quantum Mechanics* (North Holland, Amsterdam), Vol. II.

Mikitik, G. P., and Y. V. Sharlai, 1999, Phys. Rev. Lett. **82**, 2147.

Mikitik, G. P., and Y. V. Sharlai, 2004, Phys. Rev. Lett. **93**, 106403.

Mikitik, G. P., and Y. V. Sharlai, 2007, Low Temp. Phys. **33**, 439.

Min, H., B. Sahu, S. K. Banerjee, and A. H. MacDonald, 2007, Phys. Rev. B **75**, 155115.

Miyasato, T., N. Abe, T. Fujii, A. Asamitsu, S. Onoda, Y. Onose, N. Nagaosa, and Y. Tokura, 2007, Phys. Rev. Lett. **99**, 086602.

Morandi, G., 1988, *Quantum Hall Effect: Topological Problems in Condensed-Matter Physics* (Bibliopolis, Naples).

Morozov, S. V., K. S. Novoselov, M. I. Katsnelson, F. Schedin, L. A. Ponomarenko, D. Jiang, and A. K. Geim, 2006, Phys. Rev. Lett. **97**, 016801.

Morpurgo A. F., and F. Guinea, 2006, Phys. Rev. Lett. **97**, 196804.

Mucciolo E. R., C. Chamon, and C. M. Marcus, 2002, Phys. Rev. Lett. **89**, 146802.

Murakami S., N. Nagosa, and S.-C. Zhang, 2003, Science **301**, 1348.

Murakami S., N. Nagosa, and S.-C. Zhang, 2004, Phys. Rev. B **69**, 235206.

Nagaosa, N., J. Sinova, S. Onoda, A. H. MacDonald, and N. P. Ong, 2010, Rev. Mod. Phys. **82**, 1539.

Nayak, C., S. H. Simon, A. Stern, M. Freedman, and S. D. Sarma, 2008, Rev. Mod. Phys. **80**, 1083.

Nenciu, G., 1991, Rev. Mod. Phys. **63**, 91.

Niu, Q., 1990, Phys. Rev. Lett. **64**, 1812.

Niu, Q., and L. Kleinman, 1998, Phys. Rev. Lett. **80**, 2205.

Niu, Q., and G. Sundaram, 2001, Physica E (Amsterdam) **9**, 327.

Niu, Q., and D. J. Thouless, 1984, J. Phys. A **17**, 2453.

Niu, Q., D. J. Thouless, and Y.-S. Wu, 1985, Phys. Rev. B **31**, 3372.

Niu, Q., X. Wang, L. Kleinman, W.-M. Liu, D. M. C. Nicholson, and G. M. Stocks, 1999, Phys. Rev. Lett. **83**, 207.

Novoselov, K. S., A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, 2005, Nature (London) **438**, 197.

Novoselov, K. S., E. McCann, S. V. Morozov, V. I. Fal'ko, M. I. Katsnelson, U. Zeitler, D. Jiang, F. Schedin, and A. K. Geim, 2006, Nat. Phys. **2**, 177.

Nozières, P., and C. Lewiner, 1973, J. Phys. (France) **34**, 901.

Nunes, R. W., and X. Gonze, 2001, Phys. Rev. B **63**, 155107.

Nunes, R. W., and D. Vanderbilt, 1994, Phys. Rev. Lett. **73**, 712.

Obermair, G., and G. Wannier, 1976, Phys. Status Solidi B **76**, 217.

Ohta, T., A. Bostwick, T. Seyller, K. Horn, and E. Rotenberg, 2006, Science **313**, 951.

Olson, J. C., and P. Ao, 2007, Phys. Rev. B **75**, 035114.

Onoda, M., S. Murakami, and N. Nagaosa, 2004a, Phys. Rev. Lett. **93**, 083901.

Onoda, M., S. Murakami, and N. Nagaosa, 2006, Phys. Rev. E **74**, 066610.

Onoda, M., and N. Nagaosa, 2002, J. Phys. Soc. Jpn. **71**, 19.

Onoda, S., S. Murakami, and N. Nagaosa, 2004b, Phys. Rev. Lett. **93**, 167602.

Onoda, S., N. Sugimoto, and N. Nagaosa, 2006, Phys. Rev. Lett. **97**, 126602.

Onoda, S., N. Sugimoto, and N. Nagaosa, 2008, Phys. Rev. B **77**, 165103.

Onsager, L., 1952, Philos. Mag. **43**, 1006.

Ortiz, G., and R. M. Martin, 1994, Phys. Rev. B **49**, 14202.

Panati, G., H. Spohn, and S. Teufel, 2003, Commun. Math. Phys. **242**, 547.

Peierls, R. E., 1933, Z. Phys. **80**, 763.

Prange, R., and S. Girvin, 1987, *The Quantum Hall Effect* (Springer, New York).

Pu, Y., D. Chiba, F. Matsukura, H. Ohno, and J. Shi, 2008, Phys. Rev. Lett. **101**, 117208.

Qi, X.-L., T. Hughes, and S.-C. Zhang, 2008, Phys. Rev. B **78**, 195424.

Qi, X.-L., Y.-S. Wu, and S.-C. Zhang, 2006, Phys. Rev. B **74**, 085308.

Rammal, R., and J. Bellissard, 1990, J. Phys. (France) **51**, 1803.

Rashba, E. I., 1960, Sov. Phys. Solid State **2**, 1224.

Resta, R., 1992, Ferroelectrics **136**, 51.

Resta, R., 1994, Rev. Mod. Phys. **66**, 899.

Resta, R., 1998, Phys. Rev. Lett. **80**, 1800.

Resta, R., 2000, J. Phys.: Condens. Matter **12**, R107.

Resta, R., and S. Sorella, 1999, Phys. Rev. Lett. **82**, 370.

Resta, R., and D. Vanderbilt, 2007, in *Physics of Ferroelectrics: A Modern Perspective*, edited by K. Rabe, C. H. Ahn, and

J.-M. Triscone (Springer-Verlag, Berlin), p. 31.

Rice, M. J., and E. J. Mele, 1982, Phys. Rev. Lett. **49**, 1455.

Rycerz, A., J. Tworzydio, and C. W. J. Beenakker, 2007, Nat. Phys. **3**, 172.

Sakurai, J. J., 1993, *Modern Quantum Mechanics*, 2nd ed. (Addison Wesley, Reading, MA).

Sales, B. C., R. Jin, D. Mandrus, and P. Khalifah, 2006, Phys. Rev. B **73**, 224435.

Sawada, K., and N. Nagaosa, 2005, Phys. Rev. Lett. **95**, 237402.

Schellnhuber, H. J., and G. M. Obermair, 1980, Phys. Rev. Lett. **45**, 276.

Seki, T., Y. Hasegawa, S. Mitani, S. Takahashi, H. Imamura, S. Maekawa, J. Nitta, and K. Takanashi, 2008, Nature Mater. **7**, 125.

Semenoff, G. W., 1984, Phys. Rev. Lett. **53**, 2449.

Shankar, R., and H. Mathur, 1994, Phys. Rev. Lett. **73**, 1565.

Shapere, A., and F. Wilczek, 1989, Eds., *Geometric Phases in Physics* (World Scientific, Singapore).

Sharma, P., and C. Chamon, 2001, Phys. Rev. Lett. **87**, 096401.

Shi, J., and Q. Niu, 2006, e-print arXiv:cond-mat/0601531.

Shi, J., G. Vignale, D. Xiao, and Q. Niu, 2007, Phys. Rev. Lett. **99**, 197202.

Shindou, R., 2005, J. Phys. Soc. Jpn. **74**, 1214.

Shindou, R., and L. Balents, 2006, Phys. Rev. Lett. **97**, 216601.

Shindou, R., and L. Balents, 2008, Phys. Rev. B **77**, 035110.

Shindou, R., and K.-I. Imura, 2005, Nucl. Phys. B **720**, 399.

Sih, V., R. C. Myers, Y. K. Kato, W. H. Lau, A. C. Gossard, and D. D. Awschalom, 2005, Nat. Phys. **1**, 31.

Simon, B., 1983, Phys. Rev. Lett. **51**, 2167.

Sinitsyn, N. A., 2008, J. Phys.: Condens. Matter **20**, 023201.

Sinitsyn, N. A., A. H. MacDonald, T. Jungwirth, V. K. Dugaev, and J. Sinova, 2007, Phys. Rev. B **75**, 045315.

Sinitsyn, N. A., Q. Niu, J. Sinova, and K. Nomura, 2005, Phys. Rev. B **72**, 045346.

Sipe, J. E., and J. Zak, 1999, Phys. Lett. A **258**, 406.

Slater, J. C., 1949, Phys. Rev. **76**, 1592.

Smit, J., 1958, Physica (Amsterdam) **24**, 39.

Souza, I., J. Íñiguez, and D. Vanderbilt, 2002, Phys. Rev. Lett. **89**, 117602.

Souza, I., and D. Vanderbilt, 2008, Phys. Rev. B **77**, 054438.

Souza, I., T. Wilkens, and R. M. Martin, 2000, Phys. Rev. B **62**, 1666.

Stone, A. D., 2005, Phys. Today **58** (8), 37.

Strange, P., 1998, *Relativistic Quantum Mechanics: With Applications in Condensed Matter and Atomic Physics* (Cambridge University Press, Cambridge).

Streda, P., 1982, J. Phys. C **15**, L717.

Strinati, G., 1978, Phys. Rev. B **18**, 4104.

Su, W. P., J. R. Schrieffer, and A. J. Heeger, 1979, Phys. Rev. Lett. **42**, 1698.

Sundaram, G., and Q. Niu, 1999, Phys. Rev. B **59**, 14915.

Switkes, M., C. M. Marcus, K. Campman, and A. C. Gossard, 1999, Science **283**, 1905.

Tabor, M., 1989, *Chaos and Integrability in Nonlinear Dynamics: An Introduction* (Wiley-Interscience, New York).

Tagantsev, A. K., 1986, Phys. Rev. B **34**, 5883.

Tagantsev, A. K., 1991, Phase Transitions **35**, 119.

Talyanskii, V. I., J. M. Shilton, M. Pepper, C. G. Smith, C. J. B. Ford, E. H. Linfield, D. A. Ritchie, and G. A. C. Jones, 1997, Phys. Rev. B **56**, 15180.

Teufel, S., 2003, *Adiabatic Perturbation Theory in Quantum Dynamics* (Springer-Verlag, New York).

Thonhauser, T., D. Ceresoli, D. Vanderbilt, and R. Resta, 2005,

Phys. Rev. Lett. **95**, 137205.

Thouless, D. J., 1981, J. Phys. C **14**, 3475.

Thouless, D. J., 1983, Phys. Rev. B **27**, 6083.

Thouless, D. J., 1998, *Topological Quantum Numbers in Non-relativistic Physics* (World Scientific, Singapore).

Thouless, D. J., M. Kohmoto, M. P. Nightingale, and M. den Nijs, 1982, Phys. Rev. Lett. **49**, 405.

Tian, Y. L. Ye, and X. Jin, 2009, Phys. Rev. Lett. **103**, 087206.

Umucalılar, R. O., H. Zhai, and M. O. Oktel, 2008, Phys. Rev. Lett. **100**, 070402.

Valenzuela, S. O., and M. Tinkham, 2006, Nature (London) **442**, 176.

Vanderbilt, D., and R. D. King-Smith, 1993, Phys. Rev. B **48**, 4442.

van Houten, H., C. W. J. Beenakker, J. G. Williamson, M. E. I. Broekaart, P. H. M. van Loosdrecht, B. J. van Wees, J. E. Mooij, C. T. Foxon, and J. J. Harris, 1989, Phys. Rev. B **39**, 8556.

Vignale, G., and M. Rasolt, 1988, Phys. Rev. B **37**, 10685.

Wang, X., D. Vanderbilt, J. R. Yates, and I. Souza, 2007, Phys. Rev. B **76**, 195109.

Wannier, G. H., 1962, Rev. Mod. Phys. **34**, 645.

Wannier, G. H., 1980, J. Math. Phys. **21**, 2844.

Watson, S. K., R. M. Potok, C. M. Marcus, and V. Umansky, 2003, Phys. Rev. Lett. **91**, 258301.

Wen, X. G., and Q. Niu, 1990, Phys. Rev. B **41**, 9377.

Wilczek, F., 1987, Phys. Rev. Lett. **58**, 1799.

Wilczek, F., and A. Zee, 1984, Phys. Rev. Lett. **52**, 2111.

Wilkinson, M., 1984a, Proc. R. Soc. London, Ser. A **391**, 305.

Wilkinson, M., 1984b, J. Phys. A **17**, 3459.

Wilkinson, M., and R. J. Kay, 1996, Phys. Rev. Lett. **76**, 1896.

Winkler, R., 2003, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer, New York).

Wu, T. T., and C. N. Yang, 1975, Phys. Rev. D **12**, 3845.

Wunderlich, J., B. Kaestner, J. Sinova, and T. Jungwirth, 2005, Phys. Rev. Lett. **94**, 047204.

Xiao, D., D. Clougherty, and Q. Niu, 2007, unpublished.

Xiao, D., J. Shi, D. P. Clougherty, and Q. Niu, 2009, Phys. Rev. Lett. **102**, 087602.

Xiao, D., J. Shi, and Q. Niu, 2005, Phys. Rev. Lett. **95**, 137204.

Xiao, D., J. Shi, and Q. Niu, 2006, Phys. Rev. Lett. **96**, 099702.

Xiao, D., W. Yao, and Q. Niu, 2007, Phys. Rev. Lett. **99**, 236809.

Xiao, D., Y. Yao, Z. Fang, and Q. Niu, 2006, Phys. Rev. Lett. **97**, 026603.

Yafet, Y., 1963, in *Solid State Physics: Advances in Research and Applications*, edited by F. Seitz and D. Turnbull (Academic, New York), Vol. 14, p. 1.

Yang, S. A., G. S. D. Beach, C. Knutson, D. Xiao, Q. Niu, M. Tsoi, and J. L. Erskine, 2009, Phys. Rev. Lett. **102**, 067201.

Yao, W., D. Xiao, and Q. Niu, 2008, Phys. Rev. B **77**, 235406.

Yao, Y., L. Kleinman, A. H. MacDonald, J. Sinova, T. Jungwirth, D.-S. Wang, E. Wang, and Q. Niu, 2004, Phys. Rev. Lett. **92**, 037204.

Yao, Y., Y. Liang, D. Xiao, Q. Niu, S.-Q. Shen, X. Dai, and Z. Fang, 2007, Phys. Rev. B **75**, 020401.

Ye, J., Y. B. Kim, A. J. Millis, B. I. Shraiman, P. Majumdar, and Z. Tesanovic, 1999, Phys. Rev. Lett. **83**, 3737.

Zak, J., 1964a, Phys. Rev. **136**, A776.

Zak, J., 1964b, Phys. Rev. **134**, A1602.

Zak, J., 1964c, Phys. Rev. **134**, A1607.

Zak, J., 1977, Phys. Rev. B **15**, 771.

Zak, J., 1989, Phys. Rev. Lett. **62**, 2747.

Zeng, C., Y. Yao, Q. Niu, and H. H. Weitering, 2006, Phys. Rev. Lett. **96**, 037204.

Zhang, C., A. M. Dudarev, and Q. Niu, 2006, Phys. Rev. Lett. **97**, 040401.

Zhang, Y., Y.-W. Tan, H. L. Stormer, and P. Kim, 2005, Nature (London) **438**, 201.

Zheng, W., J. Wu, B. Wang, J. Wang, Q. Sun, and H. Guo, 2003, Phys. Rev. B **68**, 113306.

Zhou, F., B. Spivak, and B. Altshuler, 1999, Phys. Rev. Lett. **82**, 608.

Zhou, H.-Q., S. Y. Cho, and R. H. McKenzie, 2003, Phys. Rev. Lett. **91**, 186803.

Zhou, S. Y., G.-H. Gweon, A. V. Fedorov, P. N. First, W. A. de Heer, D.-H. Lee, F. Guinea, A. H. C. Neto, and A. Lanzara, 2007, Nature Mater. **6**, 770.