10 Dynamical Mean-Field Theory for Correlated Topological Phases

Michael Potthoff

Department of Physics, University of Hamburg,
Notkestraße 9-11, 22607 Hamburg

The Hamburg Centre for Ultrafast Imaging,
Luruper Chaussee 149, 22761 Hamburg

Contents

E. Pavarini, E. Koch, A. Lichtenstein, and D. Vollhardt (eds.)
Dynamical Mean-Field Theory of Correlated Electrons
Modeling and Simulation Vol. 12
Forschungszentrum Jülich, 2022, ISBN 978-3-95806-619-9
http://www.cond-mat.de/events/correl22

1 Motivation

Consider two objects A and B. These could be physical objects, such as elementary particles, molecules, or macroscopically large bodies. They could be mathematical structures, but also poems or novels, the mood of two people, etc. It is obvious that the question "Is A equal to B?" rarely deserves a positive answer. In the case of "Is A essentially equal to B?" on the other hand, it seems to be a question worthy of discussion in almost all situations. All the complexity here is cleverly hidden in the word "essentially". This leaves room for us to define criteria by which we can decide about essential equality or inequality. An instructive example is a cat and a dog. Are these "objects" the same? No, but it may be reasonable to regard them as essentially the same, depending on the criterion chosen, e.g., that of being an animal.

Why do we bother? The relation "is essentially the same" is an equivalence relation because the following statements appear to be immediately true: "A is essentially the same as A", "if A is essentially the same as B, then B is essentially the same as A", and "if A is essentially the same as B and B is essentially the same as C, then A is essentially the same as C". Any equivalence relation can be used to classify objects, and classification is an important goal of almost any science. Classification of life forms, for example, is one of the most important goals of biology.

In physics, there are various concepts for classifying systems or the static or dynamic states of systems. These find expression, for example, in the periodic table of elements or in equilibrium phase diagrams of macroscopically large physical bodies. Despite the same chemical composition, we distinguish, for example, the condensed and the liquid state of water. Decision-making criteria in many cases, and also in the distinction between solid and liquid, can be based on symmetry arguments, i.e., on invariance under a group of transformations.

Let us discuss a well-known example from mathematics, which is frequently used in quantum physics: A separable Hilbert space is a linear space over the field of real or complex numbers, i.e., addition and scalar multiplication are defined and satisfy a couple of postulates. In addition the space is equipped with a bilinear or sesquilinear inner product, which defines a distance function for which it is a complete metric space, and has a countable basis. Two separable Hilbert spaces $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$ are essentially the same, $H_\mathrm{A} \cong H_\mathrm{B}$, if there exists an isomorphism $\mathcal{H}_\mathrm{A} \to \mathcal{H}_\mathrm{B}$, i.e., an invertible map preserving the mathematical structure. Due to the rich structure, the equivalence relation $\cong$ is very strong. This "explains" the simplicity of the classification, namely *all* (infinite-dimensional) separable Hilbert spaces are essentially the same and, e.g., isomorphic to L^2.

A more fruitful, finer classification is obtained with a "weaker" equivalence relation. Groups, for example, are mathematical structures defined with a minimum of postulates. Structure-conserving maps between two groups G_1 and G_2, group isomorphisms, therefore define a much weaker equivalence relation $G_1 \cong G_2$. In fact, a complete classification is an entirely open mathematical problem. Narrowing a bit and asking for essential equality between finite, simple groups, brings us to one of the major milestones of group theory. Here, a complete classification has been achieved in fact with a proof that is documented several hundred papers published over the last 70 years.

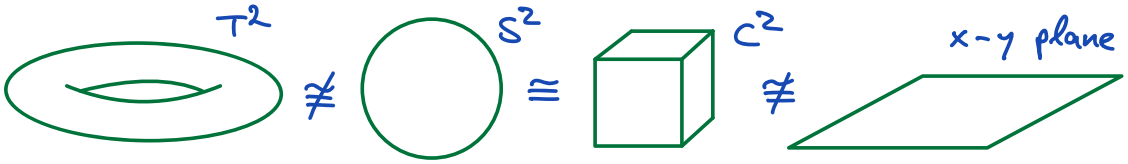

Fig. 1: *Various topological subspaces of* $\mathbb{R}^3$.

An equivalence relation that has proven very fruitful in the context of condensed-matter physics is "A can continuously and reversibly be deformed into B". Continuity is a rather weak property of functions and the main concept of topology. To give a mathematical example again, we may consider $\mathbb{R}^3$ as a topological space and forget about its capability to carry a linear structure. Among the topological subspaces of $\mathbb{R}^3$ we have, e.g., the 2-sphere $S^2 = \{x \in \mathbb{R}^3 \mid ||x|| = 1\}$, the Cartesian product of 1-spheres $S^1 \times S^1$, the surface of a cube C^2, the 2-torus ("donut") T^2, the open or closed 3-ball with the 2-sphere as its boundary, and the like, see Fig. 1. We find that $S^1 \times S^1 \cong T^2 \ncong S^2 \cong C^2 \ncong x$-$y$ plane etc.

An important tool of topological classification are *topological invariants*. These are all quantities that are preserved under continuous transformations that have a continuous inverse transformation. The number of holes is such an invariant. Two subspaces with different number of holes in fact cannot be continuously deformed into each other. This explains, why S^2 (no hole) and the torus T^2 (one hole) are topologically different. On the other hand, two subspaces with the same number of holes are not necessarily topologically equivalent. An example is the 2-sphere and the x-y plane. These may differ in another invariant.

In electron band-structure theory we are interested in the topological classification of more complex structures. Consider a system of independent electrons on an infinite periodic lattice. Its Bloch Hamiltonian $H(\boldsymbol{k})$ parametrically depends on the wave vectors $\boldsymbol{k}$ in a reciprocal unit cell. In two dimensions, for example, this parameter space forms a torus T^2 because wave vectors at the boundaries of the reciprocal unit cell must be identified. However, it is not (only) the nontrivial topology of the parameter-space geometry that is the focus. We are rather interested in the topological classification of the *bundle* of ground states of $H(\boldsymbol{k})$ over T^2. Questions behind this are: Can these ground states be deformed into each other continuously? Can we topologically classify band structures in this way? In case of band insulators, this seemingly simple question has been answered affirmatively and comprehensively. The answer is surprisingly deep and requires substantial mathematical background. It has led to a classification scheme called "the 10-fold way", in which symmetry arguments and different topological invariants play the essential role.

The goal of this contribution is twofold: First, with the Chern number, we exemplarily pick a central topological invariant and, starting from basic concepts, discuss its precise definition, its range of applicability for the construction of topological phase diagrams, its connection with time-reversal symmetry, its experimental significance, etc. This elucidates the concept of "class A" Chern insulators, i.e., band insulators with a certain topological twist, which fall into one of the mentioned 10 classes.

Second, despite the great success of topological band theory, it cannot be straightforwardly extended to *interacting* electron systems. The development of novel concepts for the topological characterization and classification of the states of strongly correlated electron systems is one the major activities in theoretical physics, and in fact it is one of the most fascinating endeavors. One promising route is paved by combining dynamical mean-field theory with topological band theory, and the discussion of this idea and of an illuminating example will take a major part of the present work.
The presentation heavily makes use of various ideas from textbooks on topology and its applications in physics [1–4], on the results of the M.Sc. thesis of David Krüger, essentially published in Ref. [5], and of the B.Sc. thesis of Lara Heyl [6].

2 Chern number

The Chern number is a topological invariant. To define this invariant and to prove its quantization, we have to introduce and discuss a couple of basic theoretical concepts, including the adiabatic theorem, the Berry connection, gauge transformations and the Berry phase, and finally the Berry curvature.

Adiabatic theorem

Consider a system Hamiltonian H which smoothly depends on r parameters $\boldsymbol{R} = (R_1, ..., R_r)$

$$H = H(\boldsymbol{R}) = H(R_1, ..., R_r)\,. \tag{1}$$

The space of parameters $\mathcal{M}$ is a smooth r-dimensional manifold. We will further assume that the ground state $\big|\Phi_0(\boldsymbol{R})\big\rangle$ of $H(\boldsymbol{R})$ is unique (nondegenerate) for all $\boldsymbol{R}$. This implies that there is a finite gap $\Delta(\boldsymbol{R})$ to the first excited state $\big|\Phi_1(\boldsymbol{R})\big\rangle$ on the entire manifold $\mathcal{M}$. For each parameter configuration $\boldsymbol{R} \in \mathcal{M}$ the Hilbert space $\mathcal{H}$ is "locally" spanned by an orthonormal eigenbasis $\{\big|\Phi_j(\boldsymbol{R})\big\rangle\}$ of $H(\boldsymbol{R})$, i.e., we have

$$H(\boldsymbol{R})\big|\Phi_j(\boldsymbol{R})\big\rangle = E_j(\boldsymbol{R})\big|\Phi_j(\boldsymbol{R})\big\rangle\,. \tag{2}$$

Suppose that the parameters evolve in time as described by a trajectory $\boldsymbol{R} = \boldsymbol{R}(t)$ in $\mathcal{M}$. This means that the Hamiltonian acquires an explicit time dependence $H(t) = H\big(\boldsymbol{R}(t)\big)$. The formal solution of Schrödinger's equation provides us with the state of the system at time t: $\big|\Psi(t)\big\rangle = \mathcal{T}\exp\big(-i\int_0^t d\tau\, H\big(\boldsymbol{R}(\tau)\big)\big)\big|\Psi(0)\big\rangle$, where $\mathcal{T}$ is the time ordering and where have set $\hbar \equiv 1$. We assume that the state at time $t = 0$ is the ground state of $H\big(\boldsymbol{R}(t=0)\big)$, i.e., at $t = 0$ we have $\big|\Psi(0)\big\rangle = \big|\Phi_0\big(\boldsymbol{R}(0)\big)\big\rangle$.
The adiabatic theorem [7,8] states that, *if the time evolution of the parameters $\boldsymbol{R}(t)$ is sufficiently slow*, the system will remain in its ground state. More precisely, the state of the quantum system at time t is the instantaneous ground state of the Hamiltonian $H\big(\boldsymbol{R}(t)\big)$, i.e., we have $\big|\Psi(t)\big\rangle \propto \big|\Phi_0\big(\boldsymbol{R}(t)\big)\big\rangle$. This is a strong statement and it is nontrivial to precisely formulate the conditions under which this holds true. Roughly, its validity requires that the typical time scale

τ_{typical} of the parameter dynamics is much larger than the inverse of the gap: $\tau_{\text{typical}} \gg 1/\Delta$. Here, we take the adiabatic theorem for granted and will focus on the phase of the state. The adiabatic theorem, as we have formulated it, tells us that the state at time t is a ground state of $H\big(\boldsymbol{R}(t)\big)$,

$$\big|\Psi(t)\big\rangle = e^{-i\gamma(t)}\,\big|\Phi_0\big(\boldsymbol{R}(t)\big)\big\rangle\;, \tag{3}$$

but it leaves the phase factor $e^{-i\gamma(t)}$ yet undetermined.

Berry connection

Schrödinger's equation can be derived from the action principle $\delta S = \delta\int dt\, L = 0$ with the Lagrangian

$$L = L\Big(|\Psi\rangle, \partial_t|\Psi\rangle, \langle\Psi|, \partial_t\langle\Psi|,\, t\Big) = \big\langle\Psi\big|i\partial_t - H\big(\boldsymbol{R}(t)\big)\big|\Psi\big\rangle\;. \tag{4}$$

In fact, $0 = \partial L/\partial\langle\Psi| - \partial_t\big(\partial L/\partial(\langle\partial_t\Psi|)\big) = \big(i\partial_t - H\big(\boldsymbol{R}(t)\big)\big)|\Psi\rangle$, while the second Euler-Lagrange equation for $\langle\Psi|$ is just the adjoint Schrödinger equation. The Lagrange formalism is a useful tool, as one can incorporate the result of the adiabatic theorem as a holonomic constraint with the goal to get a simple equation of motion for the yet unknown phase $\gamma(t)$ only. Using Eq. (3) to express the "generalized coordinates" $\big|\Psi(t)\big\rangle$ and $\big\langle\Psi(t)\big|$ in terms of $\big|\Phi_0\big(\boldsymbol{R}(t)\big)\big\rangle$ and $\big\langle\Phi_0\big(\boldsymbol{R}(t)\big)\big|$, we get an effective Lagrangian,

$$L_{\text{eff}} = e^{i\overline{\gamma}(t)}\,\Big\langle\Phi_0\big(\boldsymbol{R}(t)\big)\Big|\big(i\partial_t - H\big(\boldsymbol{R}(t)\big)\big)\,e^{-i\gamma(t)}\Big|\Phi_0\big(\boldsymbol{R}(t)\big)\Big\rangle\;, \tag{5}$$

which indeed is a functional of the unknown phase of the system state only, while the local ground states $\big|\Phi_0\big(\boldsymbol{R}(t)\big)\big\rangle$ have fixed selected phases. Since $|\Psi\rangle$ and $\langle\Psi|$ must be treated as independent in Eq. (4), γ and $\overline{\gamma}$ have to be treated as independent as well. Hence, the Lagrangian L_{eff} has the functional form $L_{\text{eff}}(\gamma, \overline{\gamma}, \partial_t\gamma, \partial_t\overline{\gamma},\, t)$ with an explicit time dependence due to $\boldsymbol{R}(t)$. Carrying out the time derivative $i\partial_t$ of the terms on the right and using Eq. (2) for $j = 0$, we arrive at

$$L_{\text{eff}} = e^{i\overline{\gamma}(t)}e^{-i\gamma(t)}\Big(\partial_t\gamma(t) - E_0\big(\boldsymbol{R}(t)\big) + \boldsymbol{A}_0\big(\boldsymbol{R}(t)\big)\cdot\dot{\boldsymbol{R}}(t)\Big)\;. \tag{6}$$

Here,

$$\boldsymbol{A}_0(\boldsymbol{R}) = i\Big\langle\Phi_0(\boldsymbol{R})\Big|\frac{\partial}{\partial\boldsymbol{R}}\Big|\Phi_0(\boldsymbol{R})\Big\rangle \tag{7}$$

is the Berry connection, which mediates between the ground states at $\boldsymbol{R}$ and $\boldsymbol{R}+d\boldsymbol{R}$ on the manifold $\mathcal{M}$, and $\boldsymbol{A}_0(\boldsymbol{R})\cdot\dot{\boldsymbol{R}} = \sum_{\rho=1}^{r} A_{0,\rho}(\boldsymbol{R})\cdot\dot{R}_\rho$ with $\dot{R}_\rho = \partial_t R_\rho$. Note that the normalization condition for the state implies that $\boldsymbol{A}_0(\boldsymbol{R})$ is real. The resulting Euler-Lagrange equations for γ and $\overline{\gamma}$ are equivalent, such that eventually $\gamma = \overline{\gamma}$. The equation for γ,

$$\partial_t\gamma(t) = E_0\big(\boldsymbol{R}(t)\big) - \boldsymbol{A}_0\big(\boldsymbol{R}(t)\big)\cdot\dot{\boldsymbol{R}}(t)\;, \tag{8}$$

is easily solved by

$$\gamma(t) = \gamma(0) + \int_0^t E_0\big(\boldsymbol{R}(\tau)\big)\,d\tau - \int_0^t \boldsymbol{A}_0\big(\boldsymbol{R}(\tau)\big)\cdot\dot{\boldsymbol{R}}(\tau)\,d\tau\;, \tag{9}$$

Finally, with Eq. (3) and choosing w.l.o.g. $\gamma(0)=0$,

$$\big|\Psi(t)\big\rangle = \exp\left(-i\int_0^t E_0\big(\boldsymbol{R}(\tau)\big)\,d\tau\right)\exp\left(i\int_{\mathcal{C},\boldsymbol{R}(0)}^{\boldsymbol{R}(t)}\boldsymbol{A}_0(\boldsymbol{R})\cdot d\boldsymbol{R}\right)\big|\Phi_0\big(\boldsymbol{R}(t)\big)\big\rangle\,. \tag{10}$$

The first phase factor is called "dynamical", the second "geometrical". For the second, the substitution rule has been used in the last step. We see that it only depends on the path $\mathcal{C}$ from $\boldsymbol{R}(0)$ to $\boldsymbol{R}(t)$ in $\mathcal{M}$, a geometrical object, but not on the time dependence of the trajectory $\boldsymbol{R}(t)$, on "velocity", "acceleration" etc.

Gauge transformations and Berry phase

After having demonstrated the relevance of the Berry connection for the adiabatic time evolution of a quantum system, we will now turn to our main question, namely its relevance for the (time-independent) topological properties of the system and reconsider the bundle of ground states $\big|\Phi_0(\boldsymbol{R})\big\rangle$ with $\boldsymbol{R}\in\mathcal{M}$. The choice of an orthonormal energy eigenbasis for each $\boldsymbol{R}$, see Eq. (2), also implies the choice of a particular phase $\varphi(\boldsymbol{R})$ for each of the ground states in $\mathcal{M}$. This choice is arbitrary, so it is important to ask what happens if we make a different choice. Our naive expectation is that all observable ground-state properties of the system are invariant under a transformation of the phase, i.e., under a gauge transformation of the form

$$\big|\Phi_0(\boldsymbol{R})\big\rangle \mapsto e^{i\varphi(\boldsymbol{R})}\,\big|\Phi_0(\boldsymbol{R})\big\rangle\,. \tag{11}$$

Such an $\boldsymbol{R}$-dependent gauge transformation is called "local", as opposed to a "global" transformation $\big|\Phi_0(\boldsymbol{R})\big\rangle \mapsto e^{i\varphi}\big|\Phi_0(\boldsymbol{R})\big\rangle$. The system's topological properties are governed by the map $\boldsymbol{R}\mapsto e^{i\varphi(\boldsymbol{R})}$ from the parameter manifold $\mathcal{M}$ into $U(1)=\big\{z\in\mathbb{C}\,\big|\,|z|=1\big\}$. The fact that this map or, equivalently, the $\boldsymbol{R}$-dependent choice of the phase factors, has observable consequences as has been pointed out by M. Berry [9].
Under a local gauge transformation, Eq. (11), the Berry connection transforms as

$$\boldsymbol{A}_0(\boldsymbol{R}) \mapsto \boldsymbol{A}_0'(\boldsymbol{R}) = \boldsymbol{A}_0(\boldsymbol{R}) - \frac{\partial\varphi(\boldsymbol{R})}{\partial\boldsymbol{R}}\,. \tag{12}$$

This implies that $\boldsymbol{A}_0(\boldsymbol{R})$ is gauge dependent. It has no direct physical meaning but depends on the *arbitrary* choice of the phases $\varphi(\boldsymbol{R})$ of the local ground states $|\Phi_0(\boldsymbol{R})\rangle$. The situation is particularly interesting, if $\mathcal{C}$ is a *closed* path in $\mathcal{M}$, given by some parametrization $\boldsymbol{R}=\boldsymbol{R}(\lambda)$, where the real parameter λ *could* stand for the time. In this case $\gamma_{\mathcal{C}}=\oint_{\mathcal{C}}\boldsymbol{A}_0(\boldsymbol{R})\cdot d\boldsymbol{R}$ is called the Berry phase. Under a gauge transformation, the Berry phase transforms as $\gamma_{\mathcal{C}}\mapsto\gamma_{\mathcal{C}}-\oint_{\mathcal{C}}\big(\partial\varphi(\boldsymbol{R})/\partial\boldsymbol{R}\big)\cdot d\boldsymbol{R}$, see Eq. (12). If certain rather general conditions are satisfied, as specified by the Poincaré lemma and regarding, in particular, the continuity of $\partial\varphi(\boldsymbol{R})/\partial\boldsymbol{R}$, we can argue that the Berry phase is invariant $\gamma_{\mathcal{C}}\mapsto\gamma_{\mathcal{C}}-\oint_{\mathcal{C}}d\varphi(\boldsymbol{R})=\gamma_{\mathcal{C}}$.
However, the integral along $\mathcal{C}$ can be nonzero in fact. Consider, for example, $\mathcal{M}=\mathbb{R}^3$ and the circle $\mathcal{C}=\big\{\boldsymbol{R}=(X,Y,Z)^T\,\big|\,X^2{+}Y^2=R^2\,,\,Z{=}0\big\}$ of radius R around $\boldsymbol{R}=0$ in the X-Y plane. Suppose $\boldsymbol{A}_0'(\boldsymbol{R})=\boldsymbol{A}_0(\boldsymbol{R})-(-Y,X,0)^T/(X^2{+}Y^2)$. *Locally*, we have $\boldsymbol{A}_0'(\boldsymbol{R})=$

$\boldsymbol{A}_0(\boldsymbol{R}) - (\partial/\partial\boldsymbol{R})\arctan(Y/X)$. However, the map $\varphi(\boldsymbol{R}) = \arctan(Y/X)$ is discontinuous on the $X = 0$ plane, and $\partial\varphi(\boldsymbol{R})/\partial\boldsymbol{R}$ is singular on the Z axis. In fact, it is straightforward to see that $\oint_{\mathcal{C}}\big(\partial\varphi(\boldsymbol{R})/\partial\boldsymbol{R}\big)\cdot d\boldsymbol{R} = 2\pi \neq 0$.
Generally, under an arbitrary gauge transformation, we have

$$\gamma_{\mathcal{C}} = \oint_{\mathcal{C}} \boldsymbol{A}_0(\boldsymbol{R})\cdot d\boldsymbol{R} \mapsto \gamma_{\mathcal{C}} + 2\pi k \tag{13}$$

with $k \in \mathbb{Z}$. This is equivalent with the statement that the Berry phase factor $e^{i\gamma_{\mathcal{C}}}$ is always gauge invariant. To verify this, we rewrite the integral as the limit of a Riemann sum

$$\begin{aligned}\exp\left(i\gamma_{\mathcal{C}}\right) &= \exp\left(i\oint_{\mathcal{C}} \boldsymbol{A}_0(\boldsymbol{R})\cdot d\boldsymbol{R}\right) = \lim_{n\to\infty}\exp\Big(i\boldsymbol{A}_0(\boldsymbol{R}_n)\cdot\Delta\boldsymbol{R}_n + \cdots + i\boldsymbol{A}_0(\boldsymbol{R}_1)\cdot\Delta\boldsymbol{R}_1\Big)\\ &= \lim_{n\to\infty}\big(1+i\boldsymbol{A}_0(\boldsymbol{R}_n)\cdot\Delta\boldsymbol{R}_n\big)\cdots\big(1+i\boldsymbol{A}_0(\boldsymbol{R}_1)\cdot\Delta\boldsymbol{R}_1\big)\,.\end{aligned} \tag{14}$$

This is a Volterra product integral. The k-th factor can be written as $1 + i\boldsymbol{A}_0(\boldsymbol{R}_k)\cdot\Delta\boldsymbol{R}_k = 1 - \big\langle\Phi_0(\boldsymbol{R}_k)\big|\Big(\big|\Phi_0(\boldsymbol{R}_k)\big\rangle - \big|\Phi_0(\boldsymbol{R}_k-\Delta\boldsymbol{R}_k)\big\rangle\Big) = \big\langle\Phi_0(\boldsymbol{R}_k)\big|\Phi_0(\boldsymbol{R}_k-\Delta\boldsymbol{R}_k)\big\rangle$, with the help of Eq. (7). We have $\boldsymbol{R}_k - \Delta\boldsymbol{R}_k = \boldsymbol{R}_{k-1}$ and, since $\mathcal{C}$ is a closed path, we can identify $\boldsymbol{R}_0$ with $\boldsymbol{R}_n$. This yields

$$\begin{aligned}\exp\left(i\gamma_{\mathcal{C}}\right) &= \lim_{n\to\infty}\big\langle\Phi_0(\boldsymbol{R}_n)\big|\Phi_0(\boldsymbol{R}_{n-1})\big\rangle\cdots\big\langle\Phi_0(\boldsymbol{R}_1)\big|\Phi_0(\boldsymbol{R}_0)\big\rangle\\ &= \lim_{n\to\infty}\mathrm{tr}\Big(\big|\Phi_0(\boldsymbol{R}_n)\big\rangle\big\langle\Phi_0(\boldsymbol{R}_n)\big|\cdots\big|\Phi_0(\boldsymbol{R}_1)\big\rangle\big\langle\Phi_0(\boldsymbol{R}_1)\big|\Phi_0(\boldsymbol{R}_0)\big\rangle\big\langle\Phi_0(\boldsymbol{R}_0)\big|\Big)\,.\end{aligned} \tag{15}$$

where the latter representation is manifestly gauge invariant, since the phases cancel out in each of the dyadic products.

Example: Toy model with two-dimensional Hilbert space

As an example, we will consider the following toy Hamiltonian on a three-dimensional parameter manifold

$$H = -\frac{1}{2}\boldsymbol{R}\cdot\boldsymbol{\tau}\,, \tag{16}$$

Here $\boldsymbol{R} \in \mathcal{M} = \mathbb{R}^3 \setminus \{0\}$ and $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)^T$ is the vector of Pauli matrices. Actually, this form of H is generic for a quantum system with a two-dimensional Hilbert space. The 2×2 matrix representing the Hamiltonian can be expanded in the Hermitian basis $\{\mathbf{1}, \tau_1, \tau_2, \tau_3\}$, which yields Eq. (16) after dropping the uninteresting term $\propto \mathbf{1}$.
When interpreting the three parameters $\boldsymbol{R}$ as the components of a field strength $\boldsymbol{B}$, the model Eq. (16) just describes a local spin $1/2$ in an external magnetic field. We thus expect a Zeeman splitting of the two eigenstates of H. In fact we have $E_0(\boldsymbol{R}) = -R/2$ and $E_1(\boldsymbol{R}) = +R/2$ where $R = |\boldsymbol{R}|$. Note that we have intentionally excluded the point $\boldsymbol{R} = 0$ from the manifold $\mathcal{M}$, as here the ground state would be degenerate and our entire construction would break down. The $\boldsymbol{R}$-dependent normalized ground state of the model is readily calculated and is conveniently expressed in terms of spherical coordinates $\boldsymbol{R} = R\big(\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta\big)^T$ as

$$\big|\Phi_0(\boldsymbol{R})\big\rangle = \big|\Phi_0(\theta,\phi)\big\rangle = \begin{pmatrix}\cos(\theta/2)\\ e^{i\phi}\sin(\theta/2)\end{pmatrix}\,. \tag{17}$$

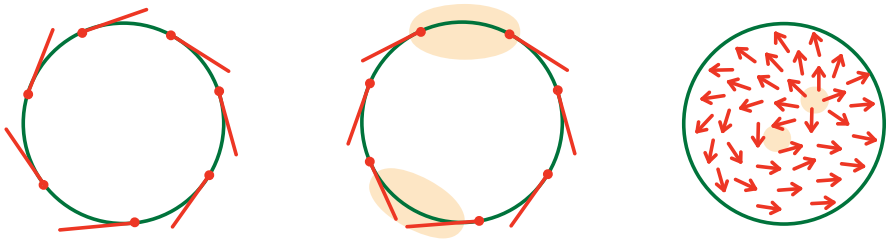

Fig. 2: *Hairy-ball theorem: There is no everywhere nonvanishing and continuous tangent vector field on an n-sphere, if n is even.* Left*: Sketch of a continuous tangent vector field for $n = 1$.* Middle*: Tangent vector field with two twirls for $n = 1$.* Right*: Tangent vector field on S^2 with (already) two twirls. At each $\boldsymbol{R} \in S^2$, the direction of the arrow with unit length specifies the phase $\varphi(\boldsymbol{R})$.*

In fact, the ground state is independent of R, and it is *almost everywhere* a smooth and invertible function of θ and ϕ. The angle variables provide us with a chart of the 2-sphere $S^2 = \{\boldsymbol{R} \,|\, R{=}1\}$. However, the state that is obtained in the limit $\theta \to \pi$, i.e., when approaching the south pole of S^2, will depend on ϕ. This means that the map $(\theta, \phi) \mapsto \boldsymbol{R} \mapsto |\Phi_0(\boldsymbol{R})\rangle$ is discontinuous at the south pole. Hence, the parameter manifold S^2 cannot be described by a single chart. A second one is necessary for a complete atlas.

Choosing a different gauge is another way to deal with the singularity since this will not affect any *physical*, gauge-invariant properties. In fact, the gauge transformation $|\Phi_0(\boldsymbol{R})\rangle \mapsto e^{-i\phi}|\Phi_0(\boldsymbol{R})\rangle$ does the job. It removes the singularity at the south pole — but it also introduces a similar one at the north pole. Clearly, the problem cannot be solved with a *continuous* gauge transformation, and the proposed one is itself *discontinuous*. As we have seen above, $\boldsymbol{R} = (X, Y, Z)^T \mapsto \varphi(\boldsymbol{R}) = \phi = \arctan(Y/X)$ is discontinuous on the entire $X = 0$ plane. Generally, there is no gauge such that the phase of $|\Phi_0(\boldsymbol{R})\rangle$ is continuous on the entire manifold $\mathcal{M}$, which essentially is the 2-sphere. The 2-sphere, with a phase factor $e^{i\varphi(\boldsymbol{R})} \in U(1) \cong S^1$ attached to each point, is like a hairy sphere, so that we can apply the "hairy-ball theorem" [4], see Fig. 2. This states that any attempt to comb a hairy billiard ball (the 2-sphere) smoothly, without twirls, must fail and that there must be at least one singular point.

On the equator $\theta = \pi/2$, oriented from west to east, and for arbitrary $\phi \in [0, 2\pi[$, the ϕ component of the Berry connection obtained from Eq. (17) is

$$\boldsymbol{A}_0(\theta, \phi)\, R\boldsymbol{e}_\phi = i\langle \Phi_0(\theta, \phi)|\partial_\phi|\Phi_0(\theta, \phi)\rangle = -\frac{1}{2}\,, \tag{18}$$

where $\boldsymbol{e}_\phi = (1/R)\partial\boldsymbol{R}/\partial\phi$. Here, we have used the definition Eq. (12) and the representation $\partial/\partial\boldsymbol{R} = \boldsymbol{e}_\phi(R\sin\theta)^{-1}(\partial/\partial\phi) + \boldsymbol{e}_\theta(\cdots) + \boldsymbol{e}_R(\cdots)$ with $\theta = \pi/2$ and $R = 1$. Computing the line integral along the equator, where $d\boldsymbol{R} = R\boldsymbol{e}_\phi\, d\phi$, yields the Berry phase

$$\gamma_{\text{equator}} = \oint_{\text{equator}} \boldsymbol{A}_0(\boldsymbol{R}){\cdot}d\boldsymbol{R} = \int_0^{2\pi} \boldsymbol{A}_0(\theta, \phi){\cdot}R\boldsymbol{e}_\phi\, d\phi = -\pi\,. \tag{19}$$

In the context of adiabatic dynamics, this has the interesting consequence that the system state $|\Psi\rangle$ picks up a sign $e^{i\gamma_C} = -1$ when $\boldsymbol{R}$ is slowly steered once around the equator C.

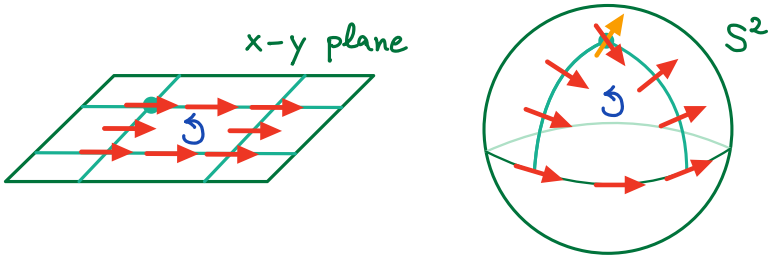

Fig. 3: *Parallel transport along a closed path on a flat and on a curved space.*

In the above-discussed alternative gauge after the transformation $\big|\Phi_0(\boldsymbol{R})\big\rangle \mapsto e^{-i\phi}\big|\Phi_0(\boldsymbol{R})\big\rangle = \big(e^{-i\phi}\cos(\theta/2), \sin(\theta/2)\big)^T$, we get a different result for the Berry connection, $\boldsymbol{A}_0'(\theta,\phi)\cdot R\boldsymbol{e}_\phi = +\frac{1}{2}$, showing again that $\boldsymbol{A}_0'(\boldsymbol{R})$ is not gauge invariant. Computing the corresponding Berry phase yields $\gamma_{\text{equator}} = +\pi$. This is consistent with our general reasoning: The Berry phase $\gamma_{\mathcal{C}}$ is gauge invariant up to an integer multiple of 2π only, but the Berry phase factor $e^{i\gamma_{\mathcal{C}}}$ is fully gauge invariant.

Berry curvature

More elegantly, we can compute the Berry phase $\gamma_{\mathcal{C}}$ with the help of the Stokes theorem

$$\gamma_{\mathcal{C}} = \oint_{\mathcal{C}} \boldsymbol{A}_0(\boldsymbol{R})\cdot d\boldsymbol{R} = \int_{\mathcal{S}} \boldsymbol{\Omega}_0(\boldsymbol{R})\cdot d\boldsymbol{S}\ . \tag{20}$$

Here $\boldsymbol{S}$ is the unit normal vector of the surface $\mathcal{S}$ on the sphere of fixed radius R that is bounded by the closed path $\mathcal{C} = \partial\mathcal{S}$, and $\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}}\times\boldsymbol{A}_0(\boldsymbol{R})$ is the curl of the Berry connection. The orientation of the path $\mathcal{C}$ is inherited from the orientation of $\mathcal{S}$ in the usual way by the right-hand rule. $\boldsymbol{\Omega}_0(\boldsymbol{R})$ is called the (ground-state) Berry curvature. In this form, the Stokes theorem is formulated for applications to a three-dimensional parameter space only, and we will stick to the above toy model. Qualitatively, however, the entire discussion holds for arbitrary r-dimensional manifolds $\mathcal{M}$, and a brief discussion of the general case will be given later.

The Berry phase γ_{C} is a quantity, which measures the degree to which parallel transport fails to preserve the geometrical data that is being transported. Parallel transport along closed loops is in a way trivial on flat spaces but leaves a detectable footprint for a curved space, see Fig. 3. In fact, Eq. (20) demonstrates that a finite Berry phase requires a finite Berry curvature $\boldsymbol{\Omega}_0(\boldsymbol{R})$. Eq. (19) shows that parallel transport of states in Hilbert space does not follow descriptive Euclidean geometry. Rather, it takes two circumnavigations of the equator instead of just one to get back to the same state.

The Berry curvature is a highly convenient quantity since it is invariant under arbitrary gauge transformations. This follows from its definition, from the transformation behavior (12) of the Berry connection, and from the fact that the curl of any gradient field vanishes. Another important property is $\partial_{\boldsymbol{R}}\boldsymbol{\Omega}_0(\boldsymbol{R}) = 0$, i.e., the Berry curvature is divergence-free. Again, this holds by definition since $\partial_{\boldsymbol{R}}\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}}\cdot\partial_{\boldsymbol{R}}\times\boldsymbol{A}_0(\boldsymbol{R}) = 0$.

Vice versa, there is the mathematical question, under which conditions a divergence-free vector field $\boldsymbol{\Omega}_0(\boldsymbol{R})$ can be expressed as the curl of another vector field on the entire manifold $\mathcal{M}$. This is answered by the following form of the Poincaré lemma: If the manifold $\mathcal{M}$ is an open ball or, more generally, a star-shaped open set in $\mathbb{R}^3$, one can find a single field $\boldsymbol{A}_0(\boldsymbol{R})$ such that $\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}} \times \boldsymbol{A}_0(\boldsymbol{R})$ *globally*, i.e., for *all* $\boldsymbol{R} \in \mathcal{M}$. Note that *locally* a representation as the curl of a vector field is always possible.

In case of topologically more complicated manifolds, a global representation may not be possible. The manifold $\mathcal{M} = \mathbb{R}^3 \backslash \{0\}$ relevant for the toy model or, since the ground state does not depend on R, the submanifold S^2, are in fact examples for manifolds that cannot be covered by a single chart and are thus topologically different from $\mathbb{R}^3$ or from $\mathbb{R}^2$, respectively. Consequently, the Stokes theorem Eq. (20) cannot be applied if $\mathcal{S} = S^2$. We would mistakenly get $\gamma_{\mathcal{C}} = 0$ since the path $\mathcal{C}$ vanishes in this case.

Let us compute the Berry curvature for our toy model concretely. Its α-th component

$$\Omega_{0,\alpha}(\boldsymbol{R}) = \frac{1}{2} \sum_{\beta\gamma} \varepsilon_{\alpha\beta\gamma}\, \Omega_{0,\beta\gamma}(\boldsymbol{R}) \tag{21}$$

can be expressed in terms of an antisymmetric real matrix

$$\begin{aligned} \Omega_{0,\beta\gamma}(\boldsymbol{R}) &= \partial_\beta A_{0,\gamma}(\boldsymbol{R}) - \partial_\gamma A_{0,\beta}(\boldsymbol{R}) = i\Big(\big\langle \partial_\beta \Phi_0(\boldsymbol{R}) \big| \partial_\gamma \Phi_0(\boldsymbol{R}) \big\rangle - \big\langle \partial_\gamma \Phi_0(\boldsymbol{R}) \big| \partial_\beta \Phi_0(\boldsymbol{R}) \big\rangle\Big) \\ &= -2\,\mathrm{Im} \big\langle \partial_\beta \Phi_0(\boldsymbol{R}) \big| \partial_\gamma \Phi_0(\boldsymbol{R}) \big\rangle = -2\,\mathrm{Im} \sum_{j\neq 0} \big\langle \partial_\beta \Phi_0(\boldsymbol{R}) \big| \Phi_j(\boldsymbol{R}) \big\rangle \big\langle \Phi_j(\boldsymbol{R}) \big| \partial_\gamma \Phi_0(\boldsymbol{R}) \big\rangle\,. \end{aligned} \tag{22}$$

For the second equality, we have used Eq. (7) and the fact that partial derivatives commute in case of a smooth $\boldsymbol{R}$ dependence. For the last, we have inserted a resolution of the identity (at $\boldsymbol{R}$). Note that the term $j = 0$ does not contribute as for $j = 0$ the matrix elements are real. Using the identity $\big\langle \Phi_j(\boldsymbol{R}) \big| \partial_\gamma \Phi_0(\boldsymbol{R}) \big\rangle = \big\langle \Phi_j(\boldsymbol{R}) \big| \partial_\gamma H(\boldsymbol{R}) \big| \Phi_0(\boldsymbol{R}) \big\rangle / \big(E_0(\boldsymbol{R}) - E_j(\boldsymbol{R})\big)$, valid for $j \neq 0$, we finally have

$$\Omega_{0,\beta\gamma}(\boldsymbol{R}) = -2\,\mathrm{Im} \sum_{j\neq 0} \frac{\big\langle \Phi_0(\boldsymbol{R}) \big| \partial_\beta H(\boldsymbol{R}) \big| \Phi_j(\boldsymbol{R}) \big\rangle \big\langle \Phi_j(\boldsymbol{R}) \big| \partial_\gamma H(\boldsymbol{R}) \big| \Phi_0(\boldsymbol{R}) \big\rangle}{\big(E_0(\boldsymbol{R}) - E_j(\boldsymbol{R})\big)^2}\,. \tag{23}$$

In case of the toy model, the j-sum merely consists of a single term, referring to the excited state $\big|\Phi_1(\boldsymbol{R})\big\rangle$. Furthermore, $\partial_\beta H(\boldsymbol{R}) = -\tau_\beta/2$, see Eq. (16), and $\big(E_0(\boldsymbol{R}) - E_1(\boldsymbol{R})\big)^2 = R^2$. Hence

$$\Omega_{0,\beta\gamma}(\boldsymbol{R}) = -\frac{1}{2}\,\mathrm{Im} \frac{\big\langle \Phi_0(\boldsymbol{R}) \big| \tau_\beta \big| \Phi_1(\boldsymbol{R}) \big\rangle \big\langle \Phi_1(\boldsymbol{R}) \big| \tau_\gamma \big| \Phi_0(\boldsymbol{R}) \big\rangle}{R^2} \tag{24}$$

and with Eq. (21)

$$\boldsymbol{\Omega}_0(\boldsymbol{R}) = \frac{1}{2} \sum_{\alpha\beta\gamma} \varepsilon_{\alpha\beta\gamma}\, \Omega_{0,\beta\gamma}(\boldsymbol{R})\, \boldsymbol{e}_\alpha = -\frac{1}{4}\,\mathrm{Im} \frac{\big\langle \Phi_0(\boldsymbol{R}) \big| \boldsymbol{\tau} \big| \Phi_1(\boldsymbol{R}) \big\rangle \times \big\langle \Phi_1(\boldsymbol{R}) \big| \boldsymbol{\tau} \big| \Phi_0(\boldsymbol{R}) \big\rangle}{R^2}\,. \tag{25}$$

We can choose the z axis to point in the direction of $\boldsymbol{R}$, such that $\big|\Phi_0(\boldsymbol{R})\big\rangle = (1,0)^T$ with Eq. (17) and $\big|\Phi_1(\boldsymbol{R})\big\rangle = e^{i\phi}(0,1)^T$. Evaluating the matrix element

$$\big\langle\Phi_0(\boldsymbol{R})\big|\boldsymbol{\tau}\big|\Phi_1(\boldsymbol{R})\big\rangle = e^{i\phi}(1,0)\,\boldsymbol{\tau}\begin{pmatrix}0\\1\end{pmatrix} = e^{i\phi}\begin{pmatrix}1\\-i\\0\end{pmatrix}, \quad \big\langle\Phi_1(\boldsymbol{R})\big|\boldsymbol{\tau}\big|\Phi_0(\boldsymbol{R})\big\rangle = e^{-i\phi}\begin{pmatrix}1\\i\\0\end{pmatrix}, \tag{26}$$

and remembering that we had assumed $\boldsymbol{R} = R\boldsymbol{e}_z$, we find

$$\boldsymbol{\Omega}_0(\boldsymbol{R}) = -\frac{1}{2}\frac{\boldsymbol{R}}{R^3}\,. \tag{27}$$

Since $\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}}\times\boldsymbol{A}_0(\boldsymbol{R})$, one may ask the question, if there is a similarly compact result for the Berry connection. In fact, one may verify by a direct calculation that, for an arbitrarily chosen unit vector $\boldsymbol{e}$, the curl of

$$\boldsymbol{A}_0(\boldsymbol{R}) = -\frac{1}{2}\frac{1}{R^2}\,\frac{\boldsymbol{e}\times\boldsymbol{R}}{1+\boldsymbol{e}\boldsymbol{R}/R} \tag{28}$$

yields the Berry curvature (27), as has already been noted by P. Dirac in 1931 [10].

Equations (27) and (28) are interesting as they allow us to interpret $\boldsymbol{A}_0(\boldsymbol{R})$ as the vector potential of the magnetic field $\boldsymbol{\Omega}_0(\boldsymbol{R})$ of a "magnetic monopole" with "magnetic point charge" $q=-1/2$ located at the origin $\boldsymbol{R}=0$. Note that $\boldsymbol{A}_0(\boldsymbol{R})$ is singular for all $\boldsymbol{R}$ with $\boldsymbol{R}=-R\boldsymbol{e}$, i.e., on the ray (Dirac string) starting at $\boldsymbol{R}=0$ to infinity in $-\boldsymbol{e}$ direction. Different choices for $\boldsymbol{e}$ just correspond to different gauges. The singular point $\boldsymbol{R}=-\boldsymbol{e}$, where the ray intersects S^2, can be moved around on S^2 by choosing different gauges but it cannot be removed for any particular gauge. Hence, $\boldsymbol{\Omega}_0(\boldsymbol{R})=\partial_{\boldsymbol{R}}\times\boldsymbol{A}_0(\boldsymbol{R})$ holds locally but not *globally* on S^2 or on $\mathbb{R}^3\backslash\{0\}$. This is consistent with our earlier discussion that the Poincaré lemma does not apply to S^2.

There is also a descriptive meaning of the Berry phase $\gamma_{\mathcal{C}}$. Equation (20) tells us that $\gamma_{\mathcal{C}}$ is the flux of the "magnetic field" through a surface that is bounded by the closed path $\mathcal{C}$. The right-hand rule tells us that, for $\mathcal{C}$ being the equator, oriented from west to east, $\mathcal{S}$ is the upper hemisphere. Using Eq. (27) with $d\boldsymbol{S}=R^2(\boldsymbol{R}/R)\,dS$, we see that $\gamma_{\mathcal{C}} = -\frac{1}{2}\int_S(\boldsymbol{R}/R^3)d\boldsymbol{S} = -\pi$ for any great circle on the sphere of arbitrary radius. This generalizes our earlier result, Eq. (19). We can also integrate over the lower hemisphere $\pi/2<\theta<\pi$ to compute the Berry phase via the Stokes theorem. According to Eq. (27), the collected flux is the same: $\int_S\boldsymbol{\Omega}_0(\boldsymbol{R})d\boldsymbol{S} = -\pi$. However, in applying the Stokes theorem with the same path $\mathcal{C}$ (the equator) we have to take care of an additional sign, $\oint_{\mathcal{C}}\boldsymbol{A}_0(\boldsymbol{R})d\boldsymbol{R} = -\int_{\mathcal{S}}\boldsymbol{\Omega}_0(\boldsymbol{R})\,d\boldsymbol{S}$, since the orientations of $\mathcal{C}$ and $\mathcal{S}$ are not consistent in this case (the right-hand rule is violated). There is a second point to be considered: The Stokes theorem holds if $\boldsymbol{A}_0(\boldsymbol{R})$ is smooth on (an open set containing) the lower hemisphere, which is not the case for the Berry connection derived from Eq. (17). In the above-discussed alternative gauge after the transformation $|\Phi_0(\boldsymbol{R})\rangle\mapsto e^{-i\phi}|\Phi_0(\boldsymbol{R})\rangle$, the Berry connection $\boldsymbol{A}_0'(\boldsymbol{R})$ is in fact smooth on the lower hemisphere. However, $\oint_{\mathcal{C}}\boldsymbol{A}_0'(\boldsymbol{R})d\boldsymbol{R} = \pi$, opposed to Eq. (19), but consistent with the Stokes theorem: $\int_{\mathcal{S}}\boldsymbol{\Omega}_0(\boldsymbol{R})\,d\boldsymbol{S} = -\oint_{\mathcal{C}}\boldsymbol{A}_0'(\boldsymbol{R})\,d\boldsymbol{R} = -\pi$.

Chern number

We are now in a position to define the Chern number and to understand why this is a topological invariant. Given the Hamiltonian $H(\boldsymbol{R})$ of a quantum system with a nondegenerate ground state for all $\boldsymbol{R}$ on its parameter manifold $\mathcal{M}$, our goal is to show that the first Chern number, defined as

$$C = \frac{1}{2\pi} \oint_{\mathcal{S}} \boldsymbol{\Omega}_0(\boldsymbol{R}) \cdot d\boldsymbol{S} \,, \tag{29}$$

is quantized: $C \in \mathbb{Z}$.

Here, we consider a *closed* two-dimensional surface $\mathcal{S}$ embedded in a two- or higher-dimensional manifold $\mathcal{M}$. The Chern number will generally be dependent on our choice for $\mathcal{S}$. For simplicity, we still stick to a three-dimensional parameter space, e.g., to $\mathcal{M} = \mathbb{R}^3 \backslash \{0\}$, which applies to our toy model. The Berry curvature $\boldsymbol{\Omega}_0(\boldsymbol{R})$ is divergence-free. Hence, locally $\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}} \times \boldsymbol{A}_0(\boldsymbol{R})$. If $\mathcal{M}$ could be covered by a single chart, i.e., if it is essentially given by $\mathbb{R}^3$, then the Poincaré lemma would tell us that $\boldsymbol{\Omega}_0(\boldsymbol{R})$ is *globally* given as the curl of the Berry connection $\boldsymbol{A}_0(\boldsymbol{R})$ defined on the *entire* manifold $\mathcal{M}$. In this case $C = 0$, Namely, the Stokes theorem gives $\oint_{\mathcal{S}} \partial_{\boldsymbol{R}} \times \boldsymbol{A}_0(\boldsymbol{R}) d\boldsymbol{S} = \oint_{\partial \mathcal{S}} \boldsymbol{A}_0(\boldsymbol{R}) d\boldsymbol{R} = 0$ since the boundary of $\mathcal{S}$ vanishes.

Consider now the case of a nontrivial $\mathcal{M}$ and assume that the closed surface $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$, where $\mathcal{S}_1$ and $\mathcal{S}_2$ are surfaces, which both share the same closed path $\mathcal{C}$ as their boundary, except for the different orientation: $\partial \mathcal{S}_1 = -\partial \mathcal{S}_2 = \mathcal{C}$. Let $\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}} \times \boldsymbol{A}_{0,1}(\boldsymbol{R})$ on $\mathcal{S}_1$ and $\boldsymbol{\Omega}_0(\boldsymbol{R}) = \partial_{\boldsymbol{R}} \times \boldsymbol{A}_{0,2}(\boldsymbol{R})$ on $\mathcal{S}_2$. On an environment of $\mathcal{C}$, the Berry connections $\boldsymbol{A}_{0,1}(\boldsymbol{R})$ and $\boldsymbol{A}_{0,2}(\boldsymbol{R})$ satisfy $\partial_{\boldsymbol{R}} \times \big(\boldsymbol{A}_{0,1}(\boldsymbol{R}) - \boldsymbol{A}_{0,2}(\boldsymbol{R})\big) = 0$, and thus differ by a gradient field, i.e., transform into each other via a gauge transformation. Thus, the corresponding Berry phase factors are equal, $e^{i\gamma_{1,\mathcal{C}}} = e^{i\gamma_{2,\mathcal{C}}}$, while the Berry phases are equal modulo $2\pi k$ with $k \in \mathbb{Z}$:

$$\oint_{\mathcal{C}} \boldsymbol{A}_{0,1}(\boldsymbol{R}) \cdot d\boldsymbol{R} - \oint_{\mathcal{C}} \boldsymbol{A}_{0,2}(\boldsymbol{R}) \cdot d\boldsymbol{R} = 2\pi k \,. \tag{30}$$

We choose the orientation of $\mathcal{C}$ consistent with the orientation of $\mathcal{S}_1$ according to the right-hand rule. The Stokes theorem then gives

$$\int_{\mathcal{S}_1} \boldsymbol{\Omega}(\boldsymbol{R}) \cdot d\boldsymbol{S}_1 = \oint_{\mathcal{C}} \boldsymbol{A}_{0,1}(\boldsymbol{R}) \cdot d\boldsymbol{R} \quad \text{and} \quad \int_{\mathcal{S}_2} \boldsymbol{\Omega}(\boldsymbol{R}) \cdot d\boldsymbol{S}_2 = -\oint_{\mathcal{C}} \boldsymbol{A}_{0,2}(\boldsymbol{R}) \cdot d\boldsymbol{R} \,, \tag{31}$$

where the sign in the second equation comes from the fact that the orientations of $\mathcal{C}$ and $\mathcal{S}_2$ are opposite. With Eq. (30) and Eq. (31) we conclude that

$$\oint_{\mathcal{S}} \boldsymbol{\Omega}(\boldsymbol{R}) \cdot d\boldsymbol{S} = \int_{\mathcal{S}_1} \boldsymbol{\Omega}(\boldsymbol{R}) \cdot d\boldsymbol{S}_1 + \int_{\mathcal{S}_2} \boldsymbol{\Omega}(\boldsymbol{R}) \cdot d\boldsymbol{S}_2 = \oint_{\mathcal{C}} \boldsymbol{A}_{0,1}(\boldsymbol{R}) \cdot d\boldsymbol{R} - \oint_{\mathcal{C}} \boldsymbol{A}_{0,2}(\boldsymbol{R}) \cdot d\boldsymbol{R} \in 2\pi\mathbb{Z}. \tag{32}$$

This is exactly the proposition made and thus concludes the proof.

Inserting the result we have obtained for the toy model Eq. (27) in the definition of the Chern number Eq. (29), we find a nontrivial (nonzero) invariant

$$C = \frac{1}{2\pi} \oint_{\mathcal{S}} \left(-\frac{1}{2} \frac{\boldsymbol{R}}{R^3} \right) \cdot d\boldsymbol{S} = -\frac{1}{2} \frac{1}{2\pi} 4\pi = -1 \in \mathbb{Z} \,, \tag{33}$$

if we choose the closed surface $\mathcal{S}$ such that it once encloses the "magnetic monopole" at $\boldsymbol{R} = 0$. Taking a sphere centered around $\boldsymbol{R} = 0$ is the most simple choice, and the result Eq. (33) is

obtained straightforwardly. Let us emphasize, however, that it also holds for any continuous deformation of the sphere as long as $\boldsymbol{R} = 0$ is included, i.e., if there is no gap closure during the deformation.

We conclude the discussion with two general points: First, it should be stressed that the Chern number is assigned to a *bundle* $\{|\Phi_0(\boldsymbol{R})\rangle\}_{\mathcal{S}}$ of ground states of Hamiltonians $H(\boldsymbol{R})$ with parameters $\boldsymbol{R}$ belonging to a closed two-dimensional manifold $\mathcal{S}$. This can be indicated by writing $C(\mathcal{S})$ or, more generally, $C(\mathcal{S}, \boldsymbol{\lambda})$, where we made explicit the dependence of the Hamiltonian $H = H(\boldsymbol{R}, \boldsymbol{\lambda})$ on additional parameters $\boldsymbol{\lambda}$, which are unrelated to $\boldsymbol{R}$ and which are called *control parameters*. Examples for $\boldsymbol{\lambda}$ will be seen in the next section. A nonzero Chern number can be interpreted as a nontrivial twist of this bundle. We can now state that the bundle $\{|\Phi_0(\boldsymbol{R}, \boldsymbol{\lambda}_1)\rangle\}_{\mathcal{S}}$ cannot be deformed continuously into the bundle $\{|\Phi_0(\boldsymbol{R}, \boldsymbol{\lambda}_2)\rangle\}_{\mathcal{S}}$, if the corresponding Chern numbers differ from each other, i.e., if $C(\mathcal{S}, \boldsymbol{\lambda}_1) \neq C(\mathcal{S}, \boldsymbol{\lambda}_2)$. In passing from $\boldsymbol{\lambda}_1$ to $\boldsymbol{\lambda}_2$ in control-parameter space, one would necessarily have to go through a critical point $\boldsymbol{\lambda}_c$, where there is a gap closure at one (or several) points $\boldsymbol{R} \in \mathcal{S}$, since the Chern number cannot change as long the ground state is nondegenerate.

Second, the entire construction can be generalized from three-dimensional parameter manifolds $\mathcal{M}$ to manifolds of arbitrary dimension r using the calculus of differential forms and the generalized Stokes theorem. In this language, the Berry connection is a 1-form, and the Berry curvature Ω a 2-form, which can be integrated over a closed two-dimensional submanifold of $\mathcal{M}$, and are thus relevant for $r = 3$. By construction Ω is *closed*, i.e., $d\Omega = 0$. In case of a topologically nontrivial manifold $\mathcal{M}$, however, it is not necessarily *exact*, i.e., of the form $\Omega = dA$. For a bundle of Hamiltonians $H(\boldsymbol{R})$ with parameters $\boldsymbol{R}$ living in an $r = 2n+1$-dimensional manifold $\mathcal{M}$, and for a given closed $2n$-dimensional submanifold $\mathcal{S}$ of $\mathcal{M}$, one can associate the n-th Chern number $C_{2n}(\mathcal{S})$. It is defined as a $2n$-dimensional integral over the $2n$-form $\frac{i^n}{(2\pi)^n}\frac{1}{n!}\int_{\mathcal{S}} \mathrm{tr}(\Omega^n)$ over $\mathcal{S}$, where $\Omega = dA + A^2$ is the Berry-curvature 2-form given in terms of the nonabelian (matrix-valued) Berry-connection 1-form A. Let us also mention that the case of even r must be treated differently. The details of the mathematical formulation are too complicated to be discussed here. One may have a look at Refs. [3, 11–13].

3 Chern insulator

Our goal is to apply the concept of the Chern number as a topological invariant to a system of noninteracting electrons on a lattice. An additional Hubbard interaction is discussed later.

Generic model of a noninteracting insulator

A lattice model for noninteracting electrons is uniquely characterized by a hopping matrix $\boldsymbol{t}$ and has the Hamiltonian

$$H = \sum_{i,i'=1}^{L} \sum_{\alpha,\alpha'=1}^{M} t_{i\alpha,i'\alpha'}\, c^\dagger_{i\alpha} c_{i'\alpha'} \,. \tag{34}$$

Here $i = 1, ..., L$ labels the sites of a D-dimensional Bravais lattice of position vectors $\boldsymbol{r}_i$ with periodic boundary conditions, and $\alpha = 1, ..., M$ is an orbital index. For simplicity we consider spinless fermions. Furthermore, $M = 2$ different orbitals per site are necessary at least to allow for a band-insulating state with a gapped band structure. We choose $M = 2$ for simplicity.

We recap a few simple facts from band theory: The lattice is spanned by basis vectors $\boldsymbol{a}_r$ with $r = 1, ..., D$. Hence, $\boldsymbol{r}_i = \sum_{r=1}^{D} n_{ir}\boldsymbol{a}_r$ with $n_{ir} \in \mathbb{Z}$. The discrete Fourier transformation $c^\dagger_{\boldsymbol{k}\alpha} = L^{-1/2}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\boldsymbol{r}_i} c^\dagger_{i\alpha}$ block diagonalizes the Hamiltonian: $H = \sum_{\boldsymbol{k}\alpha\alpha'} t_{\alpha\alpha'}(\boldsymbol{k})\, c^\dagger_{\boldsymbol{k}\alpha} c_{\boldsymbol{k}\alpha'}$. Here, the wave vectors $\boldsymbol{k}$ lie in the unit cell of the reciprocal lattice spanned by the basis vectors $\boldsymbol{b}_s$ with $s = 1, ..., D$, defined as the unique solution of the inhomogeneous linear system of equations $\boldsymbol{a}_r\boldsymbol{b}_s = 2\pi\delta_{rs}$. Periodic boundary conditions for the real-space lattice imply that the wave-vector components are discrete: $k_s = (m_s/L_s)\boldsymbol{b}_s$ with $m_s = 0, ..., L_s{-}1$. There are $L = L_1 L_2 \cdots L_D$ wave vectors in any reciprocal unit cell. Typically, one chooses the first Brillouin zone (1BZ). Furthermore, we have $t_{\alpha\alpha'}(\boldsymbol{k}{+}\boldsymbol{g}) = t_{\alpha\alpha'}(\boldsymbol{k})$ for any reciprocal lattice vector $\boldsymbol{g} = \sum_{s=1}^{D} m_s\boldsymbol{b}_s$.

The last point leads to the important observation that *in the thermodynamic limit* $L \to \infty$ the 1BZ is actually a smooth D-dimensional manifold without a boundary, namely a D-dimensional torus $T^D \cong S^1 \times \cdots \times S^1$. This is nontrivial in the sense that the Poincaré lemma does not apply when integrating over the 1BZ. Furthermore, for each $\boldsymbol{k} \in$ 1BZ $\cong T^D$ the "Bloch Hamiltonian" $\boldsymbol{t}(\boldsymbol{k})$ is a $2{\times}2$ matrix (in general $M{\times}M$). We conclude that for $D = 2$ we have exactly the situation as discussed above in the toy-model example: A bundle of Hamiltonians with two-dimensional Hilbert spaces attached to the points of a smooth nontrivial parameter manifold 1BZ $\cong T^2$ (replacing the 2-sphere).

However, the present situation is much more pleasant: The different Bloch Hamiltonians in the bundle actually all derive from one and the same Hamiltonian of the system. Furthermore, the closed two-dimensional surface $\mathcal{S} =$ 1BZ, over which the integration is performed, is an intrinsic property of the system, namely of its geometry. There is no need to indicate the dependence of the Chern number $C(\mathcal{S}, \boldsymbol{\lambda})$ on the choice for $\mathcal{S}$. There is no choice, since $\mathcal{S}$ is given *a priori*. Hence, $C(\boldsymbol{\lambda})$ is a material property and depends on material (control) parameters only. It can be used to topologically classify the ground states in $\boldsymbol{\lambda}$ space.

For each $\boldsymbol{k} \in$ 1BZ we can expand the Bloch Hamiltonian in the basis of Hermitian $2{\times}2$ matrices as $\boldsymbol{t}(\boldsymbol{k}) = d_0(\boldsymbol{k})\mathbf{1} + \boldsymbol{d}(\boldsymbol{k}){\cdot}\boldsymbol{\tau}$. Actually, this form is quite convenient to *define* interesting models, rather than specifying the real-space hopping matrix $t_{i\alpha,i'\alpha'}$. As an example, we will consider the Qi-Wu-Zhang (QWZ) model [14, 2] on the square lattice. This is defined by $d_0(\boldsymbol{k}) = 0$ and

$$\boldsymbol{d}(\boldsymbol{k}) = \begin{pmatrix} t\sin k_x \\ t\sin k_y \\ m + t\cos k_x + t\cos k_y \end{pmatrix} . \tag{35}$$

Here, $t > 0$ is the hopping parameter, which can be fixed to specify the energy scale. Further, m is the so-called "mass" parameter. The eigenvalues of $\boldsymbol{t}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k}){\cdot}\boldsymbol{\tau}$ are given by $\epsilon_\pm(\boldsymbol{k}) = \pm|\boldsymbol{d}(\boldsymbol{k})| = \pm t\sqrt{\sin^2 k_x + \sin^2 k_y + \big(m/t{+}\cos k_x{+}\cos k_y\big)^2}$. We see that there is a (fully occupied) valence band and an (unoccupied) conduction band separated by a finite

gap $\Delta(\boldsymbol{k}) = 2|\boldsymbol{d}(\boldsymbol{k})| > 0$ in the entire 1BZ ($-\pi < k_x, k_y \leq \pi$), except for special values of the mass m, for which $|\boldsymbol{d}(\boldsymbol{k})| = 0$. We immediately see that a gap closure requires $\boldsymbol{k} = \boldsymbol{k}_c$ with $\boldsymbol{k}_c \in \big\{(0,0),\, (0,\pi),\, (\pi,0),\, (\pi,\pi)\big\}$, i.e., it can take place at critical wave vectors $\boldsymbol{k}_c$ given by the high-symmetry points (HSPs) in the 1BZ. Furthermore, from the condition $m/t + \cos k_x + \cos k_y = 0$, evaluated at the HSPs, we find the critical mass parameters $m_c = -2t,\, 0,\, 2t$, at which there is a band closure (at $\boldsymbol{k}_c = (0,0)$, $\boldsymbol{k}_c = (\pi,0),\, (0,\pi)$, and $\boldsymbol{k}_c = (\pi,\pi)$, respectively).

Band and Chern insulators

Our toy model Eq. (16) is recovered with the substitution $\boldsymbol{d} \leftrightarrow -\frac{1}{2}\boldsymbol{R}$. Hence, we immediately get the Berry curvature (note that ground and excited states switch their roles and that the prefactor $\frac{1}{2}$ is the same, see Eq. (23))

$$\boldsymbol{\Omega}_0(\boldsymbol{d}) = \frac{1}{2}\frac{\boldsymbol{d}}{d^3}\,. \tag{36}$$

At least this is the Berry curvature when regarding $\boldsymbol{d} = (d_x, d_y, d_z)^T$ as the parameters. One would then have to choose, more or less arbitrarily, a closed two-dimensional surface embedded in three-dimensional $\boldsymbol{d}$ space to compute the Chern number. The *physical* parameter manifold $\mathcal{S}$, however, is the 1BZ. This yields the Chern number as a material property. If the resulting Chern number is nonzero, we refer to such a material as a "Chern insulator", as opposed to a "trivial" band insulator with $C = 0$.

For noncritical mass parameters, Eq. (35) defines a smooth map $\boldsymbol{d} : \mathcal{S} \to \mathbb{R}^3\backslash\{0\}$. Since $\mathcal{S} = \text{1BZ} \cong T^2$ is a closed two-dimensional surface, its image under $\boldsymbol{d}$,

$$\mathcal{D} \equiv \big\{\boldsymbol{d} = \boldsymbol{d}(\boldsymbol{k}) \in \mathbb{R}^3\backslash\{0\} \;\big|\; \boldsymbol{k} = (k_x, k_y) \in \text{1BZ}\big\}\,, \tag{37}$$

is a closed two-dimensional surface as well, which is embedded in $\mathbb{R}^3\backslash\{0\}$. We can define its infinitesimal vectorial surface element as $d\boldsymbol{S} = \frac{\partial \boldsymbol{d}(\boldsymbol{k})}{\partial k_x} \times \frac{\partial \boldsymbol{d}(\boldsymbol{k})}{\partial k_y}\, dk_x dk_y$, which is perpendicular to $\mathcal{D}$. This can be used, together with the substitution rule for functions of several variables, to compute the Chern number as

$$C = \frac{1}{2\pi} \oint_{\mathcal{D}} \left(\frac{1}{2}\frac{\boldsymbol{d}}{d^3}\right) d\boldsymbol{S} = \frac{1}{4\pi} \oint_{\mathcal{S}} \frac{\boldsymbol{d}(\boldsymbol{k})}{d(\boldsymbol{k})^3} \cdot \frac{\partial \boldsymbol{d}(\boldsymbol{k})}{\partial k_x} \times \frac{\partial \boldsymbol{d}(\boldsymbol{k})}{\partial k_y}\, dk_x dk_y\,. \tag{38}$$

The two-dimensional integral in Eq. (38) with the QWZ $\boldsymbol{d}(\boldsymbol{k})$ from Eq. (35) is easily evaluated numerically, and one can verify in this way that $C \in \mathbb{Z}$.

In fact, $C = 0$ for $-\infty < m < -2t$, $C = +1$ for $-2t < m < 0$, $C = -1$ for $0 < m < 2t$, and $C = 0$ for $2t < m < \infty$. This provides us with a topological phase diagram. Systems in the phase diagram with different Chern numbers cannot be deformed continuously into each other. Upon varying m, one necessarily has to pass a critical value m_c, where there is a band closure at a HSP in the 1BZ and thus a degenerate ground state.

Wrapping number

Now that we have the topological phase diagram, we may consider continuous deformations of the Hamiltonian that do not lead to a gap closure and thus ensure that the Chern number does not change. For example, we may treat the modulus of $\boldsymbol{d}(\boldsymbol{k})$ as a $\boldsymbol{k}$-dependent control parameter. A continuous deformation $d(\boldsymbol{k}) = |\boldsymbol{d}(\boldsymbol{k})| \mapsto 1$ does not close the gap $\Delta(\boldsymbol{k}) = 2d(\boldsymbol{k})$. Hence, the Hamiltonian with normalized $d(\boldsymbol{k}) = 1$ has the same Chern number as the original one.
This deformation is advantageous, since this implies $\partial|d(\boldsymbol{k})|^2/\partial k_x = 0$ (and the same for k_y), and thus $\boldsymbol{d}(\boldsymbol{k})$ becomes perpendicular to $\partial\boldsymbol{d}(\boldsymbol{k})/\partial k_x$, and to $\partial\boldsymbol{d}(\boldsymbol{k})/\partial k_y$ as well. Therewith, $\boldsymbol{d}(\boldsymbol{k})$ becomes parallel to the surface element $d\boldsymbol{S}$, and $\boldsymbol{d}(\boldsymbol{k})d\boldsymbol{S} = \pm dS$ with $dS > 0$. Eq. (38) then reads

$$C = \pm\frac{1}{4\pi}\oint_{\mathcal{D}} dS\,, \tag{39}$$

where $\mathcal{D}$ is the image of the 1BZ $\cong T^2$ under the continuous map $T^2 \to \mathbb{R}^3\backslash\{0\}$ with $\boldsymbol{k} \mapsto \boldsymbol{d}(\boldsymbol{k})$, see Eq. (37).
Since the 1BZ $\cong T^2$ is a closed two-dimensional surface and since $d(\boldsymbol{k}) = 1$, the image of T^2 is a closed two-dimensional surface embedded in the 2-sphere S^2. This implies that the image $\mathcal{D}$ must cover the entire S^2 — once or several times. The Chern number is just the area of this surface, divided by 4π. Since the surface area of S^2 is just 4π, we see that the Chern number is just the count, how often S^2 is covered. Its sign reflects the orientation of the image with its surface normal on S^2 pointing inwards or outwards. The Chern number gives us the information how often T^2 wraps around S^2. It can be interpreted as a *wrapping number*.

Time-reversal symmetry

It is important to understand that a nonzero Chern number requires that time-reversal symmetry is broken. The identity and time reversal form a group of transformations isomorphic to $\mathbb{Z}_2$. According to Wigner's theorem [8], time reversal is represented in Hilbert space by a unitary or anti-unitary operator $\mathcal{K}$. In single-particle quantum mechanics, it is defined via its action of the position, momentum and spin operator as $\mathcal{K}\boldsymbol{r}\mathcal{K}^\dagger = \boldsymbol{r}$, $\mathcal{K}\boldsymbol{p}\mathcal{K}^\dagger = -\boldsymbol{p}$, and $\mathcal{K}\boldsymbol{s}\mathcal{K}^\dagger = -\boldsymbol{s}$. This implies $\mathcal{K}[x, p_x]\mathcal{K}^\dagger = -[x, p_x]$, and since $[x, p_x] = i\hbar$ we must have $\mathcal{K}i\hbar\mathcal{K}^\dagger = -i\hbar$. This means $\mathcal{K}$ is anti-unitary, i.e., anti-linear and unitary with $\mathcal{K}^\dagger\mathcal{K} = 1$. For a many-body system with an even (odd) number of spin-1/2 fermions we have $\mathcal{K}^2 = 1$ ($\mathcal{K}^2 = -1$). Here, we consider spinless fermions, where $\mathcal{K}^2 = 1$. On the Fock space of a spinless-fermion model, the anti-unitary operator $\mathcal{K}$ can be defined to act trivially on the annihilation and creation operators with respect to the site-orbital basis, $\mathcal{K}c_{i\alpha}\mathcal{K}^\dagger = c_{i\alpha}$ (and the same for $c^\dagger_{i\alpha}$). Since $c^\dagger_{\boldsymbol{k}\alpha} = L^{-1/2}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\boldsymbol{r}_i}c^\dagger_{i\alpha}$, anti-linearity of $\mathcal{K}$ implies $\mathcal{K}c_{\boldsymbol{k}\alpha}\mathcal{K}^\dagger = c_{-\boldsymbol{k}\alpha}$.
Let us assume that the parameter-dependent Hamiltonian $H(\boldsymbol{R})$ of a spinless-fermion system is time-reversal symmetric for all $\boldsymbol{R}$, i.e., $H(\boldsymbol{R}) = \mathcal{K}H(\boldsymbol{R})\mathcal{K}^\dagger$. Hence, we may assume that the eigenstates of $H(\boldsymbol{R})$ can be chosen among the eigenstates of $\mathcal{K}$. In fact, we may choose the phases of $\big|\Phi_j(\boldsymbol{R})\big\rangle$ such that $\mathcal{K}\big|\Phi_j(\boldsymbol{R})\big\rangle = +\big|\Phi_j(\boldsymbol{R})\big\rangle$. These relations can be used to analyze

the matrix element in the representation Eq. (23) of the Berry curvature. We have

$$\begin{aligned}\big\langle \Phi_0(\boldsymbol{R})\big|\partial_\beta H(\boldsymbol{R})\big|\Phi_j(\boldsymbol{R})\big\rangle &= \big\langle \mathcal{K}\Phi_0(\boldsymbol{R})\big|\partial_\beta(\mathcal{K}H(\boldsymbol{R})\mathcal{K}^\dagger)\big|\mathcal{K}\Phi_j(\boldsymbol{R})\big\rangle \\ &= \big\langle \Phi_0(\boldsymbol{R})\big|\mathcal{K}^\dagger\mathcal{K}\partial_\beta H(\boldsymbol{R})\mathcal{K}^\dagger\mathcal{K}\big|\Phi_j(\boldsymbol{R})\big\rangle^* \\ &= \big\langle \Phi_0(\boldsymbol{R})\big|\partial_\beta H(\boldsymbol{R})\big|\Phi_j(\boldsymbol{R})\big\rangle^* , \end{aligned} \tag{40}$$

where we have used that $\langle \mathcal{A}u|v\rangle = \langle u|\mathcal{A}^\dagger v\rangle^*$ for any anti-linear operator $\mathcal{A}$ and arbitrary states $|u\rangle, |v\rangle$. Hence, the matrix element is real. With Eq. (23) we conclude that the Berry curvature vanishes for any time-reversal symmetric Hamiltonian. The Chern number is zero.
Consider a multi-orbital spinless-fermion model, where $H = \sum_{\boldsymbol{k}\alpha\alpha'} t_{\alpha\alpha'}(\boldsymbol{k}) c^\dagger_{\boldsymbol{k}\alpha} c_{\boldsymbol{k}\alpha'}$ is specified by the Bloch Hamiltonian $\boldsymbol{t}(\boldsymbol{k})$. We have

$$\mathcal{K}H\mathcal{K}^\dagger = \mathcal{K}\sum_{\boldsymbol{k}\alpha\alpha'} t_{\alpha\alpha'}(\boldsymbol{k})\, c^\dagger_{\boldsymbol{k}\alpha} c_{\boldsymbol{k}\alpha'}\mathcal{K}^\dagger = \sum_{\boldsymbol{k}\alpha\alpha'} t^*_{\alpha\alpha'}(\boldsymbol{k})\, c^\dagger_{-\boldsymbol{k}\alpha} c_{-\boldsymbol{k}\alpha'} = \sum_{\boldsymbol{k}\alpha\alpha'} t^*_{\alpha\alpha'}(-\boldsymbol{k})\, c^\dagger_{\boldsymbol{k}\alpha} c_{\boldsymbol{k}\alpha'} , \tag{41}$$

and thus H is invariant under time reversal, $\mathcal{K}H\mathcal{K}^\dagger = H$, if $\boldsymbol{t}(\boldsymbol{k}) = \boldsymbol{t}^*(-\boldsymbol{k})$. Specifically, for the QWZ model $\boldsymbol{t}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\boldsymbol{\tau}$ with $\boldsymbol{d}(\boldsymbol{k})$ given by Eq. (35). We have

$$\boldsymbol{t}^*(-\boldsymbol{k}) = \boldsymbol{d}(-\boldsymbol{k})\boldsymbol{\tau}^* = \big(-d_x(\boldsymbol{k}), -d_y(\boldsymbol{k}), +d_z(\boldsymbol{k})\big)\cdot\big(\tau_x, -\tau_y, \tau_z\big) , \tag{42}$$

showing that time-reversal symmetry is broken. Thus, a topologically nontrivial phase of the model signalled by a nonzero Chern number is not excluded.

The quantum Hall effect

Typically, a topological invariant is a rather abstract, mathematical concept, and its only significance is to discriminate topologically different phases that cannot be deformed into each other continuously. Sometimes, on top of this important classification property, a topological invariant can be an easily accessible physical observable. The Chern number of a two-dimensional Chern insulator is exactly such a case, and it plays the central role in the quantum Hall effect.
The quantum Hall effect (QHE) is observed in a two-dimensional electron gas, which can be realized, e.g., in gallium-arsenide heterostructures. Applying a magnetic field perpendicular to the heterostructure, the transverse (Hall) resistivity R_{xy} can be measured as function of the magnetic field strength. From the standard theory of electromagnetism and from Drude theory one expects a linear dependence. At sufficiently low temperatures, however, plateaus in the field dependence are found [15]. In fact, the zero-temperature Hall conductivity is quantized: $1/R_{xy} = \nu e^2/h$ with integers ν characteristic for the Hall plateaus. This striking quantization effect has been employed for a new practical standard for electrical resistance as it comes with an extremely high accuracy. A theoretical derivation [16] shows that the integer ν is given by the first Chern number. This means that the quantum Hall effect is eventually traced back to topological properties of the two-dimensional electron gas.
The QWZ model must be seen as a simplified model related to the quantum anomalous Hall effect (QAHE), i.e., the quantum version of the anomalous Hall effect. The latter requires a combination of magnetic polarization and spin-orbit coupling to generate a Hall effect rather than an external magnetic field. In fact, broken time-reversal symmetry rather than an orbital magnetic field and Landau levels is more fundamental for the concept of a Chern insulator [17].

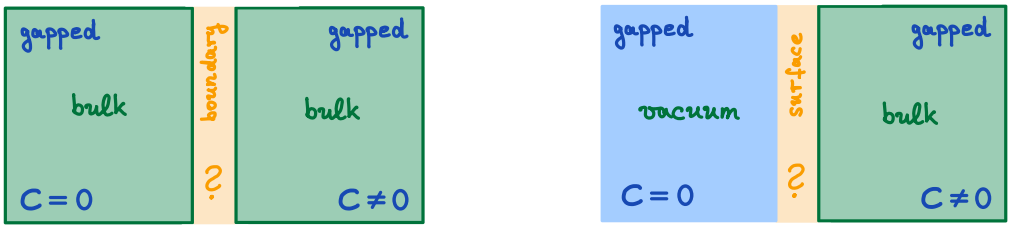

Fig. 4: *Gapless modes must exist at the boundary between two bulk systems with different Chern numbers or at the surface of a Chern insulator.*

Bulk-boundary correspondence

Another important and experimentally observable consequence of the existence of the Chern number as a topological invariant in systems with broken time-reversal symmetry is the necessary presence of gapless modes at the *surfaces* of a Chern insulator. The finite and quantized conductivity in the quantum-Hall systems is in fact carried by these boundary modes. Their number is exactly given by the Chern number, and each carries a flux quantum e^2/h.
The existence of gapless boundary modes can be explained with the so-called bulk-boundary correspondence. This is a rather deep concept in topology but can be sketched in simple terms as follows:
Consider a D-dimensional system that is spatially cut into two subsystems along a $D-1$-dimensional boundary, see Fig. 4. For each of the two D-dimensional subsystems, the bulk band structure shall be gapped and topologically characterized by a Chern number. We assume that the two Chern numbers are different and imagine an extended boundary which smoothly interpolates between the bulk Hamiltonians of the two subsystems. Then, somewhere along a path starting in the bulk of the first subsystem through the boundary and ending in the bulk of the second, the energy gap must vanish. Otherwise it would be impossible for the topological invariant to change. Hence, we conclude that at the boundary between two insulators with different Chern numbers, or, since the vacuum has vanishing Chern number, at the surface of a Chern insulator, there must be low-energy electronic states crossing the gap. These boundary modes have the fascinating and experimentally detectable property of being topologically protected. Their existence is a fundamental consequence of a gapped and topologically nontrivial band structure and thus not affected by arbitrary local perturbations at the boundary.

4 Electron correlations

Up to this point we have concentrated on the topological properties of noninteracting lattice-fermion models. We will now shift the focus to interacting models, such as the QWZ model with an additional Hubbard interaction as an instructive prototype and example for concrete calculations. To this end, we first define the *spinful* QWZ model with Hamiltonian $H^{(0)} = \sum_{\boldsymbol{k}\alpha\alpha'\sigma} t_{\alpha\alpha'}(\boldsymbol{k})\, c^\dagger_{\boldsymbol{k}\alpha\sigma} c_{\boldsymbol{k}\alpha'\sigma}$, where $\sigma = \uparrow, \downarrow$ denotes the spin projection and where $\boldsymbol{t}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\tau}$ with $\boldsymbol{d}(\boldsymbol{k})$ given by Eq. (35). The QWZ+U model is then given by $H = H^{(0)} + H^{(1)}$ where the (intra-orbital) Hubbard term reads as $H^{(1)} = U \sum_{i\alpha} n_{i\alpha\uparrow} n_{i\alpha\downarrow}$ with $n_{i\alpha\sigma} = c^\dagger_{i\alpha\sigma} c_{i\alpha\sigma}$.

Generally, a lattice-fermion model is defined via its hopping matrix $t_{i\alpha,i'\alpha'}$ (see Eq. (34)) or by the corresponding Bloch Hamiltonian $\boldsymbol{t}(\boldsymbol{k})$ that is obtained after Fourier transformation to reciprocal space and can trivially be made spinful as above (if this is not already the case) and equipped with a Hubbard or with the full Coulomb interaction. We will restrict ourselves to local interactions, such as the Hubbard interaction, since in those cases one can straightforwardly apply dynamical mean-field theory (DMFT).

Dynamical mean-field theory

For the following discussion, we recall some important facts related to DMFT. An overview is given by Ref. [18].
(i) DMFT is an approximate mean-field theory for D-dimensional interacting lattice-fermion models with local interactions. As opposed to other mean-field approaches, it becomes exact in the limit of infinite spatial dimensions $D \to \infty$. This requires a proper scaling of the hopping-matrix elements with D. In case of the hypercubic lattice, for example, the nearest-neighbor hopping must be scaled as $t = t^*/\sqrt{D}$ to maintain the balance between the noninteracting and the interacting parts, $H^{(0)}$ and $H^{(1)}$, of the Hamiltonian in the $D \to \infty$ limit.
(ii) The central quantity of DMFT is the single-particle Green function $\boldsymbol{G}(\boldsymbol{k},\omega)$. In the noninteracting case, its elements are given by

$$G^{(0)}_{\alpha\alpha'}(\boldsymbol{k},\omega) = \left(\frac{1}{\omega+\mu-\boldsymbol{t}(\boldsymbol{k})}\right)_{\alpha\alpha'}, \tag{43}$$

assuming a spin-diagonal and spin-independent hopping. Here, μ is the chemical potential. The interacting Green function is obtained from $\boldsymbol{G}^{(0)}(\boldsymbol{k},\omega)$ in terms of the local, i.e., $\boldsymbol{k}$-independent self-energy $\boldsymbol{\Sigma}(\omega)$ as

$$\boldsymbol{G}(\boldsymbol{k},\omega) = \frac{1}{\left(\boldsymbol{G}^{(0)}(\boldsymbol{k},\omega)\right)^{-1} - \boldsymbol{\Sigma}(\omega)}\,. \tag{44}$$

Locality of the self-energy is the decisive assumption of DMFT and is an exact property in the limit $D \to \infty$.
(iii) Operationally, DMFT self-consistently maps the lattice model onto an Anderson impurity model with the same local interaction term. The self-consistency cycle may be started with a guess for $\boldsymbol{\Sigma}(\omega)$. Via Eq. (44) this fixes $\boldsymbol{G}(\boldsymbol{k},\omega)$ and the local Green function $\boldsymbol{G}_{\text{loc}}(\omega) = \frac{1}{L}\sum_{\boldsymbol{k}} \boldsymbol{G}(\boldsymbol{k},\omega)$ in particular. The latter is used to define the noninteracting Green function matrix $\mathcal{G}(\omega)$ of a (multi-impurity) Anderson impurity model as

$$\mathcal{G}_{\alpha\alpha'}(\omega) = \left(\frac{1}{\left(\boldsymbol{G}_{\text{loc}}(\omega)\right)^{-1} + \boldsymbol{\Sigma}(\omega)}\right)_{\alpha\alpha'}, \tag{45}$$

and thereby fixes its noninteracting Hamiltonian $H^{(0)}_{\text{imp}}$ or action. The interacting impurity problem $H_{\text{imp}} = H^{(0)}_{\text{imp}} + H^{(1)}$ must be solved for the same interaction, chemical potential and temperature T (here assumed as $T = 0$) with the goal to get an updated self-energy that can be fed back to the start of the cycle. The DMFT solution of the lattice problem is obtained by iterating the cycle until self-consistency is reached. Clearly, an "impurity solver" is needed to run the algorithm (see Ref. [18], for examples).

Topological Hamiltonian

A noninteracting electron system on a lattice is specified by the Bloch Hamiltonian $\boldsymbol{t}(\boldsymbol{k})$. We assume that (i) $\boldsymbol{t}(\boldsymbol{k})$ is gapped, i.e., that there is an energy μ_0 such that, *for all* $\boldsymbol{k} \in$ 1BZ, its eigenvalues $\varepsilon_m(\boldsymbol{k}) < \mu_0$ if $m = 1, ..., m_0$ ("occupied bands") and $\varepsilon_m(\boldsymbol{k}) > \mu_0$ if $m = m_0{+}1, ..., M$ ("unoccupied bands"). We further assume that (ii) the total electron number is $N = 2m_0L$. Then, the state $|\chi_0(\boldsymbol{k})\rangle = \prod_{m\le m_0}\prod_\sigma c^\dagger_{\boldsymbol{k}m\sigma}|\text{vac.}\rangle$ can be seen as the ground state of $H(\boldsymbol{k}) = \sum_{m\sigma}\varepsilon_m(\boldsymbol{k})\, c^\dagger_{\boldsymbol{k}m\sigma}c_{\boldsymbol{k}m\sigma} = \sum_{\alpha\alpha'\sigma} t_{\alpha\alpha'}(\boldsymbol{k})c^\dagger_{\boldsymbol{k}\alpha\sigma}c_{\boldsymbol{k}\alpha'\sigma}$ at a given $\boldsymbol{k} \in$ 1BZ in the invariant subspace with $2m_0$ electrons. The ground state of $H^{(0)}$ is $|\Phi_0\rangle = \prod_{\boldsymbol{k}} |\chi_0(\boldsymbol{k})\rangle$.
Apart from the trivial spin degeneracy which can be disregarded here, we have a Hamiltonian $H^{(0)}(\boldsymbol{k})$ with a smooth dependence on the parameters of a two-dimensional closed manifold (the 1BZ) with a unique, i.e., nondegenerate ground state for all $\boldsymbol{k}$ and thus a gap to the lowest excited state. As discussed in the preceding section, we can thus define the corresponding Berry curvature and infer that its integral over the parameter manifold is given by 2π times an integer, the Chern number.
Within DMFT, this argument holds true if we replace $\boldsymbol{t}(\boldsymbol{k})$ by $\boldsymbol{t}_{\text{top}}(\boldsymbol{k}) = \boldsymbol{t}(\boldsymbol{k}) + \boldsymbol{\Sigma}(\omega{=}0)$, *provided that* $\boldsymbol{t}_{\text{top}}(\boldsymbol{k})$ is still gapped and that $\boldsymbol{\Sigma}(\omega)$ does not diverge as $\omega \to 0$. The bundle of matrices $\boldsymbol{t}_{\text{top}}(\boldsymbol{k})$ is called the "topological Hamiltonian" [19]. Note that the self-energy at zero frequency $\boldsymbol{\Sigma}(\omega{=}0)$ is a Hermitian matrix in the orbital indices as can be deduced from the fundamental properties of the single-particle Green function and Dyson's equation (44). We also note that the Slater determinants $|\chi_0(\boldsymbol{k})\rangle$ do not have a direct physical meaning in case of an interacting electron system and must be seen as auxiliary quantities. Nevertheless, the topological Hamiltonian provides us with a means for computing an "interacting Chern number" and thereby with a topological classification of gapped interacting electron systems on a lattice.
We will proceed with the concept of topological Hamiltonian in the context of DMFT. However, in the extended context of interacting topological insulators there are some partly open questions. However, we will postpone their discussion to the final section.

5 Exact interplay of correlations and topology in $D = \infty$

Among the various mean-field theories, DMFT has an exceptional standing, since it is internally consistent and nonperturbative, and since there is a nontrivial limit, in which this mean-field approach becomes exact [20]. This is the limit of infinite spatial dimensions. For the Hubbard model $H = \sum_{ii'\sigma} t_{ii'}\, c^\dagger_{i\sigma}c_{i'\sigma} + U\sum_i n_{i\uparrow}n_{i\downarrow}$ on a D-dimensional hypercubic lattice, one has to properly scale the nearest-neighbor hopping amplitudes to keep a nontrivial balance between the tight-binding and the Hubbard-interaction term in the limit $D \to \infty$, namely $t = t^*/\sqrt{D}$ with fixed t^* (e.g. $t^*{=}1$). In this limit, DMFT provides the exact solution and, while one can profit from certain simplifications such as the locality of the self-energy [21], the remaining physics is highly nontrivial. This is demonstrated, for example, with the famous phase diagram for the Mott metal-insulator transition [18]. Here, we consider the $D \to \infty$ limit of a multi-orbital

Hubbard model, which already for $U = 0$ has a nontrivial topological phase diagram including band and Chern insulator phases. Our goal is to construct the model such that its correlations *and* its topological properties remain nontrivial, such that we are able to study, *on a numerically exact level*, the topological properties of a paradigmatic strongly interacting electron system.

Spinful QWZ model on a D-dimensional hypercubic lattice

To this end, we start from the noninteracting QWZ model, for which an extension to arbitrary even dimensions $D = 2, 4, 6, ...$ is known [12], namely $H^{(0)} = \sum_{k\alpha\alpha'\sigma} t_{\alpha\alpha'}(\boldsymbol{k})\, c^\dagger_{\boldsymbol{k}\alpha\sigma} c_{\boldsymbol{k}\alpha'\sigma}$, where $\alpha = 1, ..., M$ is the orbital index and $t_{\alpha\alpha'}(\boldsymbol{k})$ are the elements of the spin-independent Bloch Hamiltonian

$$\boldsymbol{t}(\boldsymbol{k}) = \left(m + t\sum_{r=1}^{D} \cos k_r\right) \boldsymbol{\gamma}_D^{(0)} + t\sum_{r=1}^{D} \sin k_r\, \boldsymbol{\gamma}_D^{(r)} \,. \tag{46}$$

Here, $t = t^*/\sqrt{D}$ is the hopping and m the mass parameter. We set t^*=1 to fix the energy scale. This Bloch Hamiltonian is an $M\times M$ matrix for each wave vector $\boldsymbol{k} = (k_1, ..., k_D) \in$ 1BZ with $-\pi < k_r \leq \pi$ and is constructed with the help of the $M\times M$ matrices $\gamma_D^{(\mu)}$, where $\mu = 0$ or $\mu = r = 1, ..., D$. These satisfy the anticommutation relations of a Clifford algebra

$$\left\{\boldsymbol{\gamma}_D^{(\mu)},\, \boldsymbol{\gamma}_D^{(\nu)}\right\} = 2\delta^{(\mu\nu)} \,. \tag{47}$$

This algebra is a very convenient one, e.g., for the computation of the eigenvalues of $\boldsymbol{t}(\boldsymbol{k})$. We have $\boldsymbol{t}(\boldsymbol{k}) = \sum_{\mu=0}^{M} d_\mu(\boldsymbol{k})\boldsymbol{\gamma}^{(\mu)}$ with $d_0(\boldsymbol{k}) \equiv m + t\sum_{r=1}^{D} \cos k_r$ and $d_r(\boldsymbol{k}) \equiv t \sin k_r$ for $r = 1, ..., D$. Using Eq. (47) we immediately get

$$\boldsymbol{t}(\boldsymbol{k})^2 = \sum_\mu d_\mu(\boldsymbol{k})\boldsymbol{\gamma}^{(\mu)} \sum_\nu d_\nu(\boldsymbol{k})\boldsymbol{\gamma}^{(\nu)} = \frac{1}{2}\sum_{\mu\nu} d_\mu(\boldsymbol{k}) d_\nu(\boldsymbol{k}) \left(\boldsymbol{\gamma}^{(\mu)}\boldsymbol{\gamma}^{(\nu)} + \boldsymbol{\gamma}^{(\nu)}\boldsymbol{\gamma}^{(\mu)}\right) = \sum_\mu d_\mu(\boldsymbol{k})^2 \,. \tag{48}$$

This means that there are two $M/2$-fold degenerate bands with dispersions given by $\epsilon_\pm(\boldsymbol{k}) = \pm\left(d_0(\boldsymbol{k})^2 + \sum_r d_r(\boldsymbol{k})^2\right)^{1/2}$, when disregarding the spin degree of freedom. Note that, since the γ matrices are traceless, $\boldsymbol{t}(\boldsymbol{k})$ is traceless as well and, therefore, there must be as many negative as positive eigenvalues.

The orbital-dependent density of states (DOS) is given in terms of the free retarded Green function as

$$\rho_\alpha(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,\frac{1}{L}\sum_{\boldsymbol{k}} G^{(0)}_{\alpha\alpha}(\boldsymbol{k}, \omega + i0^+) \,. \tag{49}$$

Again, thanks to the Clifford algebra, we can easily compute

$$G^{(0)}_{\alpha\alpha}(\boldsymbol{k}, \omega) = \left(\frac{1}{\omega + \mu^{(0)} - \boldsymbol{\epsilon}(\boldsymbol{k})}\right)_{\alpha\alpha} = \left(\frac{1}{\omega - \sum_\mu d_\mu(\boldsymbol{k})\boldsymbol{\gamma}_D^\mu}\right)_{\alpha\alpha} = \left(\frac{\omega + \sum_\mu d_\mu(\boldsymbol{k})\boldsymbol{\gamma}_D^\mu}{\omega^2 - \sum_\mu d_\mu(\boldsymbol{k})^2}\right)_{\alpha\alpha} . \tag{50}$$

We consider a half-filled system and have set the chemical potential of the free system to zero. For $D = 2$, one obtains the QWZ model discussed Sec. 3 when choosing $\boldsymbol{\gamma}_2^{(1)} = \boldsymbol{\tau}_x$ and $\boldsymbol{\gamma}_2^{(2)} = \boldsymbol{\tau}_y$, and the so-called chiral element $\boldsymbol{\gamma}_2^{(0)} = -i\boldsymbol{\tau}_x\boldsymbol{\tau}_y = \boldsymbol{\tau}_z$ for the γ matrices. Note that they satisfy the anticommutation relations Eq. (47), i.e., the Pauli matrices form a two-dimensional representation of the algebra, $M = 2$. We have $\boldsymbol{t}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\cdot\tau$ with $\boldsymbol{d}(\boldsymbol{k})$ given by Eq. (35).

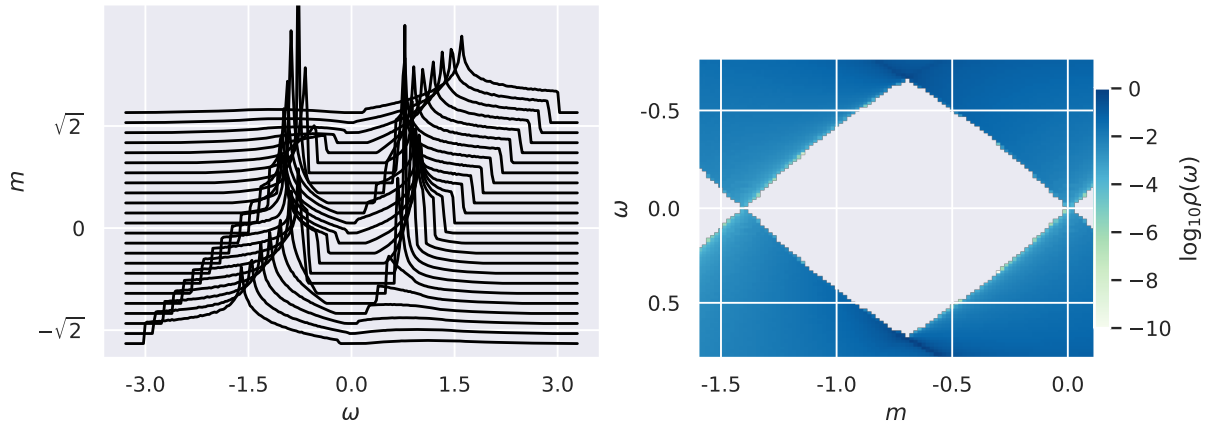

Fig. 5: *Density of states on the A-orbitals for the $D = 2$ QWZ model for different m (left). Right: Density of states (color code) at low excitation energies. $t = t^*/\sqrt{2}$ with $t^* = 1$ sets the energy scale. Figure taken from Ref. [5].*

The matrices $\boldsymbol{\gamma}^{(r)}$ have vanishing diagonal elements, such that there is a contribution from $\boldsymbol{\gamma}^{(0)}$ only. We have $\gamma^{(0)}_{\alpha\alpha} \equiv z_\alpha{=}{+}1$ for "A orbitals" $\alpha = 1, ..., M/2$ and $\gamma^{(0)}_{\alpha\alpha} \equiv z_\alpha{=}{-}1$ for "B orbitals" $\alpha = (M/2){+}1, ..., M$. This holds for general D and M (for the $D = 2$ QWZ model: $M = 2$) and implies that the orbital-resolved free Green functions $G^{(0)}_{\alpha\alpha}(\boldsymbol{k}, \omega)$ for $\alpha = 1, ..., M$ can be divided into two classes with representatives $G^{(0)}_A(\boldsymbol{k}, \omega)$ and $G^{(0)}_B(\boldsymbol{k}, \omega)$.
The local DOS on the A orbitals is displayed in Fig. 5 (left) as a function of the mass parameter m. We note the symmetry $\rho_A(\omega) \to \rho_A(-\omega)$ for $m \to -m$. Furthermore, we have $\rho_A(-\omega) = \rho_B(\omega)$. In the right plot for small frequencies, the evolution of the gap with m can be read off. Gap closures are found at $m = 0$ and $m = \pm\sqrt{2}$ (in units of $t^* = 1$, i.e., $t = 1/\sqrt{2}$), as discussed earlier.

Bulk-boundary correspondence

For $D = 2$ dimensions and for mass parameter $m = 1$ the system is a Chern insulator with $C = -1$. Hence, according to the bulk-boundary correspondence, there must be a gap closure at the surfaces of the bulk system. This is uncovered, e.g., by a calculation for the system in a ribbon geometry, i.e., we assume that the system is infinitely extended (with periodic boundary conditions) is x direction, while in y direction it consists of a finite number d of one-dimensional chains labelled by a chain index $\lambda = 1, ..., d$. As the vacuum (for $\lambda < 1$ and $\lambda > d$) has Chern number $C = 0$, we expect gapless edge modes localized at the two edges.
For the actual calculation we can only profit from translational symmetry in x direction. We perform a two-dimensional Fourier transformation of the bulk Bloch Hamiltonian, Eq. (46), or specifically, for $D = 2$, of $\boldsymbol{t}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\tau}$ with $\boldsymbol{d}(\boldsymbol{k})$ given by Eq. (35) to real space, cut the hopping parameters in Eq. (34) at the two edges, and perform a one-dimensional Fourier transform from x space back to k_x space. The result is a $d{\times}d$ matrix for each wave "vector" k_x, which must be diagonalized numerically. The d eigenvalues as functions of k_x represent the band structure of the QWZ model in a ribbon geometry.

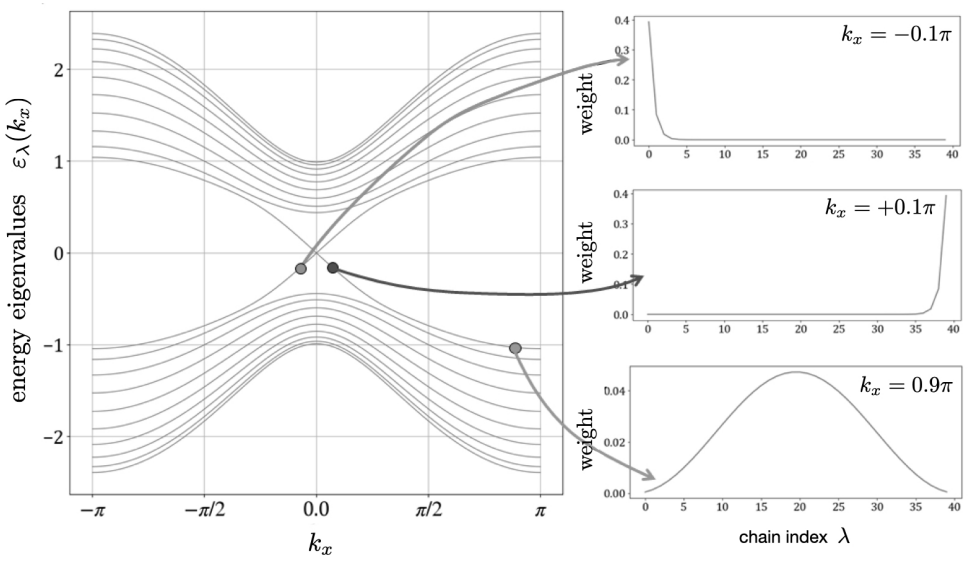

Fig. 6: *Dispersions of the $D = 2$ QWZ model in a ribbon geometry at $m = 1$ (left). Calculations for $d = 10$ chains of infinite length in x direction with edges at $\lambda = 1$ and $\lambda = 10$ in y direction. Right: Weight of various eigenstates in the different chains $\lambda = 1, ..., d$ of a ribbon with thickness $d = 40$, as indicated. Calculations for $t = t^*/\sqrt{2}$ with t^*=1. Figure taken from Ref. [6].*

They are displayed in Fig. 6 (left) for a ribbon thickness of $d = 10$. In fact, we find two low-energy eigenmodes visible around $k_x = 0$, which split off the "bulk continuum", which is here represented by 8 bands only, at wave vectors $k_x \approx \pm 0.36\pi$. Each of these modes bridges the bulk band gap. Furthermore, the analysis of the corresponding eigenstates in the right part of Fig. 6 shows that one edge mode is localized at the top edge $\lambda = 1$ and the other one at the opposite edge $\lambda = 40$. Note that here the calculations have been done for a thicker ribbon with $d = 40$. The corresponding weight factors $|\langle i, \alpha | k_x, \lambda \rangle|^2$ decay exponentially with increasing distance to the edges, opposed to bulk states which extend over the entire ribbon. Let us emphasize once more, that these two gapless edge modes must necessarily exist and that they are topologically protected against local perturbations.

Topological phase diagram of the $D = 2$ QWZ+U model

We now add a local Hubbard-type interaction $H^{(1)} = (U/2) \sum_{i\alpha\sigma} n_{i\alpha\sigma} n_{i\alpha-\sigma}$ to the Hamiltonian and consider the resulting two-dimensional QWZ+U model at half filling. Within the DMFT, the self-energy is site-diagonal, i.e., k-independent. Furthermore, as the $H^{(1)}$ term is an intra-orbital interaction only, it is diagonal in orbital space, $\Sigma_{\alpha\beta}(k, \omega) = \Sigma_{\alpha\beta}(\omega) = \delta_{\alpha\beta}\Sigma_\alpha(\omega)$. Analogous to the discussion of the density of states above, the orbital-dependent diagonal elements $\Sigma_\alpha(\omega)$ can be divided into two classes A and B. With the definition $\Sigma_\pm(\omega) = \frac{1}{2}\big(\Sigma_A(\omega) \pm \Sigma_B(\omega)\big)$, we have the following decomposition

$$\boldsymbol{\Sigma}(\omega) = \Sigma_+(\omega)\mathbf{1} + \Sigma_-(\omega)\,\boldsymbol{\gamma}^{(0)}_{D=2}\,, \tag{51}$$

where $\boldsymbol{\gamma}^{(0)}_2 = \boldsymbol{\tau}_z$ for $D = 2$. We see that the topological Hamiltonian

$$\boldsymbol{t}_{\text{top}}(\boldsymbol{k}) = \boldsymbol{t}(\boldsymbol{k}) + \boldsymbol{\Sigma}(\boldsymbol{k}, \omega{=}0) = \boldsymbol{t}(\boldsymbol{k}) + \Sigma_+(0)\mathbf{1} + \Sigma_-(0)\boldsymbol{\tau}_z \tag{52}$$

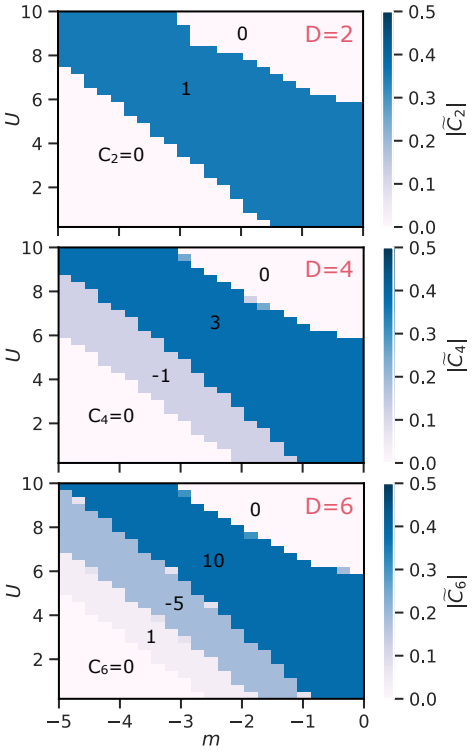

Fig. 7: *Topological m-U phase diagram of the interacting QWZ+U model for dimensions $D = 2, 4, 6$ as obtained by DMFT. The $D/2$-th Chern number C_D for the different phases is directly indicated in the plots. For $D = 2$, the first Chern number $C = C_2$. The color scale codes the modulus of the normalized Chern number defined as $\widetilde{C}_D = C_D/2^D$ (see text). Energy units: $t = t^*/\sqrt{D}$ with $t^*=1$. Figure taken from Ref. [5].*

acquires a nontrivial additional (staggered) orbital structure due to the second term, which adds to the term in Eq. (46) proportional to $\boldsymbol{\gamma}^{(0)}_{D=2}$. We thus expect a nontrivial modification of the $U = 0$ topological phase diagram.

The actual calculation by means of DMFT is conceptually simple: The lattice model is self-consistently mapped onto two Anderson impurity models, one for the A orbitals and another one the B orbitals, which are related to each other by A-B sublattice symmetry. As our intention is to cover the entire m-U parameter space, a simplified DMFT scheme, the two-site DMFT [22], is helpful. This has successfully been cross-checked against accurate DMFT results obtained with the standard Lanczos solver [23].

The top panel of Fig. 7 displays the topological phase diagram in the m-U control-parameter space for dimension $D = 2$. The phase diagram is shown for negative m only. The $m > 0$ range is obtained by reflection at the $m = 0$ axis with a simultaneous sign change $C \mapsto -C$. At $U = 0$ and when increasing m, we thus pass from $C = 0$ over $C = +1$ (for $m > -\sqrt{2}$) and $C = -1$ (for $m > 0$) to $C = 0$ (for $m > \sqrt{2}$), i.e., from a band insulator over two topologically different Chern insulators back to a band insulator again.

Right at $m = m_c = 0$, i.e., at the topological phase transition, the bulk system is a so-called semi-metal. The semi-metal state at this critical mass parameter does not have a (one-dimensional) Fermi surface. The gap is rather closed at isolated $\boldsymbol{k}$ points in the 1BZ, namely at $\boldsymbol{k}_c = (\pi, 0)$ and $\boldsymbol{k}_c = (0, \pi)$.

For $U > 0$, the Chern number is computed from the topological Hamiltonian, Eq. (52), and is thus affected by the self-energy at $\omega = 0$. As $\boldsymbol{\Sigma}(\omega)$ is diagonal in orbital space, this merely amounts to a renormalization of the chemical potential, $\mu \to \mu + \Sigma_+(\omega{=}0)$, and of the mass parameter, $m \to m + \Sigma_-(\omega{=}0)$. Since m acts as the strength of a staggered orbital field, $m = 0$ means that the orbital polarization $n_A - n_B$ vanishes and that $\Sigma_A(\omega) = \Sigma_B(\omega)$. Hence, $\Sigma_-(\omega) = 0$ and $\Sigma_+(\omega) = \Sigma_A(\omega) = \Sigma_B(\omega)$. Furthermore, the $\omega = 0$ value of $\Sigma_+(\omega) = U/2$ is exactly compensated by the chemical potential $\mu = U/2$. Hence, for finite but not too large U and at $m = 0$, the topological Hamiltonian equals the noninteracting one, and thus the system remains in a (correlated) semi-metal state.

For $U > U_c \approx 6$, the self-consistently determined DMFT self-energy develops a pole at $\omega = 0$. This implies that the concept of the topological Hamiltonian breaks down. In fact, at U_c there is a transition from the correlated semi-metal to a topologically trivial Mott insulator with $C = 0$, which is somewhat reminiscent of the paradigmatic metal-to-Mott-insulator transition predicted by DMFT [18].

For $m \neq 0$, the system becomes orbitally polarized. At constant $U < U_c$ and for decreasing m, the Chern-insulator phase with $C = 1$ extends up to a U-dependent critical mass parameter $m_c(U)$, where a topological phase transition to the band insulator takes place. Upon decreasing m for $U > U_c$ the Mott insulator is stable up to a second U-dependent critical mass parameter $m_{\text{Mott}}(U)$, where a transition to the correlated $C{=}1$ Chern-insulator takes place. The phase diagram shown in Fig. 7 suggests that the Mott insulator and the band insulator (both with $C{=}0$) cannot be connected continuously, as they are separated in parameter space by topologically nontrivial phases with $C \neq 0$.

Arbitrary even dimensions D

Since DMFT is a (rather questionable) approximation in case of a two-dimensional lattice model, it is worthwhile to ask if and how these results change for the same model on a higher-dimensional lattice. Eventually, in the $D \to \infty$ limit, the DMFT becomes exact, so that one might expect exact results for a strongly correlated and topologically nontrivial model at the same time. This extension to lattices of arbitrary dimensions D requires some effort. We concentrate on even dimensions, for odd D a somewhat different but similar approach is necessary. Let us mention that the $D \to \infty$ limit is not unique and that approaching it via even or via odd D yields different results.

The QWZ model on the hypercubic lattice with arbitrary even D, defined via Eq. (46), requires an explicit representation for the Hermitian and traceless generators $\gamma_D^{(\mu)}$ with $\mu = 0, 1, ..., D$ of the Clifford algebra Eq. (47). There is in fact a general recursive prescription [12], namely

$$\begin{aligned} \boldsymbol{\gamma}_{D+2}^{(0)} &= \boldsymbol{\tau}_z \otimes \mathbf{1} \quad , & \boldsymbol{\gamma}_{D+2}^{(r)} &= \boldsymbol{\tau}_x \otimes \boldsymbol{\gamma}_D^{(r)} \quad \text{for } r = 1, ..., D \\ \boldsymbol{\gamma}_{D+2}^{(D+1)} &= \boldsymbol{\tau}_x \otimes \boldsymbol{\gamma}_D^{(0)}, & \boldsymbol{\gamma}_{D+2}^{(D+2)} &= \boldsymbol{\tau}_y \otimes \mathbf{1} \, , \end{aligned} \tag{53}$$

where $\mathbf{1}$ is the $2^{D/2}$-dimensional unity. Explicitly, $\boldsymbol{\gamma}^{(0)} = \text{diag}(+1, +1, ..., -1, -1, ...)$. The number of orbitals $M = 2^{D/2}$ scales exponentially with D. Since a lower-dimensional faithful

representation is not possible [12, 24], the model becomes rather artificial. This is the price we have to pay to get a correlated and topologically nontrivial model solvable by DMFT.

Noninteracting electronic structure

At $U = 0$ and arbitrary even D, the band structure is easily obtained by squaring $\boldsymbol{t}(\boldsymbol{k})$, as shown above, and we get the dispersions of the two $M/2$-fold degenerate bands $\epsilon_\pm(\boldsymbol{k}) = \pm\big(t^2 \sum_r \sin^2 k_r + (m + t\sum_r \cos k_r)^2\big)^{1/2}$. We set $t = t^*/\sqrt{D}$ and $t^* = 1$. Due to the point-group symmetries, band closures are found at the high-symmetry points (HSPs) $\boldsymbol{k}_\mathrm{c} = \boldsymbol{k}_{n_0} \equiv (0, ..., 0, \pi, ..., \pi)$ in the 1BZ, and for the $\binom{D}{n_0}$ inequivalent permutations of the components, where n_0 counts the number of vanishing entries k_r. To get a band closure, the condition $m = (D-2n_0)\, t$ must be met or, equivalently,

$$m/t^* = (D-2n_0)/\sqrt{D}\,. \tag{54}$$

Thanks to the Clifford algebra, we can easily compute the moments $M_\alpha^{(n)} = \int d\omega\, \rho_\alpha^n(\omega)$ of the local partial density of states (DOS) of the orbital α for small n. We find the normalization condition $M_\alpha^{(0)} = 1$, the center of gravity $M_\alpha^{(1)} = m\gamma_{\alpha\alpha}^{(0)} = \pm m$, and the α-independent variance $M_\alpha^{(2)} - \big(M_\alpha^{(1)}\big)^2 = t^2 D = 1$. Hence, the scaling of the hopping leads to a proper balance between H_0 and H_1 for $D\to\infty$. The mass parameter m must not be scaled to maintain a nontrivial model.

The dispersions for the $D = 4$ model are shown in Fig. 8 (left). For $D = 4$ we have $M = 2^{D/2} = 4$, and thus the degeneracy of each band is $M/2 = 2$. Comparing results for different m (see color code), we find gap closures for $m = -2, -1, 0, 1, 2$ at the high-symmetry points (HSPs) $\boldsymbol{k}_\mathrm{c}$ with $n_0 = 4, 3, 2, 1, 0$, respectively. This is consistent with Eq. (54), which yields $m = (4-2n_0)/2 = 2-n_0$ for $D = 4$. The respective number of equivalent HSPs is 1, 4, 6, 4, 1, and the total number of wave vectors getting critical when varying m from $m = -\infty$ to $m = +\infty$ is $2^D = 16$.

The right panel in Fig. 8 displays the A-orbital DOS in the $D \to \infty$ limit for mass parameter $m = -1.5$. It can be calculated analytically (see Ref. [5])

$$\rho_\alpha(\omega) = \frac{1}{2}\frac{1}{t^*\sqrt{\pi}}\Theta\big(|\omega| - t^*/\sqrt{2}\big)\,\mathrm{sign}\,\omega \sum_{s=\pm}\left(\frac{\omega}{\sqrt{\omega^2 - t^{*2}/2}} + s z_\alpha\right) e^{-\frac{\left(s\sqrt{\omega^2 - t^{*2}/2} - m\right)^2}{t^{*2}}}. \tag{55}$$

Recall the definition $z_\alpha = +1, -1$ for A, B orbitals. Interestingly, the DOS has an m-independent gap given by $\Delta = \sqrt{2}t^*$, opposed to any finite-D DOS. This point requires discussion, see below.

Topological phase diagrams at $U = 0$

The middle panel of Fig. 8 displays the topological phase diagrams of the noninteracting QWZ model for $D = 2, 4, ...$, when using m as a control parameter. For each even dimension D, the red points mark the critical mass parameters given by Eq. (54). In fact they separate topologically inequivalent states of the system characterized by different Chern numbers. Actually, for arbitrary even D, the relevant invariant is the $(D/2)$-th Chern number C_D, which is obtained

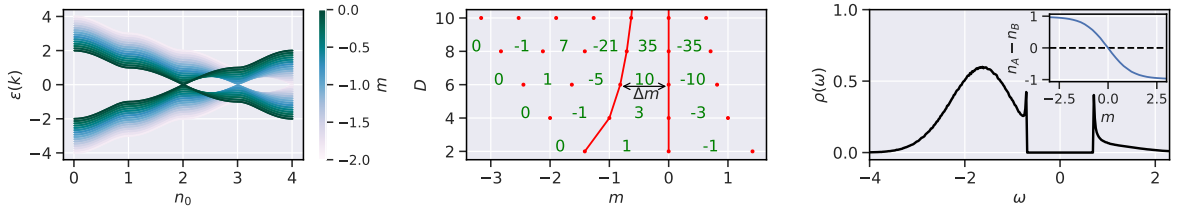

Fig. 8: *Dispersions of the two degenerate bands $\epsilon_\pm(\boldsymbol{k})$ of the $D = 4$ model* (left) *along straight shortest lines in the 1BZ connecting HSPs characterized by n_0. Results for different m, see color code.* Middle: *Different topological phases with Chern numbers $C_D(n_0)$ (green), separated by critical m-values (red dots) for different D.* Right: *$U = 0$ DOS on the A-orbitals at $m = -1.5$ for $D = \infty$. Inset: orbital polarization as function of m for $D = \infty$. Nearest-neighbor hopping: $t = t^*/\sqrt{D}$, $t^* = 1$ sets the energy scale. Figure taken from Ref. [5].*

as a D-dimensional integration of the D-form, which has been discussed above, over the 1BZ. (Note that here we depart from the conventional notation $C_{D/2}$ for the $(D/2)$-th Chern number.) The $(D/2)$-th Chern number of a topological phase for an m that satisfies, for some $n_0 \in \{0, ..., D-1\}$, the condition

$$\frac{D-2n_0-2}{\sqrt{D}} < m/t^* < \frac{D-2n_0}{\sqrt{D}} \,, \tag{56}$$

i.e., that lies between the critical mass parameters with band closures at $\boldsymbol{k}_{2n_0-2}$ and $\boldsymbol{k}_{n_0}$, respectively, is given by [12, 25, 26]

$$C_D(n_0) = (-1)^{n_0+\frac{D}{2}} \binom{D-1}{n_0} \,. \tag{57}$$

For the extreme cases

$$m/t^* < -\frac{D}{\sqrt{D}} \quad \text{and for} \quad \frac{D}{\sqrt{D}} < m/t^* \tag{58}$$

we have band insulators with $C_{D/2}(n_0 = D-1) = 0$ and $C_{D/2}(n_0{=}0) = 0$, respectively.

Fig. 8 (middle) shows that the m-distance between two neighboring transitions shrinks to zero for $D \to \infty$. Eq. (56) gives us $\Delta m = 2t^*/\sqrt{D}$. Hence, for high D the system is arbitrarily close to criticality for *any* m, and in the limit $D \to \infty$, the set of critical m's becomes dense in any finite m-interval. The phase diagram exhibits a *continuum* of topologically different phases. This observation has a couple of implications.

First, it means that the definition of a critical point in the 1BZ becomes elusive for $D \to \infty$, since $\epsilon_\pm(\boldsymbol{k}) = \epsilon_\pm(\boldsymbol{k}')$ if $\|\boldsymbol{k}-\boldsymbol{k}'\| = 0$, where we have defined $\|\boldsymbol{k}\|^2 \equiv \lim_{D\to\infty} D^{-1} \sum_{r=1}^{D} k_r^2$. $\|\cdot\|$ is a semi-norm, i.e., $\|\boldsymbol{k}\| = 0$ does not imply $\boldsymbol{k} = 0$. Hence, the concept of band closures at isolated points in the 1BZ breaks down. For a given m, we have $\epsilon_\pm(\boldsymbol{k}) = 0$ at a critical wave vector $\boldsymbol{k} = \boldsymbol{k}_c(m)$ but also at all wave vectors $\boldsymbol{k}$ with $\|\boldsymbol{k}-\boldsymbol{k}_c(m)\| = 0$. Furthermore, while the number $\binom{D}{n_0}$ of equivalent critical HSPs at a given critical m and the total number 2^D of HSPs in the 1BZ diverge, their ratio approaches a constant when $D \to \infty$.

Second, since all values of the mass parameter are critical, one would expect the absence of a gap in the DOS for any m. To understand the fact that the DOS displayed in Fig. 8 (right) has

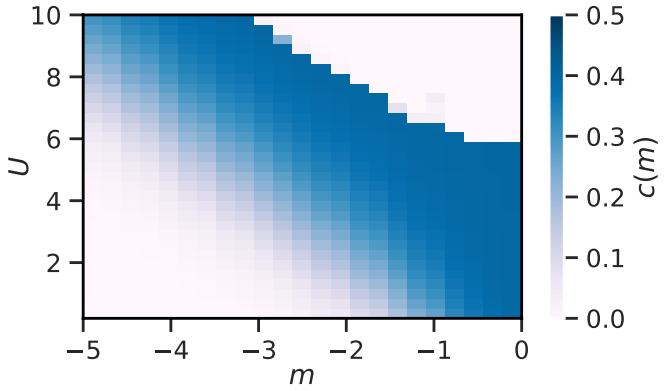

Fig. 9: *The topological phase diagram of the QWZ+U model in the $D\to\infty$ limit. The color codes the topologically invariant Chern density $c(m,U)$ in the m-U plane. Only the $m\leq 0$ range is plotted as the phase diagram is symmetric under $m\to -m$. Figure taken from Ref. [5].*

an m-independent gap $\Delta = 2t^*$, one can analytically analyze the DOS at a critical m at any finite D. For sufficiently small frequencies (see Ref. [5]),

$$\rho_\alpha(\omega) = c(D,n_0)|\omega|^{D-1}/t^{*D} \tag{59}$$

with a coefficient $c(D,n_0)$ which, for any n_0 and thus for any critical m, tends to zero exponentially fast with $D\to\infty$. This demonstrates that the contribution of wave vectors with $\epsilon_\pm(\boldsymbol{k})=0$ to the DOS vanishes in the $D\to\infty$ limit. For $D\to\infty$, we conclude that the DOS does not indicate whether the system is a (Chern) insulator or a semi-metal. The distinction between insulator and semi-metal becomes meaningless in this limit.

A third important observation is related to the Chern number. Eq. (57) shows that for $D\to\infty$ only the modulus of the Chern number, and only after proper normalization, has a well-defined limit. Since $\sum_{n_0=0}^{D-1}|C_D(n_0)| = 2^{D-1}$, we thus introduce a *Chern density* as $c(n_0) = \lim_{D\to\infty}|C_D(n_0)|/2^{D-1}$. To express the Chern density as a function of m, we employ $n_0 = \big(D-m\sqrt{D}/t^*\big)/2$, see Eq. (54). Furthermore, since $\Delta m\mapsto 0$ we write $dm\equiv 2t^*/\sqrt{D}$. With the Moivre-Laplace theorem, which states that the binomial distribution converges to the normal distribution, we then find

$$c(n_0) = \lim_{D\to\infty}\sqrt{\frac{2}{\pi D}}e^{-2\frac{(D/2-n_0)^2}{D}} = c(m)\,dm \tag{60}$$

with

$$c(m) = \frac{1}{t^*\sqrt{2\pi}}e^{-\frac{1}{2}\frac{m^2}{t^{*2}}}\,. \tag{61}$$

The Chern density is a smooth Gaussian density. It is normalized by construction and turns out to have unit variance. This is important, as it shows that not only dynamic correlation effects but also nontrivial topological properties survive the $D\to\infty$ limit when using the standard scaling of the hopping.

Topological phase diagram of the interacting system for $D\to\infty$

Fig. 9 displays the topological phase diagram of the QWZ+U model in the limit $D\to\infty$. The control parameters are m and U, and the topological invariant, the Chern density, is indicated by

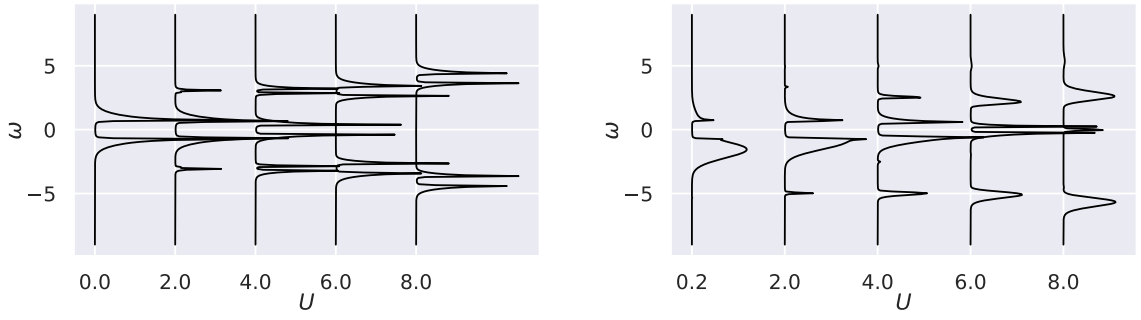

Fig. 10: *Local spectral function on $\alpha = A$-orbitals for $m = 0$ (left) and $m = -1.5$ (right) for various U in the $D \to \infty$ limit as obtained by two-site DMFT. Figure taken from Ref. [5].*

the color code. The results have been obtained numerically employing the simplified two-site DMFT scheme [22, 5] and have been cross-checked against full DMFT employing the Lanczos solver [23]. The Chern density is obtained from the topological Hamiltonian, where we merely have to take into account a renormalization of the chemical potential, $\mu \mapsto \mu + \Sigma_+(\omega{=}0)$, and of the mass parameter, $m \mapsto m + \Sigma_-(\omega{=}0)$.

We first discuss the $U = 0$ case. As a function of m, the Chern density is a Gaussian, see Eq. (61). This means that the system smoothly evolves from a conventional band insulator, with $c(m, 0) \to 0$ in the limit $m \to -\infty$, to a Chern insulator (or semimetal, this cannot be distinguished). The invariant increases and takes the maximum value $c(m, 0) = 1/\sqrt{2\pi}$ at $m = 0$. The phase diagram for $m > 0$ is simply obtained by reflection at the $m = 0$ axis (also for $U > 0$). Already the $U = 0$ phase diagram is quite unconventional, since we have an infinitely fine classification, where *each* value of the (modulus of the) control parameter m defines a separate topological phase. At the same time, each value for m is also critical.

With increasing U at $m = 0$, the Chern density $c(0, U)$ stays at its maximum, until at $U = U_\mathrm{c} \approx 6t^*$ the system undergoes a correlation-driven transition to a Mott-insulating phase. The Mott phase is topologically trivial with $c = 0$. Approaching U_c either from above or from below, the transition is characterized by a continuously vanishing renormalization factor $z \mapsto 0$, where $z \equiv 1/\big(1 - \partial\Sigma_\alpha(\omega{=}0)/\partial\omega\big)$ is independent of the orbital type α. The quantity z plays the role of a band-gap renormalization [27]. Examples for the U-dependent interacting spectral function on the A orbitals are given in Fig. 10 for $m = 0$ and for $m < 0$.

For $U > U_\mathrm{c}$, the Mott phase extends to $m < 0$ and is bounded for all m by a line of critical interactions $U_\mathrm{c}(m)$. For $m \to -\infty$ we observe that $U_\mathrm{c}(m)$ develops into a linearly increasing function of $|m|$. This is due the fact that the system becomes fully orbital-polarized, see the inset in Fig. 8 (right). Hence, the self-energy becomes static and approaches constants $\Sigma_A \to U$, $\Sigma_B \to 0$, such that the renormalization of m is trivial: $m \to m + \Sigma_-(\omega{=}0) \to m + U/2$. As a consequence, the band insulator with $c = 0$ cannot be smoothly connected to the Mott insulator with $c = 0$ without passing topologically nontrivial states with $c > 0$.

The whole phase diagram Fig. 9 can be understood as being the $D \to \infty$ limit of m-U phase diagrams at finite D, see Fig. 7, where we have used the same color coding as in Fig. 9 for the

Chern density, and where the (signed and unnormalized) Chern numbers are given additionally and directly label the topological phases in the figure. Note the alternating sign and the monotonic increase of the modulus of the Chern number along any straight path from the band to the Mott insulator. With increasing D, the number of topologically nontrivial phases $C_D(m) \neq 0$ increases and they cover ever narrower regions in the m-U plane (recall $\Delta m = 2t^*/\sqrt{D}$ for $U = 0$), until they shrink to one-dimensional curves of constant color given by $c(m, U) = \text{const}$ in Fig. 9. In the limit $D \to \infty$, systems on these iso-Chern curves are topologically equivalent, while on paths crossing iso-Cherns one passes through a continuum of topologically different phases.

6 Concluding discussion

We have demonstrated that DMFT can be used to study the interplay between topology and correlations in a model for a Chern insulator on an infinite-dimensional lattice. While the approach is numerically exact, the underlying model necessarily involves, besides the $D \to \infty$ limit, an infinite number of orbitals and is thus somewhat artificial. There are three main lines along which the discussion could proceed:

The considered example belongs (at $U = 0$) to the so-called class A of the complete topological classification of noninteracting lattice-fermion systems. The other nine classes of the "ten-fold way" have not yet been considered (a simple modification of the theory covers class A III). One important question is, whether interacting lattice-fermion models can be constructed that remain well-defined and nontrivial in the $D \to \infty$ limit for *all* classes. In the positive case, this would demonstrate that the ten classes "survive" if one switches on a Hubbard-type interaction, and we would have explicit models as prototypes, at least in the (comparatively simple) $D \to \infty$ limit. Even in this limit, where DMFT applies, however, we would not be sure that the classification was complete – another problem to be tackled.

Complexity rises dramatically, if one attempts a topological classification of all interacting systems on finite-dimensional lattice. Clearly, already the correlation problem is hard. On the other hand, we are allowed to consider continuous deformations, preferably to models, which are more simple from the correlation point of view. This might help to proceed. Up to now a general classification has not been achieved and theory is still at the very beginning, see Ref. [28] for a discussion. Some important questions are: Is the topological phase diagram of a noninteracting system stable against weak interactions? Vice versa, can topologically nontrivial phases be generated by interactions? In the $D \to \infty$ limit, the answer is "yes", see Fig. 9. More important, are there topologically nontrivial phases in interacting models that have no analog to those of noninteracting ones, i.e., cannot be connected to a phase classified by the ten-fold way? Is a topologically nontrivial Mott insulator possible? What are the consequences of the bulk-boundary correspondence in case of interacting systems?

On a more pragmatic level and with the focus to concrete materials, one may attempt to compute topological invariants of two- and three-dimensional interacting systems by approximate numerical tools. DMFT itself and its various extensions to include nonlocal correlations to

some extent, suggest themselves and appear attractive since the Chern number, for example, can be computed by means of a formula based on the single-particle Green function only [29, 30]. However, this expression rather is inconvenient for numerical studies as it involves an additional frequency integration. Invoking the tool of continuous deformations, it has been shown [31] that a strong simplification is possible and that it is sufficient to consider the self-energy at $\omega = 0$ only, i.e., that the invariant can be computed from the topological Hamiltonian. A different concept is that of the so-called pole expansion [32, 33]. Here, the self-energy is represented by a hybridization function involving auxiliary orbitals. For a given self-energy, this allows us to map the interacting to a noninteracting problem such that the standard methods of topological classification apply. However, it has been questioned [34] that these two concepts are equivalent in case of a $\boldsymbol{k}$-dependent self-energy. Another more fundamental problem is that any scheme, which is solely based on the single-particle Green function, disregards two-particle correlations to some extent. The two-particle excitation spectra may undergo a gap closure while the one-particle gap stays open, so that a complete topological classification cannot be achieved in this way. Ref. [35] provides an example along this line of reasoning.

Acknowledgements

The author would like to thank Lara Heyl, Christian Jöns, David Krüger and Robin Quade (University of Hamburg) for numerous discussions and acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG) through the Cluster of Excellence “Advanced Imaging of Matter”-EXC 2056, Project ID No. 390715994, and through the Research Unit FOR 5249-449872909 (Project P8) is gratefully acknowledged.

References

[1] B.A. Bernevig: *Topological insulators and topological superconductors* (Princeton University Press, 2013)

[2] J. Asbóth, L. Oroszlány, and A. Pályi: *A Short Course on Topological Insulators* Lecture Notes in Physics, vol. 919 (Springer, Cham, 2016)

[3] M. Nakahara: *Geometry, topology, and physics* (IOP Publishing, Bristol, 1998)

[4] S.M. Bhattacharjee, M. Mj, and A. Bandyopadhyay (eds.): *Topology and Condensed Matter Physics* (Springer, Singapore, 2017)

[5] D. Krüger and M. Potthoff, Phys. Rev. Lett. **126**, 196401 (2021)

[6] L. Heyl: *Oberflächenzustände von Chern-Isolatoren in hohen Raumdimensionen* (B.Sc. thesis, University of Hamburg, 2021)

[7] T. Kato, J. Phys. Soc. Jpn. **5**, 435 (1950)

[8] A. Messiah: *Quantenmechanik*, Vol. II (de Gruyter, Berlin, 1979)

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

[10] P.A.M. Dirac, Proc. R. Soc. London A **133**, 60 (1931)

[11] J.E. Avron, L. Sadun, J. Segert, and B. Simon, Commun. Math. Phys. **124** (1989)

[12] E. Prodan and H. Schulz-Baldes: *Bulk and Boundary Invariants for Complex Topological Insulators: From K-Theory to Physics* (Springer, 2016)

[13] C.-K. Chiu, J.C.Y. Teo, A.P. Schnyder, and S. Ryu, Rev. Mod. Phys. **88**, 035005 (2016)

[14] X.-L. Qi, Y.-S. Wu, and S.-C. Zhang, Phys. Rev. B **74**, 085308 (2006)

[15] K. von Klitzing, Rev. Mod. Phys. **58**, 519 (1986)

[16] D.J. Thouless, M. Kohmoto, M.P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982)

[17] F.D.M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988)

[18] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)

[19] Z. Wang and B. Yan, J. Phys.: Condens. Matter **25**, 155601 (2013)

[20] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)

[21] E. Müller-Hartmann, Int. J. Mod. Phys. B **3**, 2169 (1989)

[22] M. Potthoff, Phys. Rev. B **64**, 165114 (2001)

[23] M. Caffarel and W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)

[24] S. Lang: *Algebra* (Springer, New York, 2002)

[25] M.F.L. Golterman, K. Jansen, and D.B. Kaplan, Phys. Lett. B **301**, 219 (1993)

[26] X.-L. Qi, T.L. Hughes, and S.-C.Zhang, Phys. Rev. B **78**, 195424 (2008)

[27] M. Sentef, J. Kuneš, P. Werner, and A.P. Kampf, Phys. Rev. B **80**, 155116 (2009)

[28] S. Rachel, Rep. Prog. Phys. **81**, 116501 (2018)

[29] Z. Wang, X.-L. Qi, and S.-C. Zhang, Phys. Rev. Lett. **105**, 256803 (2010)

[30] K. Ishikawa and T. Matsuyama, Z. Phys. C **33**, 41 (1986)

[31] Z. Wang and S.-C. Zhang, Phys. Rev. X **2**, 031008 (2012)

[32] S.Y. Savrasov, K. Haule, and G. Kotliar, Phys. Rev. Lett. **96**, 036404 (2006)

[33] L. Wang, H. Jiang, X. Dai, and X.C. Xie, Phys. Rev. B **85**, 235135 (2012)

[34] P. Thunström and K. Held, arXiv:1907.03899

[35] Y.-Y. He, H.-Q. Wu, Z.Y. Meng, and Z.-Y. Lu, Phys. Rev. B **93**, 195164 (2016)