
Linking Symmetry, Crystallography, Topology, and Fundamental Physics in Crystalline Solids

Elena Derunova and Mazhar N. Ali

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.74175>

Abstract

In this chapter, we briefly introduce the evolution of symmetry as a mathematical concept applied to physical systems and lay the mathematical groundwork for discussion of topological physics. We explain how topological phases, like the Berry phase, can be obtained from a gauge symmetry of a quantum system. Also, we introduce numerical tools (e.g., Chern numbers, Wilson loops) for topological analysis of chemical solids based on the crystal structure and corresponding electronic structure.

Keywords: topological physics, topological quantum chemistry, Weyl semimetals, Dirac semimetals, Hall effects, Berry phase, Berry curvature

1. Introduction

This past century saw a dramatic advancement of our understanding of the physical world driven by the dethronement of classical physics by the combined discoveries of relativistic and quantum mechanics. From those revelations, and the subsequent intensive fundamental investigations, a new age of unprecedented rapid technological progress was ushered in. These physical theories were heavily inspired by differential geometry and linear algebra, like in the case of reinterpreting gravitation as a curvature of space or in the case of reimagining objects as both particles and waves. Today, another evolution in our understanding of physics is underway, this time inspired by the ideas of topology and symmetry. While the application of these concepts is slowly beginning to extend to all branches of science, the recent ramifications of their adaptation to crystal structures, electronic structures, and electronic properties have been profound. So much so, the 2016 Nobel Prize in Physics was awarded to Duncan Haldane, J. Michael Kosterlitz, and David J. Thouless for theoretical discoveries in topological phase

transitions and topological phases of matter. Ranging from superconductivity, superfluidity, quantized Hall effects and now to new quasiparticles, the ideas of topology and symmetry are revealing new, unexpected properties and states of matter.

The evolution of physical theories matches well with the evolution of symmetry as a mathematical concept. At first, symmetry was considered just as a transformation of space which conserves certain qualities. However, mathematicians later realized that all such transformations can form a group which can be a characterization of the quality. Crystallographic groups were born from this understanding of symmetry. Conserved, in this case, are the relative positions of atoms in space because the only allowed transformations are linear transforms (rotations and translations, i.e., Galilean transformations), which saves distances between points in space. Such an approach was enough for the dominant idea, at the time, of linear space and was consistent with Newton's classical mechanics. After Einstein's revolution, however, it turned out that distances between points are not necessarily conserved in real life. Since particles in crystals can move with velocities close to the speed of light, modern transport theory in crystals cannot ignore relativistic effects, requiring an expanded conceptualization of symmetry.

This issue was mitigated in quantum mechanics with the idea of nonhomogeneous space. The main equations there are written not for a vector in space but for a wave function, i.e., one does not have to deal with a real space of points but with a Hilbert space of possible transformations of all points in the space. Used in this way, the properties of the space itself are less important than the properties of the transformations. This transformation of space can include real numbers as well as complex numbers. Since complex numbers cannot be measured and observed, physicists consider the square of the wave function at some point as a probability to detect a particle at that point. Since the idea of a fixed position in space is not valid anymore, a new understanding of symmetry is required. Previously, symmetry transformations affected points in space; however, in quantum mechanics, the transforming object is a function, and symmetry operations are actually maps between functions a.k.a. an operator. In general, an operator is not required to have an expression, but for certain special functions, an action of the operator can be expressed as simply as, for example, a multiplication by a number. This number is called an eigenvalue, and this function is called eigenfunction (also often referred to as eigenvector or eigenstate). Both are characteristic of the operator. In the case where the eigenvalue is one (or is a strictly unitary operator), the operator will, of course, not change the eigenfunction. So if the wave function is an eigenfunction of the corresponding unitary operator of a transformation, the wave function can be considered to have a symmetry based on the transformation. In practice, the determination of eigenvalues is not typically such a trivial task, especially when the operator does not have an expression. However for linearly bounded operators in a Hilbert space, there is always a representation via the scalar product. Due to the Riesz representation theorem, any linearly bounded operator can be represented as a scalar product with another function. Note that here the scalar product is not the same as the usual product of numbers. For quantum mechanical operators, it can be written using integral notation, which is part of why physicists consider these kinds of operators as observable. Another important note to remember is that a wave function in quantum mechanics is also a map to complex space. As mentioned earlier, a symmetry operator's eigenvalue should be 1,

but the complex plane has two ways to achieve this: a unit in the real part which is, of course, 1 and also a unit in the imaginary part which is i . Thus the symmetry operator can be unitary or anti-unitary, respectively. Since the wave function is the solution of Schrodinger equation, the symmetry operator must also commute with the Hamiltonian of the system (this ensures that the operator acting on the wave function returns an eigenfunction of the Hamiltonian). In this way, a symmetry group of the wave function can be generalized to a group of operators, which have eigenvalues with an absolute value equal to one. This group is called a gauge group, and this symmetry causes topological phases of the wave function, which will be explained in this chapter.

2. Preliminaries

In order to use a consistent description, we first formulate basic mathematical definitions. The set $G = \{g_1, g_2, \dots, g_n\}$ with the operation denoted by “ \cdot ” is called a “group” if the following conditions hold:

- Every element $g = g_i \cdot g_j$ belongs to G .
- There exist unique unitary element $e \in G$ such that $e \cdot g = g \cdot e = g, \forall g \in G$.
- For every element $g \in G$, there exist inverse element g^{-1} such that $g^{-1} \cdot g = g \cdot g^{-1} = e$.
- We can combine elements of G in pair in any order $(g_i \cdot g_j) \cdot g_k = g_i \cdot (g_j \cdot g_k)$.

If in addition, if the order of the operation does not matter, i.e., $g_i \cdot g_j = g_j \cdot g_i$, the group is called commutative or abelian, otherwise it is called non-commutative or non-abelian.

If we have the commutative group V with the operation “+”, we also can multiply every element $x \in V$ by a number, either real or complex, in the following way:

- $1x = x, \forall x \in V$,
- $(a + b)x = ax + bx, \forall a, b - \text{numbers}, \forall x \in V$,
- $a(x + y) = ax + ay, \forall a - \text{number}, \forall x, y \in V$.

In this case, V is called a vector space (real or complex, respectively), and any element $x \in V$ is called a vector. The mapping $A : V \rightarrow V$ which sets the relationship between elements of the space V is called an operator. If the operator A satisfies the following properties,

$$A(cx) = cA(x), \quad A(x + y) = A(x) + A(y), \quad \forall c - \text{number}, \quad \forall x, y \in V, \quad (1)$$

then it is called a linear operator, or linear transformation, of the space V . The linear transformation between two different vector spaces V_1, V_2 is defined in the same way. A set of invertible linear transformations form a group with operation $A_1 \circ A_2$, which is the composition of the operators, and this group is called a general linear group on V and denoted by $GL(V)$.

If a physical system is described by some vector v , then a map defining the relationship between a real number and that vector is called a functional. Note that since we can only add vectors and multiply them by a number, the functionals $f : V \rightarrow \mathbb{R}$ which are useful for physical applications are the linear transformations, called linear functionals. The space of all such linear functionals on V is called a dual space and is denoted by V^* . Note that without the multiplication of two vectors in V , we cannot define analogues of polynomial functions on V . To obtain nonlinear functionals on V , we will define the multiplication of two vectors that give a number as a result. This makes functionals acting on V similar to the functions acting on the space of real numbers \mathbb{R} . The multiplication denoted by " $\langle x, y \rangle$ " is called the scalar product and satisfies the following conditions (assuming that V is a complex vector space and \bar{x} is the complex conjugation):

- $\langle x, y \rangle = \overline{\langle y, x \rangle}, \forall x, y \in V,$
- $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0 \leftrightarrow x = 0, \forall x \in V,$
- $\langle a(x + y), z \rangle = a \langle x, z \rangle + a \langle y, z \rangle, \forall a \in \mathbb{C}, \forall x, y, z \in V.$

The vector space with the scalar product is called a Hilbert space. The two vectors x and y are called orthogonal if $\langle x, y \rangle = 0$. The $\sqrt{\langle x, x \rangle} \equiv \|x\|$ is called the norm of the vector x , and it returns magnitude of the vector x , like the length of a vector in real space. Usually Hilbert spaces consist of functions; therefore, elements of Hilbert spaces are denoted simply by Greek letters.

The maximal set of vectors $\{\varphi_i\}_{\text{all } i}$ in a Hilbert space such that

$$\sum_i a_i \langle \varphi_i, \varphi_j \rangle = 0 \leftrightarrow a_i = 0, \quad \forall i, \quad (2)$$

is called a basis. If the following conditions also hold:

$$\|\varphi_i\| = 1, \quad \langle \varphi_i, \varphi_j \rangle = \delta_{ij}, \quad \forall i, j, \quad (3)$$

then it is called an orthonormal basis. In this case, any vector $\psi \in V$ can be decomposed into the sum

$$\psi = \sum_i a_i \cdot \varphi_i, \quad (4)$$

where a_i are numbers. If the number of basis vectors is finite, e.g., n , then ψ can be written just as a vector (a_1, a_2, \dots, a_n) . Note that space V can have different bases and ψ can be represented as vectors in different ways. The operator on V in this case can be written just as an $n \times n$ matrix.

In general, the symmetry of the physical system described by the vectors from V should make a group G composed of operators on V which are not necessarily linear. Of course it is easier to deal with linear operators; therefore, we introduce the concept of a representation of the group. The representation of the group $G = \{g_1, g_2, \dots, g_n\}$ with the operation " \cdot " on the vector space V is the mapping $p : G \rightarrow GL(V)$, which preserves the group operation " \cdot " in the following way:

$$p(g_1 \cdot g_2) = p(g_1) \circ p(g_2). \quad (5)$$

For example, the spatial symmetries of a crystal belong to a subgroup of group $GL(\mathbb{R}^3)$, but the symmetries of quantum objects are usually represented as a subgroup of $GL(\mathbb{C}^n)$. The relationships between those representations give rise to many interesting properties of crystals.

3. Spatial symmetries

Assume we have a point $x \in \mathbb{R}^3$ and we have a group $G_x = \{g_1, g_2, \dots, g_m\}$, where g_i is a linear transformation of space that leaves x fixed. The group G_x is called a point group. Now if we act on the point x by translating by a lattice vector R :

$$T_R(x) = x + R, \quad x \in \mathbb{R}^3, \quad (6)$$

we obtain the set of points:

$$O_{T_R}(x) = \{y \in \mathbb{R}^3 : y = gx, \quad \forall g \in T_R\}, \quad (7)$$

which is called the orbit of the action of the group of translations T_R on the element x .

If then we act by every element g_i on the point $y \in O_{T_R}$, we obtain a crystal lattice, i.e., the set of points in space that remain unchanged under the action of the group $G = G_x \times T_R$. In this case, G is called a symmorphic space group. It means the quotient space

$$R \backslash T_R = \{x \in \mathbb{R}^3 : y = gx, \forall g \in T_R, \forall y \in \mathbb{R}^3\} \quad (8)$$

has a point, x , with site group symmetry that is isomorphic to the original point group G_x [1].

Otherwise, if the lattice is invariant under the action of the group of linear transformations of a space that cannot be decomposed into $G = G_x \times T_R$ at least for one point x inside the unit cell, G is called nonsymmorphic. In this case, some operations of the group G are not separable into a combination of rotation and translation by lattice vectors, i.e., they should be complex operations like glide or skew operations. Examples of nonsymmorphic symmetry are shown in **Figure 1**. The converse, in general, is not true, because some particular combinations of glides or screws can leave one point inside unit cell fixed.

Any linear transformation L of \mathbb{R}^3 can be expressed in the following form:

$$L = Ax + b, \forall x \in \mathbb{R}^3, \quad (9)$$

where A is 3×3 matrix and b is a vector in \mathbb{R}^3 . All 3×3 matrices with a nonzero determinant and the multiplication operation form a group $GL(\mathbb{R}^3)$. The matrix A should have a determinant of 1 or -1 , where 1 corresponds to proper rotations and -1 corresponds to improper

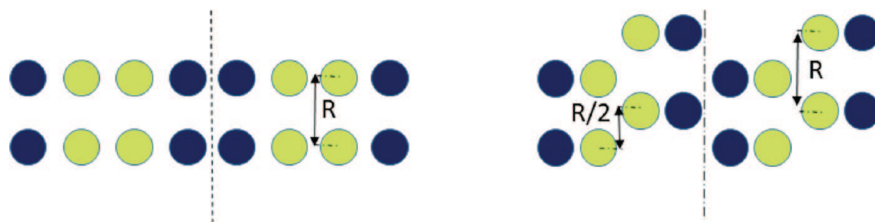


Figure 1. Symmorphic and nonsymmorphic symmetry.

rotations or reflections. All such matrices form a subgroup of $GL(\mathbb{R}^3)$, which is called an orthogonal group and is denoted by $O(\mathbb{R}^3)$. Thus, the point group G_x is a subgroup of the group $O(\mathbb{R}^3)$ [2].

4. Quantum observables

Now consider a function $\psi(x) : \mathbb{R}^3 \rightarrow \mathbb{C}$ that sets a correspondence between every point \mathbb{R}^3 and point \mathbb{C} . This function is called a wave function and represents a state of quantum system, like an electronic state in a crystal.

We can also define the sum and product of such functions and use multiplication by a number to represent interaction of the particles with each other or external forces:

$$(\psi + \phi)(x) = \psi(x) + \phi(x); \quad (\alpha\psi)(x) = \alpha \cdot \psi(x), \forall x \in \mathbb{R}^3, \alpha \in \mathbb{C} \quad (10)$$

$$(\psi\phi)(x) \equiv \langle \psi(x), \phi(x) \rangle = \int_{\mathbb{R}^3} \psi(x) \overline{\phi(x)} dx. \quad (11)$$

The space of all such functions and operations is called $L_2(\mathbb{R}^3, \mathbb{C})$. The product is a scalar product, and thus, $L_2(\mathbb{R}^3, \mathbb{C})$ is a Hilbert space.

When an observation of a physical state is carried out, it sets a correspondence between the wave function and a real number, i.e., the observation is a linear bounded functional $f : L_2(\mathbb{R}^3, \mathbb{C}) \rightarrow \mathbb{R}$. Thus, any observable property of particles should be an operator $A : L_2^*(\mathbb{R}^3, \mathbb{C}) \rightarrow L_2^*(\mathbb{R}^3, \mathbb{C})$. Luckily, Hilbert space is self-dual, i.e., $L_2^*(\mathbb{R}^3, \mathbb{C}) = L_2(\mathbb{R}^3, \mathbb{C})$, and thus, A is acting on $L_2(\mathbb{R}^3, \mathbb{C})$. In this case, due to the Riesz representation theorem [3], A can be represented as scalar product $A(\psi(x)) = \langle a(x), \psi(x) \rangle$. Thus any element of $L_2(\mathbb{R}^3, \mathbb{C})$ can be considered as a wave function and also as a functional; to distinguish this, the so-called “bra-ket” language is used—the wave functions are called “ket” vectors and denoted by $|\psi(x)\rangle \in L_2(\mathbb{R}^3, \mathbb{C})$, while the functionals are called “bra” vectors and denoted by $\langle \psi(x)| \in L_2^*(\mathbb{R}^3, \mathbb{C})$ —the scalar in this case is denoted by $\langle \psi(x)|\psi(x)\rangle$, and the action of an operator A is denoted by $\langle \psi(x)|A|\psi(x)\rangle$. Schematically the relationship between these spaces is shown in **Figure 2**.

For some functions the actions of the operator can be written as multiplication by the number

$$A(\varphi) = \lambda\varphi, \quad \lambda \in \mathbb{C}. \tag{12}$$

The number λ is called the eigenvalue, and the function $\varphi(x)$ is called the eigenfunction or eigenvector. All operators considered in quantum mechanics are assumed to be Hermitian, meaning the operator has real eigenvalues and the set of its eigenfunctions $\{\varphi_i\}_i$ form an orthogonal basis [4]. Assume the operator has finite number of eigenvalues; in this case, the action of the operator on the function from the subspace spanned by eigenfunctions is expressed by the following matrix:

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \tag{13}$$

One of the most important operators in quantum mechanics is the momentum operator $i\hbar \frac{d}{dx}$ which is particularly used when analyzing a material's electronic structure or electronic energy vs. momentum map. Its eigenfunctions $\varphi_n(x)$ are called eigenstates and denoted by $|n\rangle$. Thus, we can decompose the wave function as $\psi(x) = \sum_n c_n \varphi_n(x) = \sum_n c_n |n\rangle$. The eigenvalues form a matrix which is called the Hamiltonian of the quantum system, and they correspond to the measured energy of the system.

If we also add the normalization condition $\|\psi(x)\| = 1$, we can consider $\psi(x)$ as a probabilistic measure. That means the functional

$$\langle \psi(x) | x | \psi(x) \rangle = \int_{\mathbb{R}^3} \psi(x) x \overline{\psi(x)} dx \tag{14}$$

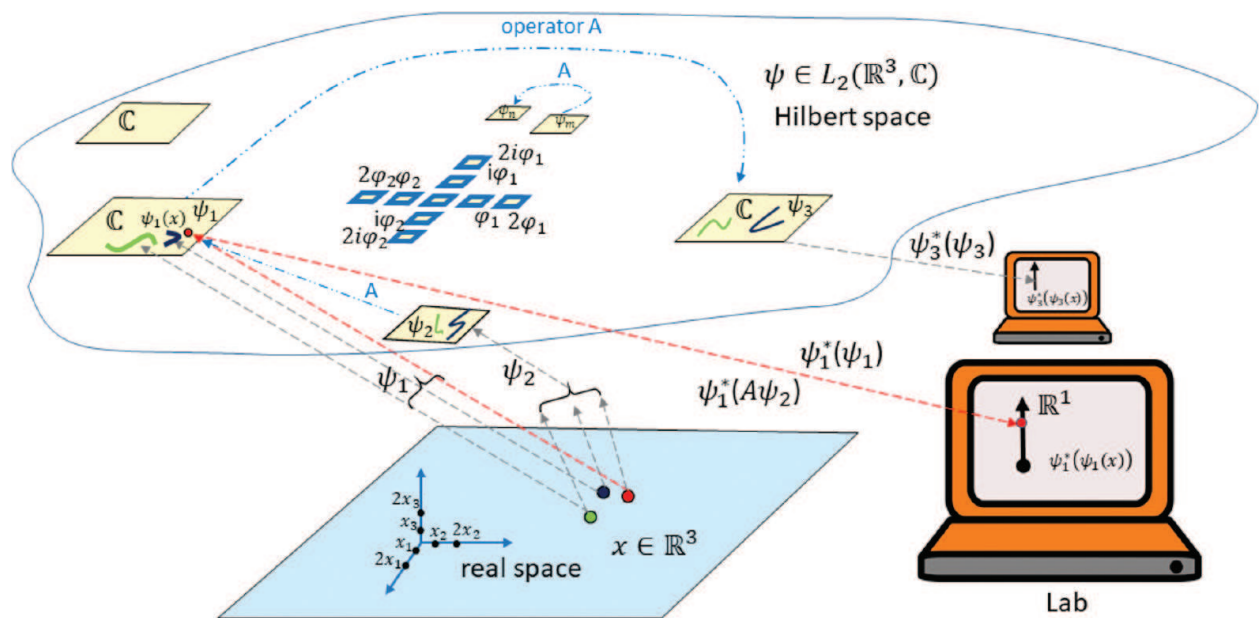


Figure 2. Different spaces used in quantum mechanics.

defines a probability to observe the particle in the position x . The $|\psi|^2$ in this case defines a probability density [4]. As we can see from **Figure 3**, different symmetries between real and imaginary parts of the wave function define different types of symmetries of the probability density.

Consider now all possible transformations of the function which preserve it as an eigenfunction of the Hamiltonian with the same eigenvalue. For the eigenfunction φ_n , the Hamiltonian acts just as multiplication by the function $\lambda_n \varphi_n$, i.e.,

$$H\varphi_n = \langle \lambda_n \varphi_n, \varphi_n \rangle = \lambda_n \langle \varphi_n, \varphi_n \rangle = \lambda_n \|\varphi_n\|^2 \quad (15)$$

Thus such a transformation should not change $\|\varphi_n\|^2$; if one wave function can be obtained from another via such a transformation, those wave functions are not distinguishable through observation. This group is called a gauge group and represents the symmetry group of the wave function [4]. Since

$$\forall a \in \mathbb{C}, \quad \|a\varphi_n(x)\|^2 = a\bar{a} \cdot \|\varphi_n(x)\|^2 = |a|^2 \cdot \|\varphi_n(x)\|^2, \quad (16)$$

multiplication by complex numbers with $|a| = 1$ forms the group of such transformations. These numbers lie on a unit circle in the complex plane, and the group of multiplications by such numbers is called $U(1)$ or the group of unitary transformations of the complex plane. The complex number $a = |a|(\cos(\alpha) + i \cdot \sin(\alpha))$ can be represented as an exponential function in the following way:

$$a = |a|e^{i\alpha} \quad (17)$$

$$|a| = 1 \rightarrow a = e^{i\alpha}. \quad (18)$$

Thus the action of the $U(1)$ gauge is just a multiplication by the function $e^{i\alpha}$. If we represent the complex plane as a stereographic projection of Riemann sphere, we can illustrate $U(1)$ action as rotation of the sphere. Schematically, it is shown in **Figure 4**.

If the wave function corresponds to a fermion, according to the Pauli principle, only two fermions with opposite sign spins can occupy the same energy state. So it is convenient to

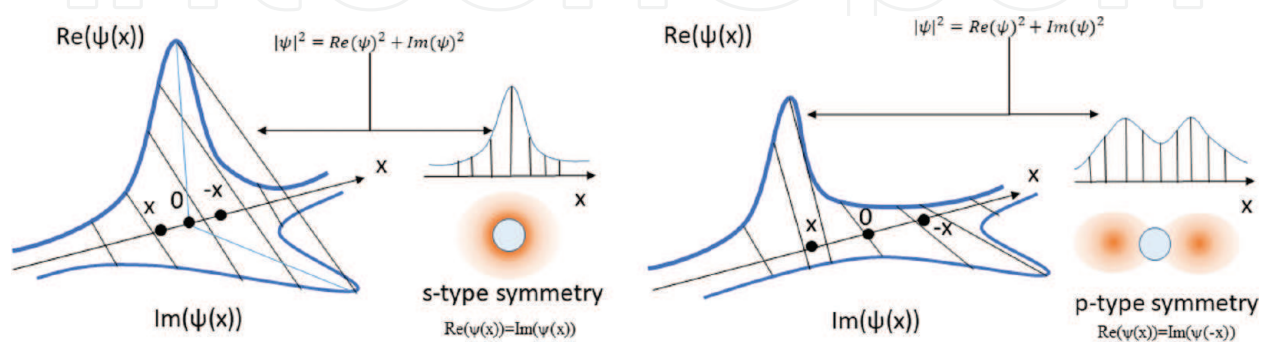


Figure 3. Symmetry of the wave function.

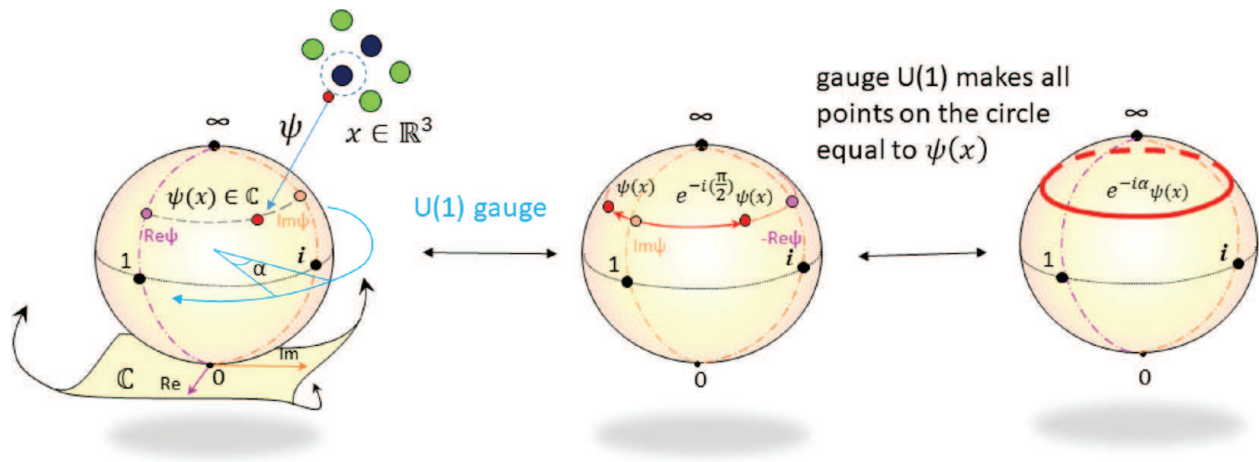


Figure 4. Action of the U (1) gauge on complex space.

consider $\psi(x)$ as having two components $\psi = (\psi_{up}, \psi_{down})$, i.e., ψ is acting to \mathbb{C}^2 . In this case the transformation of the vector (ψ_{up}, ψ_{down}) in the two-dimensional complex space \mathbb{C}^2 is described by a 2×2 complex matrix A . For the same reasons as above, this matrix should be unitary, i.e., $AA^\dagger = I$ (where A^\dagger is a Hermitian conjugated matrix). All such matrices form a group $U(2)$. The eigenvalues of such matrices lie on the unit circle that implies any matrix A from $U(2)$ can be represented in the following form:

$$A = \begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & e^{i\alpha_2} \end{pmatrix} \quad (19)$$

$$A\psi = \begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & e^{i\alpha_2} \end{pmatrix} \begin{pmatrix} \psi_{up} \\ \psi_{down} \end{pmatrix} = \begin{pmatrix} e^{i\alpha_1}\psi_{up} \\ e^{i\alpha_2}\psi_{down} \end{pmatrix}. \quad (20)$$

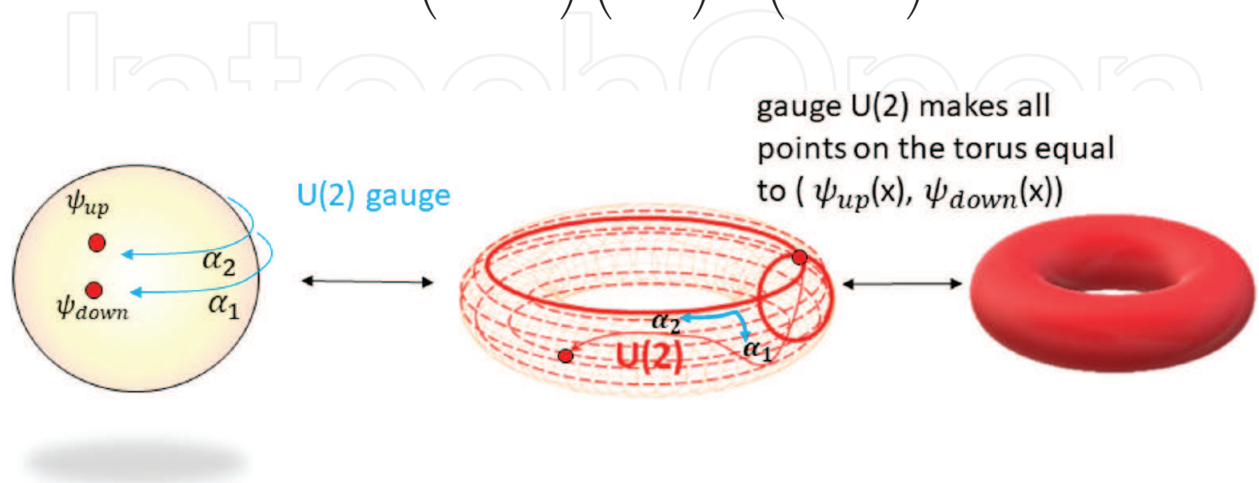


Figure 5. Action of the U (2) gauge on complex space.

Thus, if we represent (ψ_{up}, ψ_{down}) as two different points on the Riemann sphere, then the action $U(2)$ is a simultaneous rotation of the point ψ_{up} by angle α_1 and the point ψ_{down} by angle α_2 . After the full circle rotation, we arrive at the initial point making the space of parameters (α_1, α_2) a torus. This is shown schematically in **Figure 5**.

5. Geometrical phases of the Bloch states

The dynamics of the $\psi(x)$, i.e., changing $\psi(x)$ in time, is defined by the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial \psi(t, x)}{\partial t} = H\psi(t, x), \quad (21)$$

where H is the Hamiltonian and consists of all possible physical interactions (ideally) that the particle can be involved in. The wave function can be determined by solving this equation. The solution of the Eq. (21) for the eigenstates of the Hamiltonian can be written in the following form:

$$\varphi_n(t, x) = e^{-\frac{i}{\hbar} \int_0^t dt' \lambda_n(t')} \varphi_n(0, x), \quad (22)$$

For a free electron, H consists only of the kinetic energy term $-\frac{\hbar^2}{2m} \nabla^2$. Its eigenfunctions are well known as s, p, d, f, etc. (the atomic orbitals). If the electron is moving in crystal, an external periodic potential, formed by ion cores, and the average potential of all of the other electrons must be included. In this case, due to the Bloch theorem [5], the eigenfunctions of the Hamiltonian can be written as

$$\varphi_k^n(0, x) = \varphi_k^n(x) = e^{ikx} u_k^n(x), \quad u_k^n(x + R) = u_k^n(x). \quad (23)$$

The value $\hbar k$ is called the crystal momentum and is associated with an electron in the lattice. If the lattice consists of N atoms and every atom has n electrons, the full lattice Hamiltonian is the $Nn \times Nn$ matrix for the Nn electron system. Working with such high dimensional objects is not convenient. Therefore, the wave functions are categorized into bands $\{\varphi_n^k\}_{k=1..N}$ according to the local symmetry of the wave function (like shown previously in **Figure 3**), which is described by quantum numbers of the atom. After Fourier transformation, φ_n becomes a function represented in a new basis of functions which depend on crystal momentum (schematically, it is shown in **Figure 6**). This makes up the energy versus momentum band structure that is a more convenient representation of the full lattice Hamiltonian, compared with matrix notation. The band in reciprocal space, however, is not really a function; it consists of discrete points and is neither smooth nor continuous. However this band structure contains all information about the original function $\psi(x)$. Roughly speaking, the band is made up of the coordinates of the function $\varphi_n(x)$ in the basis of harmonics $\left\{ e^{i\frac{2\pi n}{l}x} \right\}_{n=1..N}$ (eigenfunctions of the translation operator [6]) where N is the number of unit cells and l is the lattice parameter.

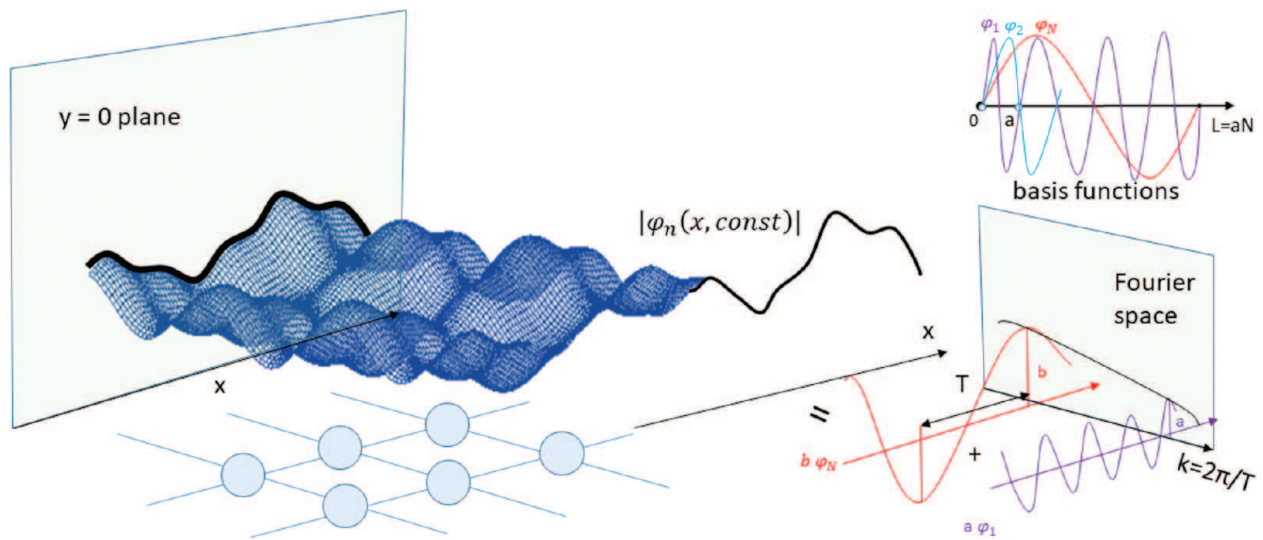


Figure 6. Representation of a function in the Fourier space.

According to the quantum adiabatic theorem during the time evolution, the system remains in the eigenstates $\varphi_n(0, x) = |n(0)\rangle$ up to phase factor or in other words $|n(0)\rangle = e^{-i\alpha_n}|n(t)\rangle$. If we consider k as a parameter changing in time then

$$\frac{d\varphi(k(t), x)}{dt} = e^{ikx} \left(i x u_k(x) \frac{dk}{dt} + \partial_k u_k(x) \frac{dk}{dt} \right) \quad (24)$$

that gives rise to an additional phase factor to the solution of Schrodinger equation [7]:

$$\varphi_n(t, x) = e^{i\gamma_n} e^{-\frac{i}{\hbar} \int_0^t dt' \lambda_n(k(t'))} |n(k(t))\rangle, \quad (25)$$

$$\gamma_n = i \int_{\text{path } C} dk \langle n(k) | \frac{\partial}{\partial k} | n(k) \rangle. \quad (26)$$

If the path C is closed, then $\gamma_n(t)$ is called the Berry phase. The expression

$$A_n(k) = i \langle n(k) | \frac{\partial}{\partial k} | n(k) \rangle \quad (27)$$

is called the Berry connection, and it is the vector field over all reciprocal space. We also can define a curl of this vector field which is called the Berry curvature:

$$\Omega_{i,j}^n(k) = \frac{\partial}{\partial k_i} A_j^n(k) - \frac{\partial}{\partial k_j} A_i^n(k) \quad (28)$$

Origin of the Berry curvature in the reciprocal space is schematically demonstrated in Figure 7.

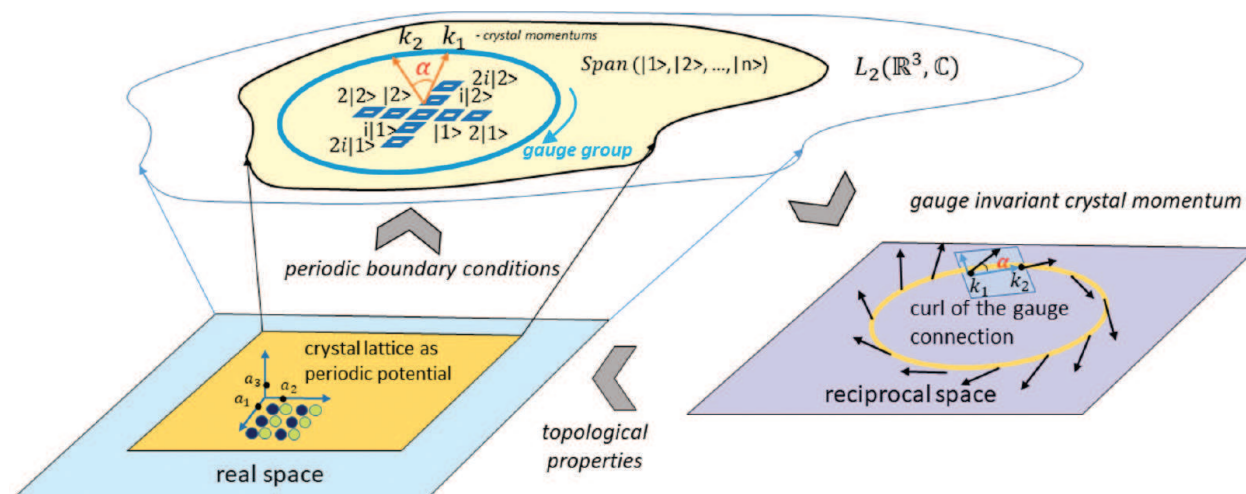


Figure 7. Berry curvature in the reciprocal space.

The Berry curvature is involved in the semiclassical equation of motion of the particle [8]:

$$\frac{dx_i}{dt} = \frac{\partial \lambda_n(k)}{\hbar \cdot \partial k^i} - \frac{dk_j}{dt} \cdot \Omega_{i,j}^n(k), \quad (29)$$

which makes the Berry curvature fundamental to various Hall effects, i.e., quantum (integer and fractional) Hall effects (QHE), the quantum anomalous Hall effect (QAHE), and the quantum spin Hall effect (QSHE) [9–11]. The QAHE is where an anomalously large current is generated orthogonal to the applied electric field without application of an external magnetic field. The QSHE is where a spin current is generated orthogonal to the applied electric field also without application of an external magnetic field.

Notice that if we change direction of time $t = -t$, we change the route from counterclockwise to clockwise in the path integrals. If time reversal symmetry is broken and the clockwise integral is *not* equal to the counterclockwise integral, it requires the Berry connection to have a nonzero curl, i.e., nonzero Berry curvature.

If time reversal symmetry is not broken, the Berry curvature still can be nonzero due to spacial symmetries. In this case, analysis can be done using topological indices resulting from the band structure.

6. Gauge action on the reciprocal space

First consider a one-dimensional case. As it was shown before, the U(1) gauge action can be represented as multiplication by the factor $e^{i\alpha}$. For an electron in a crystal, it has the following form:

$$e^{i\alpha} \varphi(x) = e^{i\alpha} e^{ikx} u_k(x) = e^{i(kx+\alpha)} u_k(x). \quad (30)$$

Thus the action of the gauge can be considered as an additional phase factor or shift by α in reciprocal space. So the gauge allows the change of parameter k in time, and all geometrical phases described above can be considered a result of the gauge symmetry. Schematically this is shown in **Figure 8**.

The gauge symmetry is the conservation of the eigenvalues and eigenstates of the momentum operator. The eigenstates of the momentum operator in a crystal are assumed to also be eigenstates of the operator of translation by a lattice vector. Orthogonality of the eigenstates implies that bands in the band structure should not intersect, i.e., not have identical E and k values. If two bands intersect that means that the corresponding eigenstates $|n\rangle, |n+1\rangle$ are not orthogonal and the corresponding matrix representing the action of the Hamiltonian has off-diagonal terms. This is contradictory to the Hermitian rules of the Hamiltonian, i.e., its eigenstates should be orthogonal. This can happen when the eigenstates of the translation operator are not a suitable set of functions to form a basis. But how can this occur?

The Hamiltonian for electrons in the crystal lattice is usually constructed from a tight-binding model and the linear combination of the atomic orbitals (LCAO) [12]. Since the crystal Hamiltonian has the following form:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + (\text{terms representing interactions}), \quad (31)$$

it is reasonable to assume that eigenstates of the electrons in the lattice do not differ strongly from the eigenstates of the free electron Hamiltonian $-\frac{\hbar^2}{2m}\nabla^2$ which are s, p, d, f, etc. orbitals. Thus we can represent the full lattice Hamiltonian as linear combinations of atomic orbitals according to the space group symmetry, which is the basic logic of *molecular* orbital theory. If the space group is nonsymmorphic, then there is also translation by a fraction of a lattice vector that generates additional eigenstates that, in combination with other symmetries, give rise to degeneracies in the band structure.

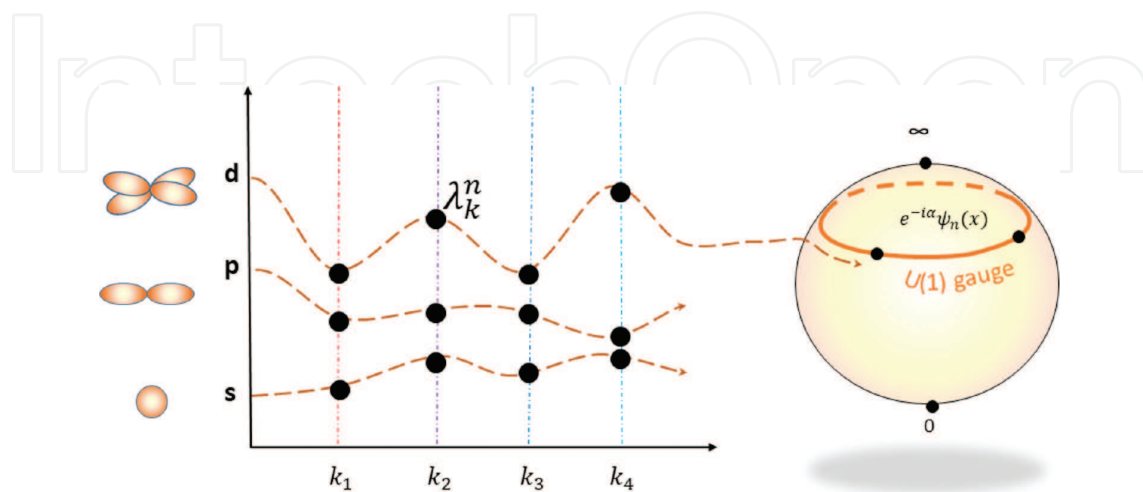


Figure 8. Gauge connection on the band structure.

To understand this we can introduce Wannier functions as bases for representation of the eigenstates of the Hamiltonian, instead of eigenstates of a translation operator, as basis. The Wannier functions can be obtained from the Bloch eigenstates in the following way [13]:

$$\phi_R^n(x) = \frac{V}{2\pi^3} \int_{BZ} dk e^{-ikR} \psi_k^n(x), \quad (32)$$

where V is the real space primitive cell volume. The Wannier functions essentially let one transform the band structure back from reciprocal space to real space, allowing relatively easy application of symmetry and the calculation of real space properties like the quantum spin Hall effect.

Now let us consider higher dimensions. The 1D band structure is the cross section of a higher-dimensional picture. In this case the anti-crossing point can remain a point in higher dimensions but may also be a line or even surface in a 3D space as illustrated below. While for the 2D case, we can still plot a 2D band structure and distinguish points from a line just visually, in 3D it becomes quite complicated. The typical procedure is to project full 3D band structure on various surfaces and analyze the series of projections. Schematically, this is shown in **Figure 9**. A degenerate point in the band structure, which remains a point on the Fermi surface, is called Dirac point or in spinful case Weyl point. An example of the calculated Dirac point of $HgTe$ [14] and the measured Dirac point (via ARPES) of Sb_2Te_3 [15] is shown in **Figure 10**.

There are ways to avoid degeneracy, however. If we consider the spinful case with $U(2)$ gauge and include a spin orbit coupling (SOC) term in the Hamiltonian, the two degenerate states become one connected state in terms of the two-component wave function. In this case, the Dirac point becomes a source of Berry curvature and thus gives nontrivial spin-dependent transport properties like the anomalous and spin Hall effect.

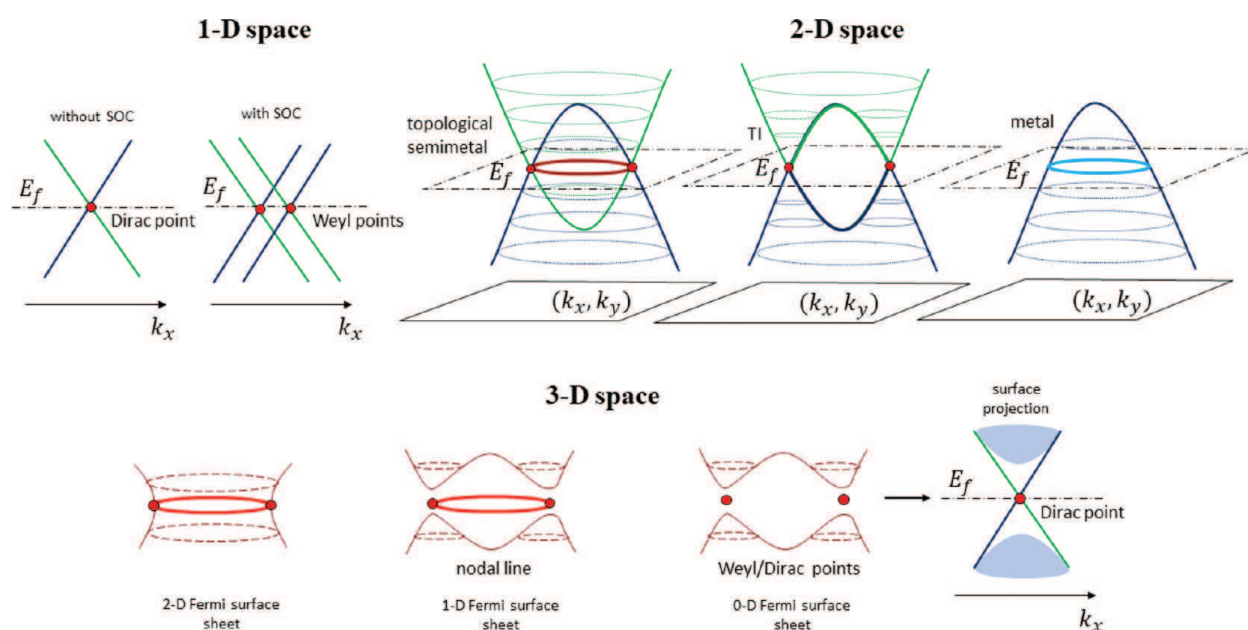


Figure 9. Degenerate manifolds in the band structure.

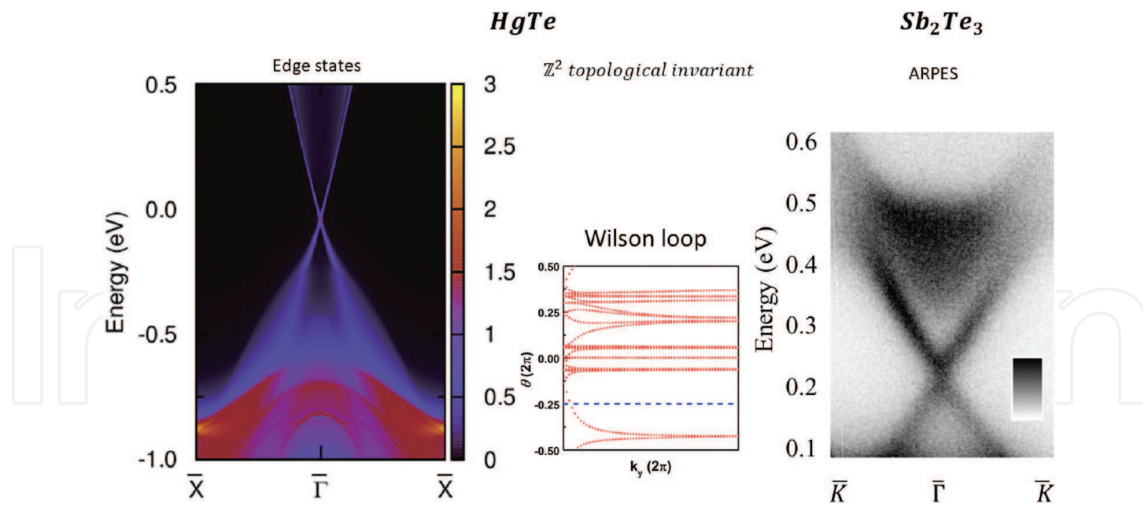


Figure 10. Numerical and experimental observation of Dirac cone [16].

Another way to do anti-crossing analysis is to use topological indices. This numerical method is the basis for the algorithmic analysis of space group symmetry and their possibilities of yielding varying topologically nontrivial band structures.

7. Topological indices

It turns out that existence of Hall conductance can be checked directly from the band structure and that it is an internal property of the bands. The indicator that physicists use to identify this topological property of the band structure is called the Chern number, which is the integral of the Berry curvature of the band over the entire Brillouin zone:

$$C_{ij}^n = \frac{1}{2\pi} \int_{BZ} \Omega_{i,j}^n(k) d_{k_i} d_{k_j} \quad (33)$$

In this case, the Hall conductance of the n^{th} band is proportional to the Chern number, quantized in units of $\frac{e^2}{h}$ and can be calculated by the formula:

$$\sigma_{ij}^n = \frac{e^2}{h} \cdot C_{ij}^n \quad (34)$$

The Chern number is a very powerful tool; it can be used not only for calculation of the Hall conductance but also to indicate a surface state. For example, the Chern number described above is the “first” Chern number, and this nonzero number indicates surface conductance for a 2D bulk insulator. For a 3D bulk insulator, higher Chern numbers can be used to indicate surface states [16].

Another way to obtain a topological index is using the Wilson loop [17]:

$$W(l) = e^{\int_l A(k) dl} \tag{35}$$

where l is a loop in k -space and $A_{ij}(k) = \langle u_{i,k}, \nabla_k u_{j,k} \rangle$ is a Berry-Wilczek-Zee connection. Note for this connection we need at least a two-band system, like (ψ_{up}, ψ_{down}) . The Wilson loop describes a parallel transport of the gauge field along the closed loop.

Mathematically, a path between two points k_1, k_2 in k -space can be parametrized by an argument t in the following way $k_t = tk_1 + (1-t)k_2, t \in [0, 1]$ —when loop $k_1 = k_2$. The Wilson loop shows how the gauge varies with crystal momentum along a closed path in k -space; the final gauge phase should be the same as initial. For example, the parameter space of the $U(1)$ gauge is a circle; thus, moving along a loop in k -space the gauge phase can be either unchanged or equal to an integer number of full circles ($2\pi n$). We can consider also the class of equivalent loops: loops that give one circle of phase, two circles, etc. These classes of equivalent loops form a group, called a fundamental group [18]. As it was shown before, the parameter space of the $U(2)$ gauge is a torus. The torus has two types of loops: one which shrinks into a point and one which does not. This is known as fundamental group of the torus. The Wilson loop distinguishes those cases and yields a \mathbb{Z}^2 topological classification. In **Figure 10**, an example of a calculated Wilson loop for the Se_2Te_3 is shown. The connection between red dashed lines indicated an index of 1 in the \mathbb{Z}^2 topological classification.

The gauge transformation is a transformation that preserves the eigenfunction, and the crystal eigenfunction is represented as a band in electronic structures. In the LCAO approach, bands

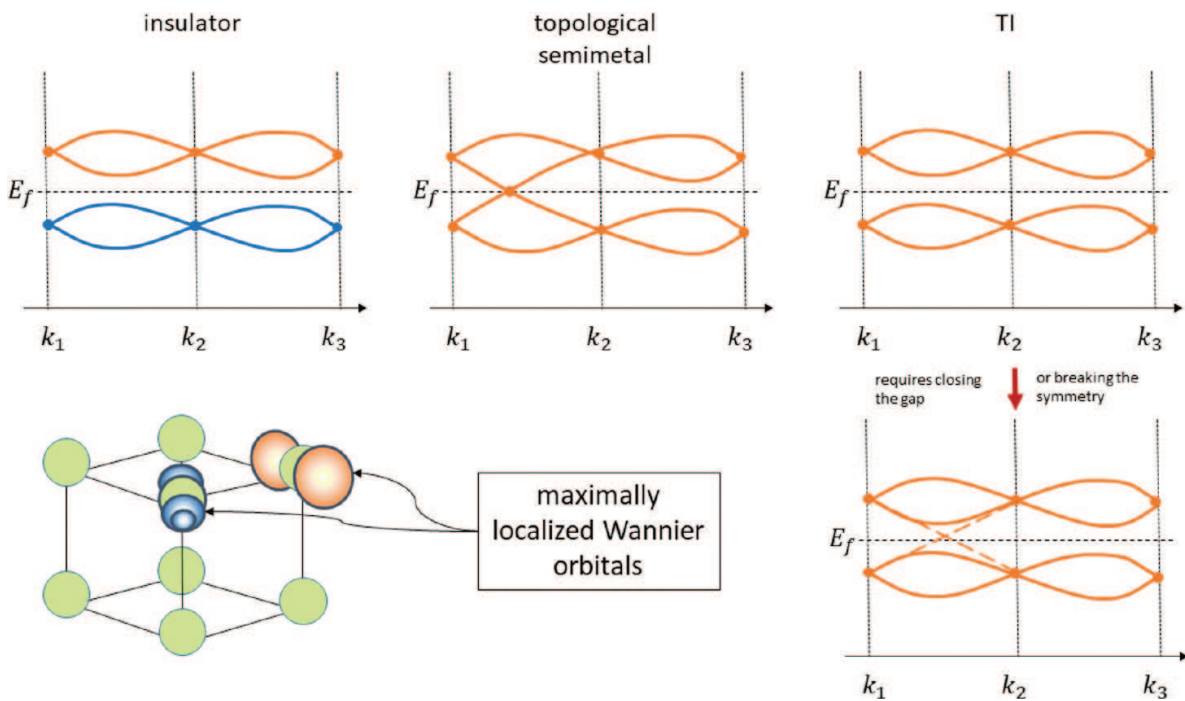


Figure 11. Connectivity of bands.

in the electronic structure can be considered also as representations of the spacial symmetry of corresponding orbitals and space group symmetry operations. Recently, a monumental and soon-to-be defining work of this field was carried out where topological analysis and classification were done for all 230 crystallographic space groups that describe all possible arrangements of atoms in space [19]. In their work, Bradlyn et al. use the fact that bands can form a connected group of bands in the band structure corresponding to Wannier functions centered at maximal Wyckoff positions. If the Fermi level occurs inside such a set of bands, the compound should be topological. If the real compound's set of bands is connected but filled by only a fraction of the number of electrons required to fully occupy the set of bands, the compound is a symmetry enforced semimetal. If such set of bands should be connected but in the real compound's band structure, the set splits into a gapped state with the Fermi level inside the gap, the compound must be a topological insulator, and the bands become connected though surface states. Schematically, this is shown in **Figure 11**.

8. Conclusion

We began this chapter by asserting that evolutions in our understanding of the physical universe have been driven by the reimagining of physical theories with different mathematical concepts and that we are currently undergoing another such evolution inspired by ideas from topology and symmetry. Starting from basic mathematical and physical concepts, we constructed the ideas of eigenfunctions, eigenstates, gauge transformations, and how symmetries affect/define them. We then showed how electrons (and their wave functions) in crystals can be understood in this manner. Finally, the power of this type of understanding was illustrated by the classification of topological phases of matter with Bradlyn's effort being the pinnacle of the body of work in the field over the last decade. Such a general work outlining the possibilities for topological phases as a matter of symmetry group will drastically change how chemists and physicists search for new topological materials; they will now be able to definitively start from a set of known possible topological outcomes, given a space group, and adjust the Fermi level using chemical/physical control to realize the type of topological state desired.

The time is arriving for topological physics to reach technological application. To this end, researchers are attempting to take advantage of the intrinsic quantum anomalous and quantum spin Hall effects (QSHE) both of which have their basis in Berry curvature which, as described earlier, can be understood from a symmetry perspective. The QSHE in particular has immediate applications to the field of spintronics, which requires large spin currents for switching the states of devices. Since anti-crossings can be sources of Berry curvature and since large Berry curvature can result in a large spin Hall effect, it follows that topological materials (which commonly have demanded or gapped anti-crossings) will be ideal candidates for spin Hall materials.

In the last 10 years, materials scientists have realized topological insulators, Dirac semimetals, Weyl semimetals, Dirac/Weyl line nodes, and compounds with three, six, and eightfold degenerate fermions Na_3Bi , Cd_3As_2 [20], $ZrSiS$ [21], WTe_2 [22], Bi_2Te_3 [23], $LuPtBi$, $YPtBi$ [24], and Ta_3Sb [25]. However almost all of this work has been done on nonmagnetic systems; the inclusion of magnetism is difficult in current density function theory calculations. For accurate

electronic structure calculations, magnetic ordering needs to be experimentally determined and verified because competing magnetically ordered states can be energetically similar. Soon, Bradlyn's type of analysis will be expanded for the 1651 magnetic space groups in three dimensions, which opens an even larger and more diverse world of possible compounds with ever more interesting properties to explore.

Author details

Elena Derunova and Mazhar N. Ali*

*Address all correspondence to: maz@berkeley.edu

Max Planck Institute for Microstructure Physics, Halle, Germany

References

- [1] Fedorov E. A. S. for X-ray, and E. Diffraction. Symmetry of Crystals. American Crystallographic Association: ACA monograph; 1971
- [2] Cotton F. Chemical Applications of Group Theory. India: Wiley; 2008
- [3] Reed M, Simon B. I: Functional Analysis. Methods of Modern Mathematical Physics: Elsevier Science; 1981
- [4] Bohm A, Loewe M. Quantum Mechanics: Foundations and Applications. Springer Study Edition. New York: Springer; 1993
- [5] Kittel C. Introduction to Solid State Physics. Wiley; 2004
- [6] Ashcroft N, Mermin N. Solid State Physics. HRW international editions, Holt, Rinehart and Winston, 1976
- [7] D Xiao, M-C Chang, Q Niu. Berry phase effects on electronic properties. Reviews of Modern Physics. 2010;**82**:1959 (2007, Jul 2010)
- [8] Haldane FDM. Berry curvature on the fermi surface: Anomalous hall effect as a topological fermi-liquid property. Physical Review Letters. November 2004;**93**:206602
- [9] Sinova J, Valenzuela SO, Wunderlich J, Back C, Jungwirth T. Spin Hall effects. Reviews of Modern Physics. Oct. 2015;**87**:1213-1260
- [10] Sinova J, Culcer D, Niu Q, Sinitsyn NA, Jungwirth T, MacDonald AH. Universal intrinsic spin hall effect. Physical Review Letters. March 2004;**92**:126603
- [11] Gradhand M, Fedorov DV, Pientka F, Zahn P, Mertig I, Györffy BL. First-principle calculations of the Berry curvature of Bloch states for charge and spin transport of electrons. Journal of Physics: Condensed Matter. May 2012;**24**:213202

- [12] Hoffmann R. *Solids and Surfaces: A Chemist's View Of Bonding In Extended Structures*. VCH Publishers; 1988
- [13] Marzari N, Mostofi AA, Yates JR, Souza I, Vanderbilt D. Maximally localized Wannier functions: Theory and applications. *Reviews of Modern Physics*. October 2012;**84**:1419-1475
- [14] Li J, He C, Meng L, Xiao H, Tang C, Wei X, Kim J, Kioussis N, Malcolm Stocks G, Zhong J. Two-dimensional topological insulators with tunable band gaps: Single-layer HgTe and HgSe. *Scientific Reports*. September 2015;**5**:14115
- [15] Johannsen JC, Autès G, Crepaldi A, Moser S, Casarin B, Cilento F, Zacchigna M, Berger H, Magrez A, Bugnon P, Avila J, Asensio MC, Parmigiani F, Yazyev OV, Grioni M. Engineering the topological surface states in the $(\text{Sb}_2)_m - \text{sb}_2\text{te}_3$ ($m = 0-3$) superlattice series. *Physical Review B*. 2015;**91**(20):201101
- [16] Bernevig B, Hughes T. *Topological Insulators And Topological Superconductors*. Princeton University Press; 2013
- [17] Alexandradinata A, Wang Z, Bernevig BA. Topological insulators from group cohomology. *Physical Review X*. 2016;**6**:021008
- [18] Tom Dieck T. *Algebraic Topology*. EMS Textbooks in Mathematics. Corrected 2nd printing edition. Zurich: European Mathematical Society; 2010. OCLC: 711865106
- [19] Bradlyn B, Elcoro L, Cano J, Vergniory MG, Wang Z, Felser C, Aroyo MI, Bernevig BA. Topological quantum chemistry. *Nature*. July 2017;**547**:298
- [20] Jenkins GS, Lane C, Barbiellini B, Sushkov AB, Carey RL, Liu F, Krizan JW, Kushwaha SK, Gibson Q, Chang T-R, Jeng H-T, Lin H, Cava RJ, Bansil A, Drew HD. Three-dimensional Dirac cone carrier dynamics in na_3Bi and cd_3as_2 . *Physical Review B*. August 2016;**94**:085121
- [21] Schoop LM, Ali MN, Straßer C, Topp A, Varykhalov A, Marchenko D, Duppel V, Parkin SSP, Lotsch BV, Ast CR. Dirac cone protected by non-symmorphic symmetry and three-dimensional Dirac line node in ZrSiS . *Nature Communications*. May 2016;**7**:11696
- [22] Ali MN, Xiong J, Flynn S, Tao J, Gibson QD, Schoop LM, Liang T, Haldolaarachchige N, Hirschberger M, Ong NP, Cava RJ. Large, nonsaturating magnetoresistance in WTe_2 . *Nature*. September 2014;**514**:205
- [23] Barua S, Rajeev KP, Gupta AK. Evidence for topological surface states in metallic single crystals of Bi_2Te_3 . *Journal of Physics: Condensed Matter*. 2015;**27**(1):015601
- [24] Yang H, Yu J, Parkin SSP, Felser C, Liu C-X, Yan B. Prediction of triple point fermions in simple half-Heusler topological insulators. *Physical Review Letters*. September 2017;**119**:136401
- [25] Bradlyn B, Cano J, Wang Z, Vergniory M, Felser C, Cava R, Bernevig BA. Beyond Dirac and Weyl fermions: Unconventional quasiparticles in conventional crystals. *Science*. 2016;**353**(6299):aaf5037

