A Short Course on Topological Insulators

J. K. Asbóth, L. Oroszlány, A. Pályi

September 18, 2019

These lecture notes provide an introduction to some of the main concepts of topological insulators, a branch of solid state physics that is developing at a fast pace. They are based on a one-semester course for MSc and PhD students at the Eötvös University, Budapest, which the authors have been giving since 2012.

Our aim is to provide an understanding of the core topics of topological insulators – edge states, bulk topological invariants, bulk–boundary correspondence – with as simple mathematical tools as possible. We restricted our attention to one- and two-dimensional band insulators. We use noninteracting lattice models of topological insulators, and build these up gradually to arrive from the simplest one-dimensional case (the Su-Schrieffer-Heeger model for polyacetylene) to two-dimensional time-reversal invariant topological insulators (the Bernevig-Hughes-Zhang model for HgTe). In each case we introduce the model first, discuss its properties, and then generalize. The prerequisite for the reader is quantum mechanics and not much else: solid state physics background is provided as we go along.

Since this is an introduction, rather than a broad overview, we try to be self-contained and give citations to the current literature only where it is absolutely necessary. For a broad overview, including pointers to the original papers and current topics, we refer the reader to review articles and books in the Introduction.

Supporting material for these lecture notes in the form of ipython notebooks will be made available online.

Despite our efforts, the book inevitably contains typos, errors, and less comprehensible explanations. We would appreciate if you could inform us of any of those; please send your comments to `janos.asboth@wigner.mta.hu`.

Acknowledgments. We are grateful for enlightening discussions on topological insulators with Anton Akhmerov, Andrea Alberti, Carlo Beenakker, and Alberto Cortijo. We thank the feedback we got from participants at the courses at Eötvös University, especially Vilmos Kocsis. A version of this course was given by one of us (J.K.A.) in the PhD programme of the University of Geneva, on invitation by Markus Büttiker, which helped shape the course.

The preparation of the first version of these lecture notes were supported by the grant TÁMOP4.1.2.A/1-11/0064. We acknowledge financial support from the Hungarian Scientific Research Fund (OTKA), Grant Nos. PD100373, K108676, NN109651, and from the Marie Curie programme of the European Union, Grant No. CIG-293834. J.K.A. and A.P. were supported by the János Bolyai Scholarship of the Hungarian Academy of Sciences

Budapest,
September 2015

János K. Asbóth, László Oroszlány, András Pályi

Introduction

The band theory of electric conduction was one of the early victories of quantum mechanics in the 1920s. It gave a simple explanation of how some crystalline materials are electric insulators, even though electrons in them can hop from one atom to the next. In the bulk of a band insulator, the electrons occupy eigenstates that form energy bands. In a band insulator, there are no partially filled bands: completely filled bands are separated by an energy gap from completely empty bands, the gap represents the energy cost of mobilizing electrons. In contrast, materials with partially filled bands are conductors, where there are plane wave states available to transmit electrons across the bulk at arbitrarily low energy. Although we now know of situations where band theory is inadequate (e.g., for Mott insulators), it remains one of the cornerstones of solid states physics.

The discovery of the Quantum Hall Effect (1980) has shown that the simple division into band insulators and metals is not the end of the story, not even in band theory. In the quantum Hall effect, a strong magnetic field confines the motion of electrons in the bulk, but the same field forces them into delocalized edge states on the surface. A two-dimensional metal in strong magnetic field is thus an insulator in the bulk, but conducts along the surface, via a discrete number of completely open edge state channels (in the language of the Landauer–Büttiker formalism). The number of edge state channels was linked to the Chern number, a topological invariant of the occupied bands (1982).

Over the last twenty years, theoretical progress over artificial systems has shown that the external magnetic field is not necessary for an insulator to have robust conducting edge states: instead, the nontrivial topology of the occupied bands is the crucial ingredient. The name *topological insulator* was coined for such systems, and their study became a blossoming branch of solid state physics. Following the theoretical prediction (Bernevig, Hughes and Zhang, 2006 [3]), electronic transport measurements confirmed that a thin layer of HgTe is a topological insulator (König et al, 2007 [12]). Since that time, a host of materials have been shown to be three-dimensional topological insulators, and thin films and quantum wires shown to be two- and one-dimensional topological insulators [1].

The intense theoretical interest in topological insulators has led to signature results, such as the so-called periodic table of topological insulators [15], which shows that similarly to phase transitions in statistical mechanics, it is the dimensionality and the basic symmetries of an insulator that decide whether it can be a topological insulator or not. Although it was derived by different ways of connecting topological insulators of various dimensions and symmetries (so-called dimensional reduction schemes), the

mathematically rigorous proof of the periodic table is still missing.

The field of topological insulators is very active, with many experimental challenges and open theoretical problems, regarding the effects of electron-electron interaction, extra crystalline symmetries, coupling to the environment, etc.

Literature

To get a quick and broad overview of topological insulators, with citations for relevant research papers, we recommend the review papers [10, 13, 5]. For a more in-depth look, there are already a few textbooks on the subject (by Bernevig and Hughes [2], by Shen [16], and one edited by Franz and Molenkamp [6]). To see the link between momentum-space topology and physics in a broader context, we direct the reader to a book by Volovik[17].

There are also introductory courses on topological insulators with a broad scope. We recommend the lectures by Charles Kane (the video recording of the version given at Veldhoven is freely available online), and the online EdX course on topology in condensed matter by a group of lecturers, with the corresponding material collected at topocondmat.org.

These lecture notes

Our aim with this set of lecture notes is to complement the literature cited above: we wish to provide a close look at some of the core concepts of topological insulators with as simple mathematical tools as possible. Using one-and two-dimensional noninteracting lattice models, we explain what edge states and what bulk topological invariants are, how the two are linked (this is known as the *bulk–boundary correspondence*), the meaning and impact of some of the fundamental symmetries.

To keep things as simple as possible, throughout the course we use noninteracting models for solid state systems. These are described using single-particle lattice Hamiltonians, with the zero of the energy corresponding to the Fermi energy. We use natural units, with $\hbar = 1$ and length measured by the lattice constant.

Contents

Chapter 1

Berry Phase, Chern Number

* To describe the theory of topological band insulators we will use the language of adiabatic phases. In this chapter we review the basic concepts: the Berry phase, the Berry curvature, and the Chern number. We further describe the relation between the Berry phase and adiabatic dynamics in quantum mechanics. Finally, we illustrate these concepts using the two-level system as a simple example.

To describe the theory of topological band insulators we will use the language of adiabatic phases. In this chapter we review the basic concepts: the Berry phase, the Berry curvature, and the Chern number. We further describe the relation between the Berry phase and adiabatic dynamics in quantum mechanics. Finally, we illustrate these concepts using a two-level system as a simple example.

For pedagogical introductions, we refer the reader to Berry's original paper [4], and papers from the American Journal of Physics [11, 8]. For the application to solid state physics, we will mostly build on Resta's lecture note [14], and the review paper [18].

1.1 Discrete Case

We begin with simple questions regarding the relative phase of states in a Hilbert space.

1.1.1 The Relative Phase of Two Quantum States is Gauge Dependent

In quantum mechanics, the state of a physical system is represented by vectors in a Hilbert space up to a multiplication by a complex phase factor (an equivalence class, a *ray* of vectors). We refer to this multiplication as a *gauge transformation*, and it changes the vector $|\Psi\rangle$,

$$\text{gauge transformation:} \quad |\Psi\rangle \to e^{i\alpha}|\Psi\rangle, \quad \text{with } \alpha \in [0, 2\pi). \tag{1.1}$$

In other words, the phase of a state is gauge dependent, and hence, not observable.

The gauge dependence also makes the *relative phase of two states* undefined. Consider two possible states of the system, with vectors $|\Psi_1\rangle$ and $|\Psi_2\rangle$, that are nonorthogonal, $\langle\Psi_1 \,|\, \Psi\rangle_2 \neq 0$. We can try to define their relative phase γ_{12}, as

$$\gamma_{12} = -\arg\langle\Psi_1 \,|\, \Psi_2\rangle\,, \tag{1.2}$$

where $\arg(z)$ denotes the phase of the complex number z, with the specification that $\arg(z) \in (-\pi, \pi]$. Clearly, the relative phase γ_{12} fulfils

$$e^{-i\gamma_{12}} = \frac{\langle\Psi_1 \,|\, \Psi_2\rangle}{|\langle\Psi_1 \,|\, \Psi_2\rangle|}. \tag{1.3}$$

However, the relative phase is not invariant under a local gauge transformation,

$$\left|\Psi_j\right\rangle \to e^{i\alpha_j}\left|\Psi_j\right\rangle \qquad\qquad e^{-i\gamma_{12}} \to e^{-i\gamma_{12}+i(\alpha_2-\alpha_1)}. \tag{1.4}$$

1.1.2 Berry Phase: the Relative Phase Along a Closed Loop is Gauge Independent

It might come as a surprise that the *relative phase of three or more states* has a gauge invariant meaning – we will call this the discrete Berry phase. Take $N \geq 3$ states in a Hilbert space, $\left|\Psi_j\right\rangle$, with $j = 1, 2, \ldots, N$, with $\left\langle\Psi_j \,\middle|\, \Psi_{j+1}\right\rangle \neq 0$ for all j (addition is modulo N). Looping around this list of states, as shown in Fig. 1.1, we define the *Berry phase of the loop* as

$$L = (|\Psi\rangle_1, |\Psi\rangle_2, \ldots, |\Psi\rangle_N)\,; \tag{1.5}$$

$$\gamma_L = -\arg e^{-i(\gamma_{12}+\gamma_{23}+\ldots+\gamma_{N1})} = -\arg\left(\langle\Psi_1 \,|\, \Psi_2\rangle\,\langle\Psi_2 \,|\, \Psi_3\rangle \ldots \langle\Psi_N \,|\, \Psi_1\rangle\right). \tag{1.6}$$

To make it explicit that the Berry phase defined above is gauge invariant, we rewrite it using gauge invariant projectors. For a quantum state represented by $|\Psi\rangle$, a gauge invariant object representing the state is the projector $\hat{P}_\Psi$ to the ray,

$$\hat{P}_\Psi = |\Psi\rangle\langle\Psi|\,; \qquad\qquad \text{gauge transformation:} \quad \hat{P}_\Psi \to \hat{P}_\Psi. \tag{1.7}$$

The Berry phase defined above rewritten using only the gauge invariant projectors to the states in the loop,

$$\gamma_L = -\arg\mathrm{Tr}\left(|\Psi_1\rangle\langle\Psi_1|\,|\Psi_2\rangle\langle\Psi_2|\ldots|\Psi_N\rangle\langle\Psi_N|\right). \tag{1.8}$$

Even though the Berry phase is not the expectation value of some operator, it is a gauge invariant quantity, and as such, it can have a direct physical significance. We will find such a significance, but first, we want to gain more intuition about its behaviour.

1.1.3 Berry Flux: Gauge Independent Contribution of Plaquettes to the Berry Phase

Consider a Hilbert space of quantum states, and a finite two-dimensional square lattice with points labelled by $n, m \in \mathbb{Z}$, $1 \leq n \leq N$, and $1 \leq m \leq M$. Assign a quantum state

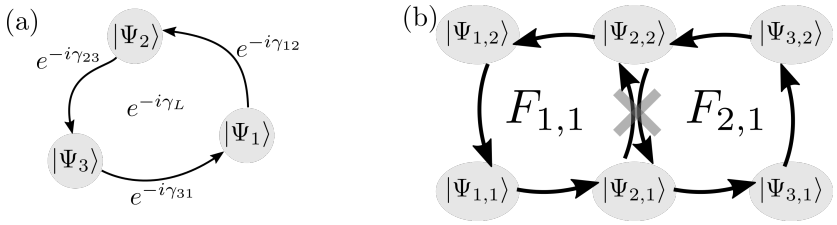

Figure 1.1: Berry phase, Berry flux and Berry curvature for discrete quantum states. (a) The Berry phase γ_L for the loop L consisting of $N = 3$ states is defined from the relative phases γ_{12}, γ_{23}, γ_{31}. (b) The Berry phase of a loop defined on a lattice of states can be expressed as the sum of the Berry phases $F_{1,1}$ and $F_{2,1}$ of the plaquettes enclosed by the loop. The plaquette Berry phase $F_{n,m}$ is also called Berry flux.

$|\Psi_{n,m}\rangle$ from the Hilbert space to each lattice site. Say you want to know the Berry phase of the loop L around this set,

$$\gamma_L = -\arg\exp\left[-i\left(\sum_{n=1}^{N-1}\gamma_{(n,1),(n+1,1)} + \sum_{m=1}^{M-1}\gamma_{(N,m),(N,m+1)} + \sum_{n=1}^{N-1}\gamma_{(n+1,M),(n,M)} + \sum_{m=1}^{M-1}\gamma_{(1,m+1),(1,m)}\right)\right] \quad (1.9)$$

as shown in Fig. 1.1. Although the Berry phase is a gauge invariant quantity, calculating it according to the recipe above involves multiplying together many gauge dependent complex numbers. The alternative route, via Eq. (1.8), involves multiplying gauge independent matrices, and then taking the trace.

There is a way to break the calculation of the Berry phase of the loop down to a product of gauge independent complex numbers. To each plaquette (elementary square) on the grid, with n,m indexing the lower left corner, we define the *Berry flux* $F_{n,m}$ of the plaquette using the sum of the relative phases around its boundary,

$$F_{nm} = -\arg\exp\left[-i\left(\gamma_{(n,m),(n+1,m)} + \gamma_{(n+1,m),(n+1,m+1)} + \gamma_{(n+1,m+1),(n,m+1)} + \gamma_{(n,m+1),(n,m)}\right)\right], \quad (1.10)$$

for $n = 1,\ldots,N$ and $m = 1,\ldots,M$. Note that the Berry flux is itself a Berry phase and is therefore gauge invariant. Alternatively, we can also write

$$F_{nm} = -\arg\left(\langle\Psi_{n,m}\,|\,\Psi_{n+1,m}\rangle\,\langle\Psi_{n+1,m}\,|\,\Psi_{n+1,m+1}\rangle\,\langle\Psi_{n+1,m+1}\,|\,\Psi_{n,m+1}\rangle\,\langle\Psi_{n,m+1}\,|\,\Psi_{n,m}\rangle\right), \quad (1.11)$$

Now consider the product of all plaquette phase factors $e^{-iF_{nm}}$,

$$\prod_{n=1}^{N-1}\prod_{m=1}^{M-1} e^{-iF_{nm}} = \exp\left[-i\sum_{n=1}^{N-1}\sum_{m=1}^{M-1}F_{nm}\right] = \exp\left[-i\sum_{n=1}^{N-1}\sum_{m=1}^{M-1}\left(\gamma_{(n,m),(n+1,m)} + \gamma_{(n+1,m),(n+1,m+1)} + \gamma_{(n+1,m+1),(n,m+1)} + \gamma_{(n,m+1),(n,m)}\right)\right] \quad (1.12)$$

Each internal edge of the lattice is shared between two plaquettes, and therefore occurs twice in the product. However, since we fixed the orientation of the plaquette phases, these two contributions will always be complex conjugates of each other, and cancel each other. Therefore the exponent in the right-hand-side of Eq. (1.12) simplifies to the exponent appearing in Eq. (1.9), implying

$$\exp\Big[-i\sum_{n=1}^{N-1}\sum_{m=1}^{M-1}F_{nm}\Big]=e^{-i\gamma_L}. \tag{1.13}$$

This result is reminiscent of the Stokes theorem connecting the integral of the curl of a vector field on an open surface and the line integral of the vector field along the boundary of the surface. In Eq. (1.13), the sum of the relative phases, i.e., the Berry phase γ_L, plays the role of the line integral, whereas the double sum of the Berry fluxes plays the role of the surface integral. There is an important difference with respect to the Stokes theorem, namely, the equality of the total Berry flux and the Berry phase is not guaranteed: Eq. (1.13) only tells that they are either equal or have a difference of 2π times an integer.

1.1.4 The Chern Number is the Complete Berry Flux Over a Surface

Consider states in a Hilbert space arranged on a grid as above, $|\Psi_{n,m}\rangle$, with $n,m\in\mathbb{Z}$, $1\le n\le N$, and $1\le m\le M$, but now imagine this grid to be on the surface of a torus. We use the same definition for the Berry flux per plaquette as in (1.11), but now with $n \bmod N+1$ in place of $n+1$ and $m \bmod M+1$ in place of $m+1$.

The product of the Berry flux phase factors of all plaquettes is now 1,

$$\prod_{m=1}^{M}\prod_{n=1}^{N}e^{-iF_{nm}}=1. \tag{1.14}$$

The same derivation can be applied as for Eq. (1.13) above, but now every edge is an internal edge, and so all contributions to the product cancel.

The Chern number Q associated to our structure is defined via the sum of the Berry fluxes of all the plaquettes forming the closed torus surface:

$$Q=\frac{1}{2\pi}\sum_{nm}F_{nm}. \tag{1.15}$$

The fact that the Chern number Q is defined via the gauge invariant Berry fluxes ensures that Q itself is gauge invariant. Furthermore, taking the arg of Eq. (1.14) proves that the Chern number Q is an integer.

It is worthwhile to look a little deeper into the discrete formula for the Chern number. We can define modified Berry fluxes $\tilde{F}_{nm}$ as

$$\tilde{F}_{nm}=\gamma_{(n,m),(n+1,m)}+\gamma_{(n+1,m),(n+1,m+1)}+\gamma_{(n+1,m+1),(n,m+1)}+\gamma_{(n,m+1),(n,m)}. \tag{1.16}$$

Since each edge is shared between two neighboring plaquettes, the sum of the modified Berry fluxes over all plaquettes vanishes,

$$\sum_{m=1}^{M}\sum_{n=1}^{N}\tilde{F}_{nm}=0. \tag{1.17}$$

If for some n,m, we have $-\pi\leq\tilde{F}_{nm}<\pi$, then $\tilde{F}_{nm}=F_{nm}$. However, $\tilde{F}_{nm}$ can be outside the range $[-\pi,\pi)$: then as the logarithm is taken in Eq. (1.10), F_{nm} is taken back into $[-\pi,\pi)$ by adding a (positive or negative) integer multiple of 2π. In that case, we say the plaquette nm contains a number $Q_{nm}\in\mathbb{Z}$ of vortices, with

$$Q_{nm}=\frac{F_{nm}-\tilde{F}_{nm}}{2\pi}\in\mathbb{Z}. \tag{1.18}$$

We have found a simple picture for the Chern number: *The Chern number Q – the sum of the Berry fluxes of all the plaquettes of a closed surface – is the number of vortices on the surface*,

$$Q=\frac{1}{2\pi}\sum_{nm}F_{nm}=\sum_{nm}Q_{nm}\in\mathbb{Z}. \tag{1.19}$$

Although we proved it here for the special case of a torus, the derivation is easily generalized to all orientable closed surfaces. We focused on the torus, because this construction can be used as a very efficient numerical recipe to discretize and calculate the (continuum) Chern number of a 2-dimensional insulator [7], to be defined in Sect. 1.2.4.

1.2 Continuum Case

We now assume that instead of a discrete set of states, $\{|\Psi_j\rangle\}$, we have a continuum, $|\Psi(\mathbf{R})\rangle$, where the $\mathbf{R}$'s are elements of some D-dimensional parameter space $\mathscr{P}$.

1.2.1 Berry Connection

We take a smooth directed curve $\mathscr{C}$, i.e., a *path* in the parameter space $\mathscr{P}$,

$$\mathscr{C}:[0,1)\to\mathscr{P},\quad t\mapsto\mathbf{R}(t). \tag{1.20}$$

We assume that all components of $|\Psi(\mathbf{R})\rangle$ are smooth, at least in an open neighborhood of the the curve $\mathscr{C}$. The relative phase between two neighbouring states on the curve $\mathscr{C}$, corresponding to the parameters $\mathbf{R}$ and $\mathbf{R}+d\mathbf{R}$, is

$$e^{-i\Delta\gamma}=\frac{\langle\Psi(\mathbf{R})\,|\,\Psi(\mathbf{R}+d\mathbf{R})\rangle}{|\langle\Psi(\mathbf{R})\,|\,\Psi(\mathbf{R}+d\mathbf{R})\rangle|};\qquad \Delta\gamma=i\,\langle\Psi(\mathbf{R})|\,\nabla_{\mathbf{R}}\,|\Psi(\mathbf{R})\rangle\cdot d\mathbf{R}, \tag{1.21}$$

obtained to first order in $d\mathbf{R}\to 0$. The quantity multiplying $d\mathbf{R}$ on the right-hand side defines the *Berry connection*,

$$\mathbf{A}(\mathbf{R})=i\,\langle\Psi(\mathbf{R})\,|\,\nabla_{\mathbf{R}}\Psi(\mathbf{R})\rangle=-\mathrm{Im}\,\langle\Psi(\mathbf{R})\,|\,\nabla_{\mathbf{R}}\Psi(\mathbf{R})\rangle\,. \tag{1.22}$$

Here $|\nabla_{\mathbf{R}}\Psi(\mathbf{R})\rangle$ is defined by requiring for every Hilbert space vector $|\Phi\rangle$, that

$$\langle\Phi|\,\nabla_{\mathbf{R}}\Psi(\mathbf{R})\rangle = \nabla_{\mathbf{R}}\,\langle\Phi|\,\Psi(\mathbf{R})\rangle\,. \tag{1.23}$$

The second equality in Eq. (1.22) follows from the conservation of the norm, $\nabla_{\mathbf{R}}\,\langle\Psi(\mathbf{R})\,|\,\Psi(\mathbf{R})\rangle = 0$.

We have seen in the discrete case that the relative phase of two states is not gauge invariant; neither is the Berry connection. Under a gauge transformation, it changes as

$$|\Psi(\mathbf{R})\rangle \to e^{i\alpha(\mathbf{R})}\,|\Psi(\mathbf{R})\rangle: \qquad \mathbf{A}(\mathbf{R}) \to \mathbf{A}(\mathbf{R}) - \nabla_{\mathbf{R}}\alpha(\mathbf{R}). \tag{1.24}$$

1.2.2 Berry Phase

Consider a closed directed curve $\mathscr{C}$, i.e., a loop, in parameter space. The Berry phase along the loop is defined as

$$\gamma(\mathscr{C}) = -\arg\exp\left[-i\oint_{\mathscr{C}}\mathbf{A}\cdot d\mathbf{R}\right] \tag{1.25}$$

The Berry phase of a closed directed curve is gauge invariant, since it can be interpreted as a limiting case of the gauge invariant discrete Berry phase, Eq. (1.6).

1.2.3 Berry Curvature is Berry Flux Density

As in the discrete case above, we would like to express the gauge invariant Berry phase as a surface integral of a gauge invariant quantity. This quantity is the *Berry curvature*. Similarly to the discrete case, we consider a two-dimensional parameter space, and for simplicity denote the parameters as x and y. We take a simply connected region $\mathscr{F}$ in this two-dimensional parameter space, with the oriented boundary curve of this surface denoted by $\partial\mathscr{F}$, and consider the continuum Berry phase corresponding to the boundary.

Smoothness of the Manifold of States

Before relating the Berry phase to the Berry curvature, an important note on the manifold $|\Psi(\mathbf{R})\rangle$ of considered states is in order. From now on, we consider a manifold of states, living in our two-dimensional parameter space, that is smooth, in the sense that the map $\vec{R} \mapsto |\Psi(\mathbf{R})\rangle\,\langle\Psi(\mathbf{R})|$ is smooth. Importantly, this condition does not necessarily imply that that the function $\mathbf{R} \mapsto |\Psi(\mathbf{R})\rangle$, which is the wavefunction in a given gauge, is smooth. (For further discussion and examples, see Sect. 1.5.1.) Nevertheless, even if $\mathbf{R} \mapsto |\Psi(\mathbf{R})\rangle$ is not smooth in a point $\mathbf{R}_0$ of the parameter space, one can always find an alternative gauge, where the wavefunction $|\Psi'(\mathbf{R})\rangle$ is (i) locally smooth (smooth in the point $\mathbf{R}_0$), and (ii) locally generates the same map as $|\Psi(\mathbf{R})\rangle$, meaning $|\Psi'(\mathbf{R})\rangle\,\langle\Psi'(\mathbf{R})| = |\Psi(\mathbf{R})\rangle\,\langle\Psi(\mathbf{R})|$ in an infinitesimal neighborhood of $\mathbf{R}_0$. An intuitive argument supporting the latter claim can be given using quantum-mechanical perturbation theory. Take the Hamiltonian $\hat{H}(\mathbf{R}) = -\,|\Psi(\mathbf{R})\rangle\,\langle\Psi(\mathbf{R})|$, which can be substituted

in the infinitesimal neighborhood of $\mathbf{R}_0$ with $\hat{H}(\mathbf{R}_0+\Delta\mathbf{R}) = \hat{H}(\mathbf{R}_0)+\Delta\mathbf{R}\cdot(\nabla\hat{H})(\mathbf{R}_0)$. According to first-order perturbation theory, the ground state of the latter is given by

$$\left|\Psi'(\mathbf{R}_0+\Delta\mathbf{R})\right\rangle = |\Psi(\mathbf{R}_0)\rangle - \sum_{n=2}^{D} |\Psi_n(\mathbf{R}_0)\rangle \langle\Psi_n(\mathbf{R}_0)|\Delta\mathbf{R}\cdot(\nabla\hat{H})(\mathbf{R}_0)|\Psi(\mathbf{R}_0)\rangle, \quad (1.26)$$

where the states $|\Psi_n(\mathbf{R}_0)\rangle$ ($n = 2,3,\ldots,D$), together with $|\Psi(\mathbf{R}_0)\rangle$, form a basis of the Hilbert space. On the one hand, Eq. (1.26) defines a function that is smooth in $\mathbf{R}_0$, hence the condition (i) above is satisfied. On the other hand, as $|\Psi'(\mathbf{R}_0+\Delta\mathbf{R})\rangle$ is the ground state of $\hat{H}(\mathbf{R}_0+\Delta\mathbf{R})$, condition (ii) is also satisfied.

Berry Phase and Berry Curvature

Now return to our original goal and try to express the Berry phase as a surface integral of a gauge invariant quantity. We start by relating the Berry phase to its discrete counterpart:

$$\oint_{\partial\mathscr{F}} \mathbf{A}\cdot d\mathbf{R} = \lim_{\Delta x,\Delta y\to 0} \gamma_{\partial\mathscr{F}}, \quad (1.27)$$

where we discretize the parameter space using a square grid of steps Δx, Δy, and express the integral as the discrete Berry phase $\gamma_{\partial\mathscr{F}}$ of a loop approximating $\partial\mathscr{F}$, in the limit of an infinitesimally fine grid. Then, from Eq. (1.27) and the Stokes-type theorem in Eq. (1.13), we obtain

$$\exp\left[-i\oint_{\partial\mathscr{F}} \mathbf{A}\cdot d\mathbf{R}\right] = \lim_{\Delta x,\Delta y\to 0} e^{-i\sum_{nm} F_{nm}}, \quad (1.28)$$

where the nm sum goes for the plaquettes forming the open surface $\mathscr{F}$. Furthermore, let us take a wavefunction $|\Psi'(\mathbf{R})\rangle$ and the corresponding Berry connection $\mathbf{A}'$ that is smooth in the plaquette nm; this could be $|\Psi(\mathbf{R})\rangle$ and $\mathbf{A}$ if that was already smooth. Then, due to the gauge invariance of the Berry flux we have

$$e^{-iF_{nm}} = e^{-iF'_{nm}}, \quad (1.29)$$

where F'_{nm} is the Berry flux corresponding to the locally smooth gauge. Furthermore, in the limit of an infinitely fine grid it holds that

$$\begin{aligned} F'_{nm} = {} & A'_x\left(x_n+\frac{\Delta x}{2}, y_m\right)\Delta x + A'_y\left(x_{n+1}, y_m+\frac{\Delta y}{2}\right)\Delta y \\ & - A'_x\left(x_n+\frac{\Delta x}{2}, y_{m+1}\right)\Delta x - A'_y\left(x_n, y_m+\frac{\Delta y}{2}\right)\Delta y. \end{aligned} \quad (1.30)$$

Taylor expansion of the Berry connection around $\mathbf{R}_{nm} = \left(x_n+\frac{\Delta x}{2}, y_m+\frac{\Delta y}{2}\right)$ to first order yields

$$F'_{nm} = \left[\partial_x A'_y(\mathbf{R}_{nm}) - \partial_y A'_x(\mathbf{R}_{nm})\right]\Delta x\Delta y. \quad (1.31)$$

Thereby, with the definition of the *Berry curvature* as

$$B_{nm} = \lim_{\Delta x, \Delta y \to 0} \frac{F'_{nm}}{\Delta x \Delta y}, \tag{1.32}$$

we obtain a quantity that is gauge invariant, as it is defined via the gauge invariant Berry flux, and is related to the Berry connection via

$$B = \partial_x A'_y(\mathbf{R}_{nm}) - \partial_y A'_x(\mathbf{R}_{nm}). \tag{1.33}$$

We can rephrase Eq. (1.31) as follows: the Berry flux for the *nm* plaquette is expressed as the product of the Berry curvature on the plaquette and the surface area of the plaquette.

Substituting Eqs. (1.29) and (1.31) into Eq. (1.28) yields

$$\exp\left[-i \oint_{\partial \mathscr{F}} \mathbf{A} \cdot d\mathbf{R}\right] = \exp\left[-i \int_{\mathscr{F}} B(x,y) dx dy\right], \tag{1.34}$$

which is the continuum version of the result (1.13). Equation (1.34) can also be rephrased as

$$\gamma(\partial \mathscr{F}) = -\arg e^{-i \int_{\mathscr{F}} B(x,y) dx dy}. \tag{1.35}$$

A Special Case Where the Usual Stokes Theorem Works

A shortcut towards a stronger result than Eq. (1.34) is offered in the special case when $|\Psi(\mathbf{R})\rangle$ is smooth in the neighborhood of the open surface $\mathscr{F}$. Then, a direct application of the two-dimensional Stokes theorem implies

$$\oint_{\partial \mathscr{F}} \mathbf{A} \cdot d\mathbf{R} = \int_{\mathscr{F}} (\partial_x A_y - \partial_y A_x) dx dy = \int_{\mathscr{F}} B dx dy \tag{1.36}$$

Summarizing Eqs. (1.34) and (1.36), we can say that line integral of the Berry connection equals the surface integral of the Berry curvature if the set of states $|\Psi(\mathbf{R})\rangle$ is smooth in the neighborhood of $\mathscr{F}$, but they might differ by an integer multiple of 2π otherwise.

The Case of the Three-Dimensional Parameter Space

We briefly discuss the case of a three-dimensional parameter space. This will be particularly useful in the context of two-level systems. Starting with the case when the gauge $|\Psi(\mathbf{R})\rangle$ on the two-dimensional open surface $\mathscr{F}$ embedded in the three-dimensional parameter space is smooth in the neighborhood of $\mathscr{F}$, we can directly apply the three-dimensional Stokes theorem to convert the line integral of $\mathbf{A}$ to the surface integral of the curl of $\mathbf{A}$ to obtain

$$\oint_{\partial \mathscr{F}} \mathbf{A} \cdot d\mathbf{R} = \int_{\mathscr{F}} \mathbf{B} \cdot d\vec{S}, \tag{1.37}$$

where the Berry curvature is defined as the vector field $\mathbf{B}(\mathbf{R})$ via

$$\mathbf{B}(\mathbf{R}) = \nabla_{\mathbf{R}} \times \mathbf{A}(\mathbf{R}), \tag{1.38}$$

which is gauge invariant as in the two-dimensional case. Even if $|\Psi(\mathbf{R})\rangle$ is not smooth on $\mathscr{F}$, the relation

$$\gamma(\partial\mathscr{F}) = -\arg e^{-i\oint_{\partial\mathscr{F}}\mathbf{A}\cdot d\mathbf{R}} = -\arg e^{-i\int_{\mathscr{F}}\mathbf{B}\cdot d\vec{S}} \tag{1.39}$$

holds, similarly to the two-dimensional result Eq. (1.34).

Note furthermore that the Berry phase $\gamma(\partial\mathscr{F})$ for a fixed boundary curve $\partial\mathscr{F}$ is not only gauge invariant, but also invariant against continuous deformations of the two-dimensional surface $\mathscr{F}$ embedded in three dimensions, as long as the Berry curvature is smooth everywhere along the way.

We also remark that although we used the three-dimensional notation here, but the above results can be generalized for any dimensionality of the parameter space.

The notation $\mathbf{A}$ and $\mathbf{B}$ for the Berry connection and Berry curvature suggest that they are much like the vector potential and the magnetic field. This is a useful analogy, for instance, $\nabla_{\mathbf{R}}\mathbf{B} = 0$, from the definition (1.38). Nevertheless, it is not true that in every problem where the Berry curvature is nonzero, there is a physical magnetic field.

1.2.4 Chern Number is the Integral of the Berry Curvature

In the discrete case, we defined the Chern number as a sum of Berry fluxes for a square lattice living on a torus (or any other orientable closed surface). Here, we take a continuum parameter space that has the topology of a torus. The motivation is the Brillouin zone of a two-dimensional lattice representing a solid crystalline material. The Brillouin zone has a torus topology, as the momentum vectors (k_x, k_y), $(k_x + 2\pi, ky)$, and $(k_x, k_y + 2\pi)$ are equivalent.

Quite naturally, in the continuum definition of the Chern number, the sum of Berry fluxes is replaced by the surface integral of the Berry curvature:

$$Q = -\frac{1}{2\pi}\int_{\mathscr{P}} B dx dy. \tag{1.40}$$

As this can be interpreted as a continuum limit of the discrete Chern number, it inherits the properties of the latter: the continuum Chern number is a gauge invariant integer.

For future reference, let us lay down the notation to be used for calculating the Chern numbers of electronic energy bands in two-dimensional crystals. Consider a square lattice for simplicity, which has a square-shaped Brillouin zone as well. Our parameter space $\mathscr{P}$ is the two-dimensional Brillouin zone now, which has a torus topology as discussed above. The parameters are the Cartesian components $k_x, k_y \in [-\pi, \pi)$ of the momentum vector $\vec{k}$. The electronic energy bands and the corresponding electron wavefunctions can be obtained from the bulk momentum-space Hamiltonian $\hat{H}(k_x, k_y)$. The latter defines the Schrödinger equation

$$\hat{H}(\mathbf{k})\,|u_n(\mathbf{k})\rangle = E_n(\mathbf{k})\,|u_n(\mathbf{k})\rangle\,, \tag{1.41}$$

where $n = 1, 2, \ldots$ is the band index, which has as many possible values as the dimension of the Hilbert space of the internal degree of freedom of our lattice model. Note that defining the Berry connection, the Berry curvature and the Chern number for the

nth band is possible only if that band is separated from other bands by energy gaps. The Berry connection of the nth band, in line with the general definition (1.22), reads

$$A_j^{(n)}(\mathbf{k}) = i \langle u_n(\mathbf{k}) | \partial_{k_j} | u_n(\mathbf{k}) \rangle , \quad \text{for} \quad j = x, y. \tag{1.42}$$

The Chern number of the nth band, in correspondence with Eqs. (1.93) and (1.33), reads

$$Q^{(n)} = -\frac{1}{2\pi} \int_{BZ} dk_x dk_y \left(\frac{\partial A_y^{(n)}}{\partial k_x} - \frac{\partial A_x^{(n)}}{\partial k_y} \right). \tag{1.43}$$

Certain approximations of the band structure theory of electrons provide low-dimensional momentum-space Hamiltonians that can be diagonalized analytically, allowing for an analytical derivation of the Chern numbers of the electronic bands. More often, however, the electronic wave functions are obtained from numerical techniques on a finite-resolution grid of (k_x, k_y) points in the Brillouin zone. In that case, the Chern number of a chosen band can still be effectively evaluated using the discrete version of its definition (1.15).

The Chern number of a band of an insulator is a topological invariant in the following sense. One can imagine that the Hamiltonian describing the electrons on the lattice is deformed adiabatically, that is, continuously and with the energy gaps separating the nth band from the other bands kept open. In this case, the Berry curvature varies continuously, and therefore its integral for the Brillouin zone, which is the Chern number, cannot change as the value of the latter is restricted to integers. If the deformation of the crystal Hamiltonian is such that some energy gaps separating the nth band from a neighboring band is closed and reopened, that is, the deformation of the Hamiltonian is not adiabatic, then the Chern number might change. In this sense, the Chern number is a similar topological invariant for two-dimensional lattice models as the winding number is for the one-dimensional SSH model.

1.3 Berry Phase and Adiabatic Dynamics

In most physical situations of interest, the set of states whose geometric features (Berry phases) we are interested in are eigenstates of some Hamiltonian $\hat{H}$. Take a physical system with D real parameters that are gathered into a formal vector $\mathbf{R} = (R_1, R_2, \ldots, R_D)$. The Hamiltonian is a smooth function $\hat{H}(\mathbf{R})$ of the parameters, at least in the region of interest. We order the eigenstates of the Hamiltonian according to the energies $E_n(\mathbf{R})$,

$$\hat{H}(\mathbf{R}) | n(\mathbf{R}) \rangle = E_n(\mathbf{R}) | n(\mathbf{R}) \rangle . \tag{1.44}$$

We call the set of eigenstates $|n(\mathbf{R})\rangle$ the *snapshot basis*.

The definition of the snapshot basis involves *gauge fixing*, i.e., specifying the otherwise arbitrary phase prefactor for every $|n(\mathbf{R})\rangle$. This can be a tricky issue: even in cases where a gauge exists where all elements of the snapshot basis are smooth functions of the parameters, this gauge might be very challenging to construct.

We consider the following problem. We assume that the system is initialized with $\mathbf{R} = \mathbf{R}_0$ and in an eigenstate $|n(\mathbf{R}_0)\rangle$ that is in the discrete part of the spectrum, i.e., $E_n(\mathbf{R}) - E_{n-1}(\mathbf{R})$ and $E_{n+1}(\mathbf{R}) - E_n(\mathbf{R})$ are nonzero. At time $t = 0$ we thus have

$$\mathbf{R}(t=0) = \mathbf{R}_0; \qquad |\psi(t=0)\rangle = |n(\mathbf{R}_0)\rangle . \tag{1.45}$$

Now assume that during the time $t = 0 \to T$ the parameter vector $\mathbf{R}$ is slowly changed: $\mathbf{R}$ becomes $\mathbf{R}(t)$, and the values of $\mathbf{R}(t)$ define a continuous curve $\mathscr{C}$. Also, assume that $|n(\mathbf{R})\rangle$ is smooth along the curve $\mathscr{C}$. The state of the system evolves according to the time-dependent Schrödinger equation:

$$i\frac{d}{dt}|\psi(t)\rangle = \hat{H}(\mathbf{R}(t))\,|\psi(t)\rangle . \tag{1.46}$$

Further, assume that $\mathbf{R}$ is varied in such a way that at all times the energy gaps around the state $|n(\mathbf{R}(t))\rangle$ remain finite. We can then choose the rate of variation of $\mathbf{R}(t)$ along the path $\mathscr{C}$ to be slow enough compared to the frequencies corresponding to the energy gap, so the *adiabatic approximation* holds. In that case, the system remains in the energy eigenstate $|n(\mathbf{R}(t))\rangle$, only picking up a phase. We are now going to find this phase.

By virtue of the adiabatic approximation, we take as Ansatz

$$|\psi(t)\rangle = e^{i\gamma_n(t)} e^{-i\int_0^t E_n(\mathbf{R}(t'))dt'}\,|n(\mathbf{R}(t))\rangle . \tag{1.47}$$

For better readability, in the following we often drop the t argument where this leads to no confusion. The time derivative of Eq. (1.47) reads

$$i\frac{d}{dt}|\psi(t)\rangle = e^{i\gamma_n} e^{-i\int_0^t E_n(\mathbf{R}(t'))dt'} \left(-\frac{d\gamma_n}{dt}\,|n(\mathbf{R})\rangle + E_n(\mathbf{R})\,|n(\mathbf{R})\rangle + i\left|\tfrac{d}{dt}n(\mathbf{R})\right\rangle \right). \tag{1.48}$$

To show what we mean by $\left|\frac{d}{dt}n(\mathbf{R}(t))\right\rangle$, we write it out explicitly in terms of a fixed basis, that of the eigenstates at $\mathbf{R} = \mathbf{R}_0$:

$$|n(\mathbf{R})\rangle = \sum_m c_m(\mathbf{R})\,|m(\mathbf{R}_0)\rangle ; \tag{1.49}$$

$$\left|\tfrac{d}{dt}n(\mathbf{R}(t))\right\rangle = \frac{d\mathbf{R}}{dt}\cdot|\nabla_{\mathbf{R}} n(\mathbf{R})\rangle = \frac{d\mathbf{R}}{dt}\sum_m \nabla_{\mathbf{R}} c_m(\mathbf{R})\,|m(\mathbf{R}_0)\rangle . \tag{1.50}$$

We insert the Ansatz (1.47) into the right hand side of the Schrödinger equation (1.46), use the snapshot eigenvalue relation (1.44), simplify and reorder the Schrödinger equation, and obtain

$$-\frac{d\gamma_n}{dt}\,|n(\mathbf{R})\rangle + i\left|\tfrac{d}{dt}n(\mathbf{R})\right\rangle = 0. \tag{1.51}$$

Multiplying from the left by $\langle n(\mathbf{R})|$, and using Eq. (1.50), we obtain

$$\frac{d}{dt}\gamma_n(t) = i\left\langle n(\mathbf{R}(t))\right|\left.\tfrac{d}{dt}n(\mathbf{R}(t))\right\rangle = \frac{d\mathbf{R}}{dt}\,i\,\langle n(\mathbf{R})\,|\,\nabla_{\mathbf{R}} n(\mathbf{R})\rangle . \tag{1.52}$$

We have found that for the curve $\mathscr{C}$ in parameter space, traced out by $\mathbf{R}(t)$, there is an adiabatic phase $\gamma_n(\mathscr{C})$, which reads

$$\gamma_n(\mathscr{C}) = \int_{\mathscr{C}} i \langle n(\mathbf{R}) \,|\, \nabla_{\mathbf{R}} n(\mathbf{R}) \rangle \, d\mathbf{R}. \tag{1.53}$$

A related result is obtained after a similar derivation, if the parameter space of the $\mathbf{R}$ points is omitted and the snapshot basis $|n(t)\rangle$ is parametrized directly by the time variable. Then, the adiabatic phase is

$$\gamma_n(t) = \int_0^t i \left\langle n(t') \,\middle|\, \partial_{t'} n(t') \right\rangle dt'. \tag{1.54}$$

Equation (1.53) allows us to formulate the key message of this section as the following. Consider the case of an adiabatic and *cyclic* change of the Hamiltonian, that is, when the curve $\mathscr{C}$ is closed, $\mathbf{R}(T) = \mathbf{R}_0$. In this case, the adiabatic phase reads

$$\gamma_n(\mathscr{C}) = \oint_{\mathscr{C}} i \langle n(\mathbf{R}) \,|\, \nabla_{\mathbf{R}} n(\mathbf{R}) \rangle \, d\mathbf{R}. \tag{1.55}$$

Therefore, the adiabatic phase picked up by the state during a cyclic adiabatic change of the Hamiltonian is equivalent to the Berry phase corresponding to the closed oriented curve representing the Hamiltonian's path in the parameter space.

Two further remarks are in order. First, on the face of it, our derivation seems to do too much. It seems that we have produced an exact solution of the Schrödinger equation. Where did we use the adiabatic approximation? In fact, Eq. (1.52) does not imply Eq. (1.51). For the more complete derivation, showing how the nonadiabatic terms appear, see [9].

The second remark concerns the measurability of the Berry phase. The usual way to experimentally detect phases is by an interferometric setup. This means coherently splitting the wavefunction of the system into two parts, taking them through two adiabatic trips in parameter space, via $\mathbf{R}(t)$ and $\mathbf{R}'(t)$, and bringing the parts back together. The interference only comes from the overlap between the states: it is maximal if $|n(\mathbf{R}(T))\rangle = |n(\mathbf{R}'(T))\rangle$, which is typically ensured if $\mathbf{R}(T) = \mathbf{R}'(T)$. The difference in the adiabatic phases γ_n and γ'_n is the adiabatic phase associated with the closed loop $\mathscr{C}$, which is the path obtained by going forward along $t = 0 \to T : \mathbf{R}(t)$, then coming back along $t = T \to 0 : \mathbf{R}'(t)$.

1.4 Berry's Formulas for the Berry Curvature

Berry provided[4] two practical formulas for the Berry curvature. Here we present them in a form corresponding to a three-dimensional parameter space. To obtain the two-dimensional case, where the Berry curvature B is a scalar, one can identify the latter with the component B_z of the three-dimensional case treated below; for generalization to higher than 3 dimensions, see the discussion in Berry's paper[4].

The first of Berry's formulas reads,

$$B_j = -\mathrm{Im}\, \varepsilon_{jkl}\, \partial_k \langle n |\, \partial_l n\rangle = -\mathrm{Im}\, \varepsilon_{jkl}\, \langle \partial_k n |\, \partial_l n\rangle + 0, \tag{1.56}$$

where the second term is 0 because $\partial_k \partial_l = \partial_l \partial_k$ but $\varepsilon_{jkl} = -\varepsilon_{jlk}$.

To obtain Berry's second formula, inserting a resolution of identity in the snapshot basis in the above equation, we obtain

$$\mathbf{B}^{(n)} = -\mathrm{Im} \sum_{n' \neq n} \left\langle \nabla n \,\middle|\, n' \right\rangle \times \left\langle n' \,\middle|\, \nabla n \right\rangle, \tag{1.57}$$

where the parameter set $\mathbf{R}$ is suppressed for brevity. The term with $n' = n$ is omitted from the sum, as it is zero, since because of the conservation of the norm, $\langle \nabla n |\, n\rangle = -\langle n |\, \nabla n\rangle$. To calculate $\langle n' |\, \nabla n\rangle$, start from the definition of the eigenstate $|n\rangle$, act on both sides with ∇, and then project unto $|n'\rangle$:

$$\hat{H}\,|n\rangle = E_n\,|n\rangle\,; \tag{1.58}$$

$$(\nabla \hat{H})\,|n\rangle + \hat{H}\,|\nabla n\rangle = (\nabla E_n)\,|n\rangle + E_n\,|\nabla n\rangle\,; \tag{1.59}$$

$$\left\langle n' \middle|\, \nabla \hat{H}\,|n\rangle + \left\langle n' \middle|\, \hat{H}\,|\nabla n\right\rangle = 0 + E_n \left\langle n' \,\middle|\, \nabla n\right\rangle. \tag{1.60}$$

Act with $\hat{H}$ towards the left in Eq. (1.60), rearrange, substitute into (1.57), and you obtain the second form of the Berry curvature, which is manifestly gauge invariant:

$$\mathbf{B}^{(n)} = -\mathrm{Im} \sum_{n' \neq n} \frac{\langle n |\, \nabla \hat{H}\,|n'\rangle \times \langle n' |\, \nabla \hat{H}\,|n\rangle}{(E_n - E_{n'})^2}. \tag{1.61}$$

This shows that the monopole sources of the Berry curvature, if they exist, are the points of degeneracy.

A direct consequence of Eq. (1.61), is that the sum of the Berry curvatures of all eigenstates of a Hamiltonian is zero. If all the spectrum of $\hat{H}(\mathbf{R})$ is discrete along a closed curve $\mathscr{C}$, then one can add up the Berry phases of all the energy eigenstates.

$$\begin{aligned}
\sum_n \mathbf{B}^{(n)} &= -\mathrm{Im} \sum_n \sum_{n' \neq n} \frac{\langle n |\, \nabla_{\mathbf{R}} \hat{H}\,|n'\rangle \times \langle n' |\, \nabla_{\mathbf{R}} \hat{H}\,|n\rangle}{(E_n - E_{n'})^2} \\
&= -\mathrm{Im} \sum_n \sum_{n' < n} \frac{1}{(E_n - E_{n'})^2} \Big(\langle n |\, \nabla_{\mathbf{R}} \hat{H}\,|n'\rangle \times \langle n' |\, \nabla_{\mathbf{R}} \hat{H}\,|n\rangle \\
&\quad + \left\langle n' \middle|\, \nabla_{\mathbf{R}} \hat{H}\,|n\rangle \times \langle n |\, \nabla_{\mathbf{R}} \hat{H}\,|n'\right\rangle \Big) = 0.
\end{aligned} \tag{1.62}$$

The last equation holds because $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ for any two vectors $\vec{a}, \vec{b}$.

1.5 Example: The Two-Level System

So far, most of the discussion on the Berry phase and the related concepts have been kept rather general. In this section, we illustrate these concepts via the simplest non-trivial example, that is, the two-level system.

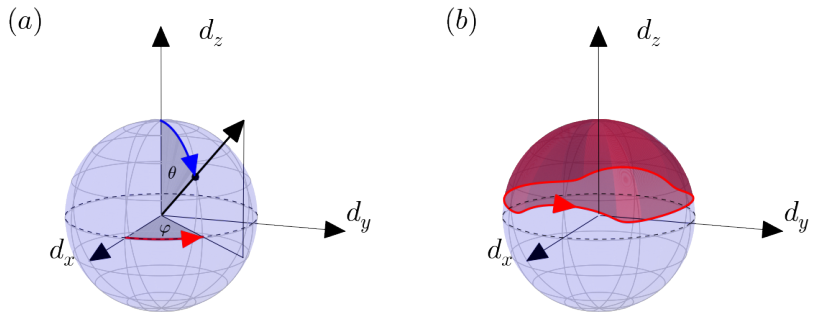

Figure 1.2: The Bloch sphere. A generic traceless gapped two-level Hamiltonian is a linear combination of Pauli matrices, $\hat{H}(\mathbf{d}) = \mathbf{d}\cdot\hat{\sigma}$. This can be identified with a point in $\mathbb{R}^3\backslash\{0\}$. The eigenenergies are given by the distance of the point from the origin, the eigenstates depend only on the direction of the vector $\mathbf{d}$, i.e., on the angles θ and φ, as defined in subfigure (a) and in Eq. (1.64) The Berry phase of a closed curve $\mathscr{C}$ is half the area enclosed by the curve when it is projected onto the surface of the Bloch sphere.

1.5.1 No Continuous Global Gauge

Consider a Hamiltonian describing a two-level system:

$$\hat{H}(\mathbf{d}) = d_x\hat{\sigma}_x + d_y\hat{\sigma}_y + d_z\hat{\sigma}_z = \mathbf{d}\cdot\hat{\sigma}, \tag{1.63}$$

with $\mathbf{d} = (d_x, d_y, d_z) \in \mathbb{R}^3\backslash\{0\}$. Here, the vector $\mathbf{d}$ plays the role of the parameter $\mathbf{R}$ in of preceding sections, and the parameter space is the punctured three-dimensional Euclidean space $\mathbb{R}^3\backslash\{0\}$, to avoid the degenerate case of the energy spectrum. Note the absence of a term proportional to σ_0: this would play no role in adiabatic phases. Because of the anticommutation relations of the Pauli matrices, the Hamiltonian above squares to a multiple of the identity operator, $\hat{H}(\mathbf{d})^2 = \mathbf{d}^2\sigma_0$. Thus, the eigenvalues of $\hat{H}(\mathbf{d})$ have to have absolute value $|\mathbf{d}|$.

A practical graphical representation of $\hat{H}(\mathbf{d})$ is the Bloch sphere, shown in Fig. 1.2. The spherical angles $\theta \in [0,\pi)$ and $\varphi \in [0,2\pi)$ are defined as

$$\cos\theta = \frac{d_z}{|\mathbf{d}|}; \qquad\qquad e^{i\varphi} = \frac{d_x + id_y}{\sqrt{d_x^2 + d_y^2}}. \tag{1.64}$$

We denote the two eigenstates of the Hamiltonian $\hat{H}(\mathbf{d})$ by $|+_\mathbf{d}\rangle$ and $|-_\mathbf{d}\rangle$, with

$$\hat{H}(\mathbf{d})\,|\pm_\mathbf{d}\rangle = \pm|\mathbf{d}|\,|\pm_\mathbf{d}\rangle\,. \tag{1.65}$$

These eigenstates depend on the direction of the 3-dimensional vector $\mathbf{d}$, but not on its

length. The eigenstate with $E = +|\mathbf{d}|$ of the corresponding Hamiltonian is:

$$|+_{\mathbf{d}}\rangle = e^{i\alpha(\theta,\varphi)} \begin{pmatrix} e^{-i\varphi/2}\cos\theta/2 \\ e^{i\varphi/2}\sin\theta/2 \end{pmatrix}, \tag{1.66}$$

while the eigenstate with $E = -|\mathbf{d}|$ is $|-_{\mathbf{d}}\rangle = e^{i\beta(\mathbf{d})}|+_{-\mathbf{d}}\rangle$. The choice of the phase factors α and β above corresponds to fixing a gauge. We will now review a few gauge choices.

Consider fixing $\alpha(\theta,\varphi) = 0$ for all θ,φ. This is a very symmetric choice, in this way in formula (1.66), we find $\theta/2$ and $\varphi/2$. There is problem, however, as you can see if you consider a full circle in parameter space: at any fixed value of θ, let $\varphi = 0 \to 2\pi$. We should come back to the same Hilbert space vector, and we do, but we also pick up a phase of -1. We can either say that this choice of gauge led to a discontinuity at $\varphi = 0$, or that our representation is not single-valued. We now look at some attempts at fixing these problems, to find a gauge that is both continuous and single valued.

As a first attempt, let us fix $\alpha = \varphi/2$; denoting this gauge by subscript S, we have

$$|+_{\mathbf{d}}\rangle_S = \begin{pmatrix} \cos\theta/2 \\ e^{i\varphi}\sin\theta/2 \end{pmatrix}. \tag{1.67}$$

The phase prefactor now gives an additional factor of -1 as we make the circle in φ at fixed θ, and so it seems we have a continuous, single valued representation. There are two tricky points, however: the North Pole, $\theta = 0$, and the South Pole, $\theta = \pi$. At the North Pole, $|(0,0,1)\rangle_S = (1,0)$ no problems. This gauge is problematic at the South Pole, however (which explains the choice of subscript): there, $|(0,0,-1)\rangle_S = (0,e^{i\varphi})$, the value of the wavefunction depends on which direction we approach the South Pole from.

We can try to solve the problem at the South Pole by choosing $\alpha = -\varphi/2$, which gives us

$$|+_{\mathbf{d}}\rangle_N = \begin{pmatrix} e^{-i\varphi}\cos\theta/2 \\ \sin\theta/2 \end{pmatrix}. \tag{1.68}$$

As you can probably already see, this representation runs into trouble at the North Pole: $|(0,0,1)\rangle_N = (e^{-i\varphi},0)$, whose value depends on the direction of approach to the North Pole.

We can try to overcome the problems at the poles by taking linear combinations of $|+_{\mathbf{d}}\rangle_S$ and $|+_{\mathbf{d}}\rangle_N$, with prefactors that vanish at the South and North Poles, respectively. A family of options is:

$$|+_{\mathbf{d}}\rangle_\chi = e^{i\chi}\sin\frac{\theta}{2}|+_{\mathbf{d}}\rangle_S + \cos\frac{\theta}{2}|+_{\mathbf{d}}\rangle_N \tag{1.69}$$

$$= \begin{pmatrix} \cos\frac{\theta}{2}(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}e^{i\chi}e^{-i\varphi}) \\ \sin\frac{\theta}{2}e^{i\varphi}(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}e^{i\chi}e^{-i\varphi}) \end{pmatrix}. \tag{1.70}$$

This is single valued everywhere, solves the problems at the Poles. However, it has its own problems: somewhere on the Equator, at $\theta = \pi/2$, $\varphi = \chi \pm \pi$, its norm disappears.

It is not all that surprising that we could not find a well-behaved gauge: there is none. By the end of this chapter, it should be clear, why.

1.5.2 Calculating the Berry Curvature and the Berry Phase

Consider the two-level system as defined in the previous section. Take a closed curve $\mathscr{C}$ in the parameter space $\mathbb{R}^3\backslash\{0\}$. We are going to calculate the Berry phase γ_- of the $|-_{\mathbf{d}}\rangle$ eigenstate on this curve:

$$\gamma_-(\mathscr{C}) = \oint_{\mathscr{C}} \mathbf{A}(\mathbf{d}) d\mathbf{d}, \tag{1.71}$$

with the Berry vector potential defined as

$$\mathbf{A}(\mathbf{d}) = i \langle -_{\mathbf{d}} | \nabla_{\mathbf{d}} | -_{\mathbf{d}} \rangle . \tag{1.72}$$

The calculation becomes straightforward if we use the Berry curvature,

$$\mathbf{B}(\mathbf{d}) = \nabla_{\mathbf{d}} \times \mathbf{A}(\mathbf{d}); \tag{1.73}$$

$$\gamma_-(\mathscr{C}) = \int_{\mathscr{S}} \mathbf{B}(\mathbf{d}) d\mathscr{S}, \tag{1.74}$$

where $\mathscr{S}$ is any surface whose boundary is the loop $\mathscr{C}$. (Alternatively, it is a worthwhile exercise to calculate the Berry phase directly in a fixed gauge, e.g., one of the three gauges introduced above.)

Specifically, we make use of Berry's gauge invariant formulation (1.61) of the Berry curvature, derived in the last chapter. In the case of the generic two-level Hamiltonian (1.63), Eq. (1.61) gives

$$\mathbf{B}^{\pm}(\mathbf{d}) = -\mathrm{Im} \frac{\langle \pm | \nabla_{\mathbf{d}} \hat{H} | \mp \rangle \times \langle \mp | \nabla_{\mathbf{d}} \hat{H} | \pm \rangle}{4\mathbf{d}^2}, \tag{1.75}$$

with

$$\nabla_{\mathbf{d}} \hat{H} = \hat{\sigma}. \tag{1.76}$$

To evaluate (1.75), we choose the quantization axis parallel to $\mathbf{d}$, thus the eigenstates simply read

$$|+_{\mathbf{d}}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \qquad |-_{\mathbf{d}}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.77}$$

The matrix elements can now be computed as

$$\langle -| \hat{\sigma}_x | + \rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1, \tag{1.78}$$

and similarly,

$$\langle -| \sigma_y | + \rangle = i; \tag{1.79}$$

$$\langle -| \sigma_z | + \rangle = 0. \tag{1.80}$$

So the cross product of the vectors reads

$$\langle -|\,\hat{\sigma}\,|+\rangle \times \langle +|\,\hat{\sigma}\,|-\rangle = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2i \end{pmatrix}. \tag{1.81}$$

This gives us for the Berry curvature,

$$\mathbf{B}^{\pm}(\mathbf{d}) = \pm \frac{\mathbf{d}}{|\mathbf{d}|} \frac{1}{2\mathbf{d}^2}. \tag{1.82}$$

We can recognize in this the field of a pointlike monopole source in the origin. Note however that this monopole exists in the abstract space of the vectors **d** and not in real space, as the magnetic monopoles of electrodynamics [CITATION?].

The Berry phase of the closed loop $\mathscr{C}$ in parameter space, according to Eq. (1.74), is the flux of the monopole field through a surface $\mathscr{S}$ whose boundary is $\mathscr{C}$. It is easy to convince yourself that this is half of the solid angle subtended by the curve,

$$\gamma_{-}(\mathscr{C}) = \frac{1}{2}\Omega_{\mathscr{C}}. \tag{1.83}$$

In other words, the Berry phase is half of the area enclosed by the image of $\mathscr{C}$, projected onto the surface of the unit sphere, as illustrated in Fig. 1.2.

What about the Berry phase of the other energy eigenstate? From Eq. (1.75), the corresponding Berry curvature $\mathbf{B}_{+}$ is obtained by inverting the order of the factors in the cross product: this flips the sign of the cross product. Therefore the Berry phases of the ground and excited state fulfil the relation

$$\gamma_{+}(\mathscr{C}) = -\gamma_{-}(\mathscr{C}). \tag{1.84}$$

One can see the same result on the Bloch sphere. Since $\langle +|\,-\rangle = 0$, the point corresponding to $|-\rangle$ is antipodal to the point corresponding to $|+\rangle$. Therefore, the curve traced by the $|-\rangle$ on the Bloch sphere is the inverted image of the curve traced by $|+\rangle$. These two curves have the same orientation, therefore the same area, with opposite signs.

1.6 Graphical Methods to Determine the Chern Number in Two-Band Models

We now discuss two useful graphical ways to determine the Chern number in two-band models, which are the simplest interesting cases. We consider a Hamiltonian of the form,

$$\hat{H}(\mathbf{k}) = \mathbf{d}(\mathbf{k})\hat{\sigma}; \quad \mathbf{k} \in \mathscr{M}; \tag{1.85}$$

where **k** is a parameter on a closed surface, i.e., a two-dimensional orientable manifold without boundaries. In the most useful example, $\mathscr{M}$ is the surface of a torus, but it

could also be, e.g., that of a sphere. We denote the parameter by $\mathbf{k}$, because in the cases of interest for us, this parameter will be the wavenumber of a two-dimensional hopping model, and $\mathbf{k}$ then lives on a toroidal Brillouin zone. We assume that the function $\mathbf{d}(\mathbf{k})$ is smooth.

In case the Hamiltonian $\hat{H}(\mathbf{k})$ is gapped, its two eigenstates define two well-defined manifolds of quantum states, two *bands*. We will denote the two eigenstates as

$$\hat{H}(\mathbf{k})\,|\pm(\mathbf{k})\rangle = \pm|\mathbf{d}(\mathbf{k})|\,|\pm(\mathbf{k})\rangle\,. \tag{1.86}$$

The gap condition is

$$\forall \mathbf{k} \in \mathscr{M}: \quad |\mathbf{d}(\mathbf{k})| > 0. \tag{1.87}$$

1.6.1 The Chern number is the number of intersections of the surface with a line of sight

One way to determine the Chern number is – similarly to what we did for the winding number in Chapter **??** – counting intersections with the line of sight to infinity. For this, consider the deformed closed surface in the space of $\mathbf{d}$ vectors, $\mathbb{R}^3\backslash\{0\}$, that is obtained by mapping the closed surface $\mathscr{M}$, using the function $\mathbf{d}(\mathbf{k})$. This is an oriented surface: its inside can be painted red, its outside, blue.

The Chern number of $|-\rangle$ (using the notation of Sect. **??**, of $|u_1(\mathbf{k})\rangle$) is the flux of $\mathbf{B}_-(\mathbf{d})$ through this torus. We have seen above that $\mathbf{B}_-(\mathbf{d})$ is the magnetic field of a monopole at the origin $\mathbf{d} = 0$. If the origin is on the inside of the torus, this flux is +1. If it is outside of the torus, it is 0. If the torus is turned inside out, and contains the origin, the flux is -1. The torus can also intersect itself, and therefore contain the origin any number of times.

One way to count the number of times the torus contains the origin is as follows. Take any line from the origin to infinity, and count the number of times it intersects the torus, with a $+1$ for intersecting from the inside, and a -1 for intersecting from the outside. The sum is independent of the shape of the line, as long as it goes all the way from the origin to infinity.

FIGURE HERE!

1.6.2 The Chern number is the number of skyrmions on the surface

In a two-band model, the Chern number can also be seen as the number of skyrmions on the closed surface $\mathscr{M}$. This is similar to the Chern number being the sum of vortices on the closed surface, as in Sect. 1.1.4.

Grid of states, obtained by cutting the surface open and flattening it.

3D unit vector at each grid point.

Surface enclosed.

1.7 Summary

Although the relative phase of two quantum states is gauge dependent, the relative phase along a loop of at least three states is gauge invariant. This is the Berry phase, which is the basis of the Berry flux, Berry curvature, and Chern number. These can be formulated for discretely and for continuously varying parameters.

- The Berry phase for a closed, discrete loop L in parameter space:

$$\gamma_L = -\arg \mathrm{Tr}\left(|\Psi_1\rangle\langle\Psi_1||\Psi_2\rangle\langle\Psi_2|\ldots|\Psi_N\rangle\langle\Psi_N|\right). \tag{1.88}$$

- The Berry flux on a plaquette p with boundary ∂p, directed counterclockwise,

$$F_p = \gamma_{\partial p}. \tag{1.89}$$

- The Chern number for a closed surface consisting of plaquettes, $\mathscr{S} = \{p|p \in \mathscr{S}\}$ is the total Berry flux over the surface, in units of 2π,

$$Q_{\mathscr{S}} = \frac{1}{2\pi}\sum_{p\in\mathscr{S}} F_p \in \mathbb{Z}. \tag{1.90}$$

- In the continuum case, the Berry phase, the Berry curvature, and the Chern number are defined using the Berry connection $\mathbf{A}$:

$$\mathbf{A}(\mathbf{R}) = i\langle\Psi(\mathbf{R})|\,\nabla_{\mathbf{R}}\Psi(\mathbf{R})\rangle = -\mathrm{Im}\,\langle\Psi(\mathbf{R})|\,\nabla_{\mathbf{R}}\Psi(\mathbf{R})\rangle. \tag{1.91}$$

- In the continuum case, the Berry phase γ and Berry curvature B are

$$\gamma(\mathscr{C}) = -\arg\exp\left[-i\oint_{\mathscr{C}}\mathbf{A}\cdot d\mathbf{R}\right]; \qquad B = \partial_x A'_y(\mathbf{R}) - \partial_y A'_x(\mathbf{R}). \tag{1.92}$$

- In the continuum case, the Chern number is

$$Q = -\frac{1}{2\pi}\int_{\mathscr{P}} B dx dy. \tag{1.93}$$

- In the simplest case of a two-level system, the Hilbert space is mapped to the Bloch sphere, the Berry phase of a loop $\mathscr{C}$ is half the solid angle subtended by the loop,

$$\gamma_-(\mathscr{C}) = \frac{1}{2}\Omega_{\mathscr{C}}. \tag{1.94}$$

Problems

Discrete Berry phase and Bloch vectors
Take an ordered set of three arbitrary, normalized states of a two-level system. Evaluate the corresponding discrete Berry phase. Each state is represented by a vector on the Bloch sphere. Show analytically that if two of the vectors coincide, then the discrete Berry phase vanishes.

Two-level system and the Berry connection
Consider the two-level system defined in Eq. (1.63), and describe the excited energy eigenstates using the gauge $|+_{\mathbf{d}}\rangle_S$ defined in Eq. (1.67). Using this gauge, evaluate and visualize the corresponding Berry connection vector field $\vec{A}(\vec{d})$. Is it well-defined in every point of the parameter space? Complete the same tasks using the gauge $|+_{\mathbf{d}}\rangle_N$ defined in Eq. (1.68).

Massive Dirac Hamiltonian
Consider the two-dimensional *massive Dirac Hamiltonian* $\hat{H}(k_x,k_y) = m\hat{\sigma}_z + k_x\hat{\sigma}_x + k_y\hat{\sigma}_y$, where $m \in \mathbb{R}$ is a constant and the parameter space is $\mathbb{R}^2 \ni (k_x,k_y)$. (a) Take a circular loop with radius κ in the parameter space, centered around the origin. Calculate the Berry phase associated to this loop and the ground-state manifold of the Hamiltonian: $\gamma_-(m,\kappa) =$?. (b) Calculate the Berry connection $B_-(k_x,k_y)$ for the ground-state manifold. (c) Integrate the Berry connection for the whole parameter space. How does the result depend on m?

Absence of a continuous global gauge
In Sect. 1.5.1, we have shown example gauges for the two-level system that were not globally smooth on the parameter space. Prove that such globally smooth gauge does not exist.

Chern number of two-band models
Consider a two-band lattice model with the Hamiltonian $\hat{H}(\mathbf{k}) = \mathbf{d}(\mathbf{k})\cdot\hat{\sigma}$. Express the Chern number of the lower-energy band in terms of $\mathbf{d}(\mathbf{k})/|\mathbf{d}(\mathbf{k})|$.

Bibliography

[1] Yoichi Ando. Topological insulator materials. *Journal of the Physical Society of Japan*, 82(10):102001, 2013.

[2] B. Andrei Bernevig. *Topological Insulators and Topological Superconductors*. Princeton University Press, 2013.

[3] B. Andrei Bernevig, Taylor L. Hughes, and Shou-Cheng Zhang. Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells. *Science*, 314:1757, 2006.

[4] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A*, 392:45–57, 1984.

[5] Jan Carl Budich and Björn Trauzettel. From the adiabatic theorem of quantum mechanics to topological states of matter. *physica status solidi (RRL) Rapid Research Letters*, 7(1-2):109–129, 2013.

[6] Marcel Franz and Laurens Molenkamp. *Topological Insulators*, volume 6. Elsevier, 2013.

[7] Takahiro Fukui, Yasuhiro Hatsugai, and Hiroshi Suzuki. Chern Numbers in Discretized Brillouin Zone: Efficient Method of Computing (Spin) Hall Conductances. *Journal of the Physical Society of Japan*, 74(6):1674–1677, 2005.

[8] Anupam Garg. Berry phases near degeneracies: Beyond the simplest case. *American Journal of Physics*, 78(7):661–670, 2010.

[9] D.J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, 2005.

[10] M. Z. Hasan and C. L. Kane. Colloquium: Topological insulators. *Rev. Mod. Phys.*, 82:3045, 2010.

[11] Barry R. Holstein. The adiabatic theorem and Berry's phase. *American Journal of Physics*, 57(12):1079–1084, 1989.

[12] Markus König, Steffen Wiedmann, Christoph Brüne, Andreas Roth, Hartmut Buhmann, Laurens W. Molenkamp, Xiao-Liang Xi, and Shou-Cheng Zhang. Quantum Spin Hall Insulator State in HgTe Quantum Wells. *Science*, 318(6):766–770, 2007.

[13] Xiao-Liang Qi and Shou-Cheng Zhang. Topological insulators and superconductors. *Rev. Mod. Phys.*, 83:1057–1110, Oct 2011.

[14] Raffaele Resta. Berry Phase in Electronic Wavefunctions. *Troisieme Cycle de la Physique en Suisse Romande*, 1996.

[15] Shinsei Ryu, Andreas P Schnyder, Akira Furusaki, and Andreas WW Ludwig. Topological insulators and superconductors: tenfold way and dimensional hierarchy. *New Journal of Physics*, 12(6):065010, 2010.

[16] Shun-Qing Shen. Topological insulators: Dirac equation in condensed matter. *Springer Series in Solid-State Sciences*, 174, 2012.

[17] Grigory E Volovik. *The universe in a helium droplet*. Oxford University Press New York, 2009.

[18] Di Xiao, Ming-Che Chang, and Qian Niu. Berry phase effects on electronic properties. *Rev. Mod. Phys.*, 82:1959–2007, Jul 2010.