DOCTORATE THESIS

Symmetry-protected topological phases: From Floquet theory to machine learning

OLEKSANDR BALABANOV

UNIVERSITY OF GOTHENBURG

Department of Physics
University of Gothenburg
Göteborg, Sweden 2020

Symmetry-protected topological phases:
From Floquet theory to machine learning
Oleksandr Balabanov

ISBN 978-91-8009-014-8 (PRINT)
ISBN 978-91-8009-015-5 (PDF)

This thesis is electronically published, available at
http://hdl.handle.net/2077/64531

Department of Physics
University of Gothenburg
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-786 00 00

Printed by Stema specialtryck AB
Borås, Sweden 2020

ABSTRACT

It is by now a well known fact that boundary states in conventional time-independent topological insulators are protected against perturbations that preserve relevant symmetries. In the first part of this thesis, accompanying Papers A - C, we study how this robustness extends to time-periodic (Floquet) topological insulators. Floquet theory allows us to go beyond ordinary time-independent perturbations and study also periodically-driven perturbations of the boundary states. The time-dependence here opens up an extra lever of control and helps to establish the robustness to a much broader class of perturbations. In Paper A, a general idea behind the topological protection of the boundary states against time-periodic perturbations is presented. In Paper B we address the experimental detection of the proposed robustness and suggest that signatures of it can be seen in the measurements of linear conductance. Our idea is explicitly illustrated on a case study: A topologically nontrivial array of dimers weakly attached to external leads. The discussed features are described analytically and confirmed numerically. All computations are performed by employing a convenient methodology developed in Paper C. The idea is to combine Landauer-Büttiker theory with the so-called Floquet-Sambe formalism. It is shown that in this way all formulas for currents and densities essentially replicate well known expressions from time-independent theory.

To find closed mathematical expressions for topological indices is in general a nontrivial task, especially in presence of various symmetries and/or interactions. The second part of the thesis introduces a computational protocol, based on artificial neural networks and a novel topological augmentation procedure, capable of finding topological indices with minimal external supervision. In Paper D the protocol is presented and explicitly exemplified on two simple classes of topological insulators in 1d and 2d. In Paper E we significantly advance the protocol to the classification of a more general type of systems. Our method applies powerful machine-learning algorithms to topological classification, with a potential to be extended to more complicated classes where known analytical methods may become inapplicable.

The thesis is meant to serve as a supplement to the work contained in Papers A-E. Here we provide an extensive introduction to Floquet theory, focused on developing the machinery for describing time-periodic topological insulators. The basic theory of artificial neural nets is also presented.

Sammanfattning

Ett välkänt resultat i teorin för topologiska kvantfaser säger att kanttillstånden i konventionella tidsoberoende topologiska isolatorer är skyddade mot störningar vilka bevarar relevanta symmetrier. I denna avhandlings första del, svarande mot artiklarna A-C, studerar vi hur denna typ av stabilitet hos kanttillstånd kan utsträckas till att gälla för tidsperiodiska (Floquet) topologiska isolatorer. Förutom vanliga tidsoberoende störningar så kan vi med hjälp av Floquet-teori analysera också tidsperiodiska störningar och deras effekter på kanttillstånden. Tidsberoendet öppnar här upp för en extra frihetsgrad som gör det möjligt att stabilisera kanttillstånden i närvaron av en mycket större klass av störningar. I artikel A presenterar vi en allmän teori för topologiskt skydd av kanttillstånd mot tidsperiodiska störningar. I artikel B diskuterar vi möjliga experimentella test av kanttillståndens stabilitet och förutsäger hur konduktansmätningar förväntas signalera stabiliteten. Som fallstudie analyserar vi en modell av en atomkedja i en topologiskt icke-trivial fas, svagt kopplad till yttre ledare. Analytiska och numeriska beräkningar bygger här på en metod introducerad i artikel C vilken kombinerar Landauer-Büttiker teori med Floquet-Sambe formalism. Metoden reducerar i allt väsentligt formler för strömmar och tätheter till motsvarande välkända uttryck i tidsoberoende teori.

Att identifiera matematiska uttryck för topologiska index vilka karakteriserar topologiska kvantfaser är ofta svårt, särskilt i närvaron av specifika symmetrier och/eller växelverkningar. I avhandlingens andra del presenteras ett nytt numeriskt beräkningsprotokoll — baserat på artificiella neurala nätverk och en ny typ av topologisk data-augmentering — vilken möjliggör identifikationen av topologiska index med minimal extern kontroll. Protokollet beskrivs i artikel D och exemplifieras med beräkningar på två klasser av enkla modeller av topologiska isolatorer, i 1d och 2d. I artikel E vidareutvecklar vi protokollet för klassifikation av mer allmänna typer av topologiska system. Vår metod utnyttjar kraftfulla maskininlärningsalgoritmer för topologisk klassifikation, med potentialen att kunna användas för än mer komplexa topologiska system där kända analytiska metoder inte längre är tillämpbara.

Denna avhandling syftar till att ge en introduktion och bakgrund till de problem som behandlas i artiklarna A-E. Vi ger här speciellt en utförlig introduktion till Floquet-teori, med fokus på dess tillämpning på tidsberoende topologiska isolatorer. Vi presenterar också grunderna för neurala nätverk.

List of papers

This thesis is based on five appended papers:

Paper A

O. Balabanov and H. Johannesson, *Robustness of symmetry-protected topological states against time-periodic perturbations*, Phys. Rev. B **96**, 035149 (2017).

Paper B

O. Balabanov and H. Johannesson, *Transport signatures of symmetry protection in 1D Floquet topological insulators*, J. Phys.: Condens. Matter **32** 015503 (2020).

Paper C

O. Balabanov, *Transport through periodically driven systems: Green's function approach formulated within frequency domain*, arXiv:1812.05755 (2018).

Paper D

O. Balabanov and M. Granath, *Unsupervised learning using topological data augmentation*, Phys. Rev. Research **2**, 013354 (2020).

Paper E

O. Balabanov and M. Granath, *Unsupervised interpretable learning of topological indices invariant under permutations of atomic bands*, (to be submitted).

My contribution

I conducted all analytic and numeric calculations in Papers A-E. The analysis was also mainly done by me, with fillings of Henrik Johannesson (HJ) in Papers A-B, and Mats Granath (MG) in Papers D-E. The writing was a joint effort of me and HJ in Papers A-B, and me and MG in Papers D-E. Everything in Paper C is entirely my work.

ACKNOWLEDGEMENTS

I am particularly grateful to my supervisor Henrik Johannesson. Under his wise guidance I developed as a scientist, he always supported my ideas and was constructive with the criticism. Henrik's help in research and outside the academia was uncountable, and I sincerely appreciate it very much.

Also, special thanks to Mats Granath and Bernhard Mehlig, with whom I collaborated for a more than a year. Our discussions were very fruitful and lead to my better understanding of various problems in physics and machine learning.

I would also like to thank my office mates and just good friends Jonathan, Lorenzo, and Yoran for always being open for discussions and sharing their thoughts on different subjects. Thanks to everyone in our department, I was very fortunate to work in such great and friendly environment.

Most importantly, thanks to my family for always being there for me. It was a long journey in my life, with many ups and downs, but I never felt alone.

CONTENTS

PART I
THESIS

1 Introduction

Quantum phases describe many-body orders in quantum matter and point to qualitatively different classes of materials. To identify and systematically study various types of orders is unarguably of key importance. The conventional scheme for classifying quantum matter and transitions between phases of quantum matter builds on the work by Landau from the 1930s [1]. Studying classical finite-temperature phase transitions, Landau came up with a very intuitive classification approach stating that any transition to a phase associated with a higher degree of order happens if the many-body states spontaneously break a symmetry of the underlying Hamiltonian. His paradigm can be intuitively illustrated on the standard example of a paramagnetic to ferromagnetic phase transition where spins on a lattice align by spontaneously breaking the continuous rotation symmetry. The difference between the two phases is then formally identified by constructing a so-called *order parameter*, required to change from zero to a finite value as one crosses the phase transition point into the ferromagnetic phase. By constructing a local order parameter (depending on local coordinates) and incorporating it into an effective field theory, Landau, together with Ginzburg [2], formulated a macroscopic theory of superconductivity, by this providing a prescription for the analysis of any symmetry-breaking phase transition. The standard Landau theory of phase transitions via symmetry-breaking laid the foundation for understanding many basic phenomena in physics and serves as a corner-stone of condensed matter theory. Despite of all its success, the Landau formalism has turned out to be very far from being complete and by now many other types of orders have been revealed and actively investigated – foremost those classified by topology and identifiable by a nonzero value of a *non-local* order parameter [3].

Topological quantum phase transitions have been extensively studied

over the past decades, with many challenging questions still remaining open. These phase transitions are not governed by any symmetry breaking and lie outside the standard Landau paradigm. A simple case in point is the well known Integer Quantum Hall Effect [4, 5, 6, 7, 8, 9, 10] where the ground state of a 2d electron gas subject to a perpendicular magnetic field undergoes multiple quantum phase transitions as the field is varied: At each critical point the Hall condensate of electrons reorders leading to the emergence of robust chiral boundary states. All quantum Hall states belong to the same symmetry class and cannot be distinguished within Landau's theory – this system manifests a fundamentally different type of order, linked to topology.

Topological quantum phases come in two basic varieties [3, 11], 'topologically ordered' and 'symmetry-protected' (SPT). Intrinsic topological order is present only in phases that can not be continuously transformed into a product state without a gap-closing phase transition. Order here originates from the intrinsic long-range entanglement of the many-body state. Differently, two SPT phases are said to be topologically distinct if they can not be continuously transformed into each other while preserving the underlying symmetries. In fact, if the symmetries are allowed to be broken all SPT phases can generically be transformed to a product state, in other words, they have no (non-symmetry-protected) topological order. Still, two different SPT phases cannot be connected via any local continuous reordering of the condensate within a given class of symmetries, defining a robust obstruction between the corresponding orders. Recently, a third class of topological quantum phases have been postulated and termed by symmetry enriched topological (SET) phases [12]. The SET phases describe topological quantum phases within which the topological and SPT orders intertwine.

To find and analyze different topological and SPT orders is in general a challenging task. Over the years many advanced analytic techniques were developed [13, 14, 15, 16, 17, 18, 19] and successfully applied to reveal various topological phases of quantum matter. Despite of all the progress made within the theory of phase transitions, there are lots of bits still missing. The types of orders lacking from the overall picture are most likely very cumbersome to identify and some of them might not be even within analytic reach. Machine Learning (ML) protocols have a potential of becoming very useful here if that is the case. Artificial neural networks [20] are of particular interest because they specifically specialize in recognizing patterns in data that are far beyond what can presently be expected to be achieved using

analytical techniques. Topological phase identification within ML is a recent research area [21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35] and currently it undergoes an exploratory stage: The focus is mainly placed on developing concrete computational protocols revealing the capabilities of various Machine Learning techniques. The protocols are first verified on basic examples of well-studied systems, with the long-term goal to gradually increase the complexity of the studied systems to a highly nontrivial level. The field is rapidly developing and by now has attracted a lot of attention within the physics community.

This doctoral thesis primarily focuses on topological band insulators (TIs), i.e. topologically nontrivial quantum phases within the class of noninteracting theories. Any ground state of a quantum system without interactions is described by a symmetrized or anti-symmetrized product state and it has no intrinsic topological order. Nevertheless, noninteracting systems exhibit a rich verity of SPT orders and they have been at the center of research activity over the past years [36], both theoretically and experimentally. Simply put, the TIs have bulk band structures similar to ordinary (trivial) band insulators: There is a nonzero band gap separating conduction from valence bands and the Fermi level lies between those energy levels. However, the system becomes necessarily gapless at the boundaries. In 2d (3d) the corresponding gapless boundary modes are confined in one (two) direction(s) and extended in the others. Ideally these modes carry dissipationless currents since there is no channel for backscattering if the underlying symmetries are maintained. The boundary states in 1d TIs, on the other hand, are completely localized, with zero energy, and hence stay put at the middle of the band gap. The symmetry constraints allow these states to be gapped out only in pairs, cf. Chapter 2, in this way protecting them against gap-preserving perturbations respecting the symmetries. The topological phase transitions within noninteracting theories are by now considered to be well understood and essentially complete. Nevertheless, there is still a lot of unexplored ground left, with new concepts and ideas still popping up in the literature. Examples in point include fairly recent proposals of non-Hermitian topological phases [37, 38], topologically robust states defined on fractal lattices [39, 40], topological phase transitions at critically [41], and others.

At the base of any topological band insulator lies some topological nontriviality of its bulk band structure: The quantum phases are put into different classes corresponding to their bulk topological features. Different classes are

labeled by different numbers, called topological invariants. The boundary properties are then established via so-called index theorems. They connect various topological invariants to the number of the symmetry-protected boundary states, generally termed the bulk-edge correspondence. This connection is often assumed without proof, however; it is not obvious at all and always has to be rigorously justified, see for example [36, 42, 43, 44].

Interestingly, topological quantum phases associated with non-triviality of the bulk band structure can be extend to quantum systems driven by external time-periodic fields [45, 46, 47]. This is done within the so-called Floquet-Bloch theory where a concept analogous to the energy band structure, dubbed the quasienergy band structure, is introduced. The Floquet TIs are exceptionally interesting because they bring the concept of SPT order to states that are out of equilibrium. Despite the conceptual similarity with the conventional undriven SPT phases, Floquet TIs have been shown to exhibit SPT phases having no analogue in the time-independent case [48]. As a matter of fact, all undriven phases are trivially time-periodic and therefore they are automatically incorporated within the Floquet formalism, but only as special cases. Being a fascinating subject, a large part of this thesis elaborates on various aspects of Floquet TIs.

To efficiently identify topologically nontrivial band structures one would want to have a simple and to some extent universal indicator of topological non-triviality. The so-called band inversions are known to be effective in this regard. They are usually thought of as a signal of a possible topological phase transition. However, it is important to bear in mind that the inversion of bands is neither a sufficient nor a necessary condition to have a topologically nontrivial phase [49]. The band inversion, in this context, means that the energy levels at high symmetry points in the Brillouin zone are flipped in respect to their natural order. In quantum chemistry, the natural order corresponds to the order of individual atomic orbitals. In effective models, commonly exploited in condensed matter physics, we cannot formally define which band order is by default the most natural. What we can definitely do, however, is to tell when the order of bands changes under a continuous variation of some parameters in the model, in this way signaling a possible topological phase transition. The band inversion is a very important concept and it supplies us with an efficient tool in search for topologically nontrivial materials. This concept will be often used throughout the thesis, in both stationary and periodically driven cases.

1.1 Structure of the thesis

The base of this doctoral thesis consists of Papers A-E appended in Part II. The papers are supplemented with a comprehensive introduction to the subject described in the main body of the thesis, Part I.

The body of the thesis starts by introducing two simple models of band insulators in Chapter 2, where basic concepts and terminology are explained on two concrete examples. In Chapter 3 we then generalize the main concepts and describe the standard ten-fold classification of SPT phases. The Floquet formalism is discussed in Chapter 4 and explicitly employed to extend the SPT classification to periodically driven systems in Chapter 5. Chapter 6 then focuses on a particular unique property of SPT phases, namely robustness of the symmetry-protected boundary states against time-periodic perturbations. It mainly summarizes Papers A-C, with the exception of Sec. 6.3 which introduces a new probe of Floquet symmetry protection based on tracking of the density evolution of a localized edge state. Chapters 7 and 8 are aimed to complement Papers D-E by providing the essentials on artificial neural networks, explaining basic ideas and some technicalities of machine learning. The thesis summary and outlook is given in Chapter 9.

2 Simple models of topological band insulators

To set the stage, let us first take a look at two simple models of topologically nontrivial systems, one in 1d and the other in 2d. Both of them are well studied and constitute textbook material [50] on noninteracting fermionic SPT phases. Since they illustrate the central concepts of topological matter we find them very suitable for starting our main discussion.

2.1 The Su-Schrieffer-Heeger model of polyacetylene

The simplified spinless Su-Schrieffer-Heeger (SSH) model of polyacetylene is a toy model describing spinless fermions placed on a 1d lattice with staggered hopping amplitudes. Within this model the basic ideas of bulk-edge correspondence and topological protection may be plainly explained [51, 52, 53, 54]. To define the model, consider spinless fermions hopping on a bipartite lattice of $M = 2N$ sites. This system is described by the following Hamiltonian,

$$H = -\sum_{j=1}^{N} \big(\gamma_{1,j}|A,j\rangle\langle B,j| + \gamma_{2,j}|B,j\rangle\langle A,j+1| + \text{h.c.}\big), \tag{2.1}$$

where A, B are the two sublattices of the bipartite lattice, and $\gamma_{1,j}$ ($\gamma_{2,j}$) are intercell (intracell) hopping amplitudes. It is common practice to assume that the system is homogeneous in its interior, i.e. far from the edges, cf. Fig. 2.1, implying that $\gamma_{1,j}$ and $\gamma_{2,j}$ are taken to be constants in the bulk.

In the entirely homogeneous case, $\gamma_{1,j} = \gamma_1$ and $\gamma_{2,j} = \gamma_2$ at *all* sites j, and under periodic boundary conditions the SSH model is described by the following Bloch Hamiltonian

$$H(k) = \begin{pmatrix} 0 & \gamma_1 + \exp(ik)\gamma_2 \\ \gamma_1^* + \exp(-ik)\gamma_2^* & 0 \end{pmatrix}, \tag{2.2}$$

and has the dispersion relation

$$\varepsilon(k) = \pm\sqrt{|\gamma_1|^2 + |\gamma_2|^2 + 2|\gamma_1||\gamma_2|\cos(k + \arg(\gamma_1) + \arg(\gamma_2))}, \tag{2.3}$$

where k is the momentum.

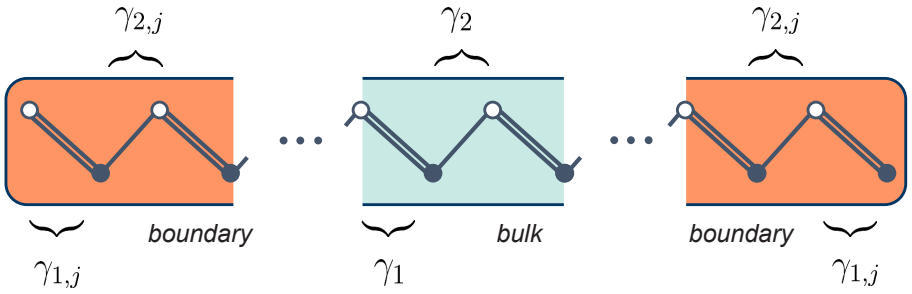

Figure 2.1: A schematic illustration of an SSH chain with intercell $\gamma_{1,j}$ and intracell $\gamma_{2,j}$ hopping amplitudes. We have here divided the chain into a homogeneous interior region (bulk) and, in general disordered, edges (boundaries).

By analyzing this dispersion relation we notice that there is an energy gap between the positive and negative energy bands, given by $\Delta = 2\left||\gamma_1| - |\gamma_2|\right|$. The band gap equals the stronger hopping amplitude minus the weaker one, and this is valid for all cases independently which one of the γ_1 and γ_2 is stronger. Let us fix γ_1 to some real value and sweep through all possible real γ_2. At first the band gap is open (for $|\gamma_2| < |\gamma_1|$), it then closes at $k = \pi$ (for $|\gamma_2| = |\gamma_1|$) and reopens again (for $|\gamma_2| > |\gamma_1|$). Under this process the two bands become inverted for the eigenstates at $k = \pi$ signaling that one may expect to have a topological phase transition here. Following this we classify all periodic SSH chains depending on the absolute values of γ_1 and γ_2. To be precise, we say that two chains are in the same class if their Hamiltonians $H(k)$ can be continuously transformed into each other without closing the energy gap. In other words, we create two classes of homogeneous space-periodic SSH chains, namely the class consisting of the systems with $|\gamma_1| > |\gamma_2|$ and the class containing the systems with $|\gamma_1| < |\gamma_2|$. It is important to emphasize that any space-periodic gapped SSH chain belongs to one of these two classes, and hence, in order to continuously transform two systems belonging to different classes into each other we have to close the band gap at some point during the transformation.

The inverted bands at $k = \pi$ hint to the fact that the SSH chains with $|\gamma_1| > |\gamma_2|$ and $|\gamma_1| < |\gamma_2|$ are topologically distinct, which can formally be justified by calculating the corresponding topological index, the so-called winding number. The winding number is defined specifically for systems

respecting a sublattice symmetry meaning that the Hamiltonian is allowed to contain only terms that couple sites from different sublattices. The winding number ν is then defined via [50]

$$\nu = \frac{1}{2\pi i}\int_0^{2\pi} dk\, \frac{d}{dk} \log\det H_{AB}(k), \tag{2.4}$$

where $H_{AB}(k)$ denotes one of the off-diagonal blocks of the matrix $H(k)$ in Eq. (2.2). The winding number calculates the number of times $\det H_{AB}(k)$ winds around the origin as the momentum k is swept through the Brillouin zone. Clearly, this topological index cannon be changed by any gap-preserving continuous deformation of $H(k)$ maintaining the sublattice symmetry. The winding number is $\nu = 0$ for $|\gamma_1| > |\gamma_2|$ and $\nu = 1$ for $|\gamma_1| < |\gamma_2|$ defining two topologically distinct classes of SSH chains.

Equivalently, the topological inequality of the SSH chains can be established by counting the states at their boundaries, generally known as bulk-edge correspondence. The bulk-edge correspondence ties the topological nontriviality of the bulk to the properties at the boundaries in the thermodynamic limit: In topologically nontrivial systems without interactions there necessarily exist gapless boundary states robust to local symmetry-preserving perturbations. The presence of robust edge states in SSH chains may be shown by the succeeding argumentation. The sublattice symmetry obligates the eigenstates of the SSH Hamiltonian to come in pairs: it is not hard to verify that a state $(\psi_{A,j}, \psi_{B,j})$ is an eigenstate with energy ε if and only if the state $(\psi_{A,j}, -\psi_{B,j})$ is also an eigenstate but with energy $-\varepsilon$. It follows that under a symmetry-preserving adiabatic change of parameters in the Hamiltonian a pair of states at zero energy can change energy only simultaneously, assuming there are no other states at zero level. It is clear that in the extreme case $\gamma_1 = 0$ there are two states with zero energies localized on opposite edges. We then adiabatically deform this chain into another one with arbitrary hopping amplitudes $|\gamma_1| < |\gamma_2|$ and this deformation is done in the following way: First we change the hopping amplitudes in the bulk region by keeping the gap open, and then transform the boundary regions one by one. The edge states are never affected by gap-preserving bulk deformations since by assumption they are non-vanishing only in the boundary regions (otherwise we can redefine the boundary regions to be larger). A symmetry-preserving deformation at one of the edges can affect just one

of the boundary states because the other one is non-vanishing only at the opposite edge with respect to the deformation. We have argued above that if the symmetry is maintained, the edge states can leave the zero energy level only simultaneously and therefore both are required to stay put at zero under any symmetry-preserving boundary deformation. It follows that the two edge states are present in any SSH chain with $|\gamma_1| < |\gamma_2|$. One may also apply the reverse argumentation here: There are no zero energy edge states present in the extreme case $\gamma_2 = 0$ and therefore they will also be absent for all chains with $|\gamma_1| > |\gamma_2|$. The presence (absence) of the edge states signifies that the SSH chains with $|\gamma_1| > |\gamma_2|$ and $|\gamma_1| < |\gamma_2|$ are indeed topologically distinct. Moreover, these edge states are robust to any local gap-preserving perturbations that do not break the sublattice symmetry. Their protection against perturbations is a unique feature of a topological phase. In Chapter 6 we will see how this intriguing property is realized in *time-periodic* SPT phases, in fact, precisely on the example of an SSH model subject to a time-periodic drive.

2.2 The Qi-Wu-Zhang model

The Qi-Wu-Zhang (QWZ) model is arguably the simplest model of a topological band insulator in 2d [50]. In the topologically nontrivial regime this model realizes a so-called Chern insulator with chiral boundary states propagating at the boundaries. The Chern phases have properties similar to quantum Hall states, and in fact they generalize quantum Hall states to cases where applied magnetic field is not necessary.

Consider a quantum system of spin-$\frac{1}{2}$ fermions hopping on a square lattice described by the following Hamiltonian:

$$\begin{aligned} H = &\frac{1}{2}\sum_{j_x=1}^{N_x-1}\sum_{j_y=1}^{N_y}\left(|j_x+1,j_y\rangle\langle j_x,j_y|\otimes(\sigma_z+i\sigma_x)+\text{h.c.}\right) \\ &+\frac{1}{2}\sum_{j_x=1}^{N_x}\sum_{j_y=1}^{N_y-1}\left(|j_x,j_y+1\rangle\langle j_x,j_y|\otimes(\sigma_z+i\sigma_y)+\text{h.c.}\right) \\ &+u_{j_x,j_y}\sum_{j_x=1}^{N_x}\sum_{j_y=1}^{N_y}|j_x,j_y\rangle\langle j_x,j_y|\otimes\sigma_z, \end{aligned} \tag{2.5}$$

where $|j_x, j_y\rangle$ are basis states localized at the lattice sites, σ_i with $i = x, y, z$ are the Pauli matrices acting on the spin degree of freedom, and u_{j_x,j_y} is a parameter of the Hamiltonian.

Under periodic boundary conditions the homogeneous QWZ system is described by the Bloch Hamiltonian

$$H(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\sigma};\ d_{x/y}(\boldsymbol{k}) = \sin(k_{x/y});\ d_z(\boldsymbol{k}) = u + \cos(k_x) + \cos(k_y), \quad (2.6)$$

yielding the dispersion relation

$$\varepsilon(\boldsymbol{k}) = \pm\sqrt{\sin^2(k_x) + \sin^2(k_y) + (u + \cos(k_x) + \cos(k_y))^2}, \quad (2.7)$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $u_{j_x,j_y} = u$ for all sites j_x and j_y.

By looking at the dispersion relation we see that the spectrum becomes gapless at $u = 0$ and $u = \pm 2$, and stays gapped otherwise. The band gap always closes at high-symmetric momentum points: at $(k_x, k_y) = (0, 0)$ for $u = -2$, at $(k_x, k_y) = (0, \pi)$ and $(k_x, k_y) = (\pi, 0)$ for $u = 0$, and at $(k_x, k_y) = (\pi, \pi)$ for $u = 2$. All these band gap closings are accompanied with the inversion of bands pointing to possible topological phase transitions.

Formally, the topological phase diagram of the homogeneous QWZ model may be obtained by calculating the topological index for 2d band insulators, the so-called Chern number C [50]. The Chern number of a two-band system in 2d counts the number of times the vector $\boldsymbol{d}(\boldsymbol{k})$, see Eq. (2.6), wraps the origin. Explicitly, the Chern number is calculated as

$$C = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} dk_x\, dk_y\ \hat{\boldsymbol{d}}(\boldsymbol{k})\cdot(\partial_{k_x}\hat{\boldsymbol{d}}(\boldsymbol{k})\times\partial_{k_y}\hat{\boldsymbol{d}}(\boldsymbol{k})), \quad (2.8)$$

with $\hat{\boldsymbol{d}}(\boldsymbol{k}) = \boldsymbol{d}(\boldsymbol{k})/|\boldsymbol{d}(\boldsymbol{k})|$. The QWZ Hamiltonian is associated with the Chern number $C = -1$ for $-2 < u < 0$ and $C = 1$ for $0 < u < 2$. For $u < -2$ or $u > 2$ the Hamiltonian is topologically trivial, the Chern number C is zero.

By the bulk-edge correspondence the non-zero Chern number counts the number of robust chiral boundary states and there are multiple ways to rigorously prove this statement for any 2d system, see for example [42, 43, 44]. Here, however, we sketch the bulk-edge correspondence focusing only on the QWZ model. Take the QWZ Hamiltonian, Eq. (2.5), with open boundaries in direction x and periodic boundary condition in direction y with k_y being a good quantum number. In the topologically nontrivial regime, say at $u = 1$

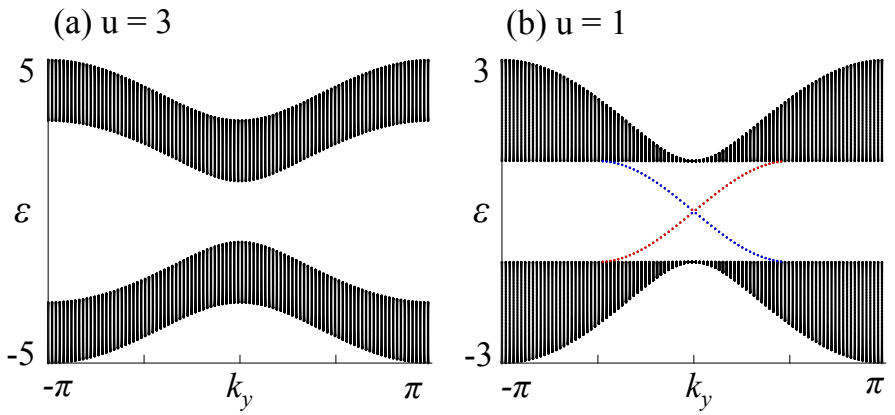

Figure 2.2: The energy spectra of the QWZ model under periodic boundary conditions in direction y and open boundaries in direction x. Here we take $N_x = 100$, $N_y = 100$, (a) $u = 3$ (trivial, $C = 0$) and (b) $u = 1$ (nontrivial, $C = 1$).

with $C = 1$, we numerically diagonalize the QWZ Hamiltonian and observe the existence of a pair of in-gap chiral modes, Fig. 2.2. The right and left propagating states, distinguished by different colors, are localized on opposite edges and therefore cannot hybridize under any bulk-gap-preserving perturbations. It follows that the chiral boundary states cannot disappear under any y-invariant disorder preserving the bulk band gap: This is simply due to continuity of the bands and the fact that the chiral modes continuously connect the valence and conduction bulk bands, and thus we need the bulk gap to close in order to destroy the boundary states. This means that a pair of chiral boundary modes is necessarily present for any value of $0 < u < 2$ and any type of y-invariant boundary perturbation. The analogous reasoning can be applied for the case $-2 < u < 0$ with $C = -1$. This type of argument can be adapted to any boundary shape, not only y-invariant (like the open boundaries in the x-direction assumed here). To see this, we consider the QWZ Hamiltonian in the thermodynamic limit, defined on a lattice of any shape and disordered at the finite boundary region. We then first adiabatically unbend a small portion of the boundary, but big enough to assume translational invariance there, and remove any disorder from that region. This process cannot affect those parts of the boundary states – if there are any – which are far from the 'cleaned' region. Now, if the bulk is topologically nontrivial, then there should be a chiral boundary mode at the

cleaned boundary because of our previous argument. The essential thing to notice here is that this state cannot simply terminate after leaving the cleaned region due to unitarity of its evolution, i.e. it has to propagate somewhere, and the gapped bulk leaves it no other option than to make a full circle around the boundary. This proves bulk-boundary correspondence for every part of the boundary except the cleaned region that was initially disordered. To show existence of the chiral state in that region as well, we simply follow the same line of reasoning by considering a boundary region far from this one. In this way the existence of a robust boundary chiral mode is established.

3 Classification of time-independent band insulators

The ideas described in Chapter 2 are very interesting and intriguing. But the selected toy models illustrating these ideas were described in a rather introductory way and many important questions were barely mentioned. Where do the topological properties come from? Why are they called topological? How are the topological indices constructed? What is their relevance to physics? We address these questions in this chapter by providing a comprehensive introduction to the classification scheme of symmetry-protected topological phases of quantum matter, here focusing only on non-interacting theories.

To build some physical insight and intuition one should first understand how geometry enters quantum theory. Maybe the most elegant and fundamental concept describing connection between quantum physics and geometry is the so-called Berry phase [55]. The Berry phase along with the Berry connection and curvature allow one to understand the logic of theories describing various topologically robust phenomena. In their original formulations these concepts are usually introduced for systems without any particular symmetries. It is less conventional, yet very fruitful, to extend these notions to apply in the presence of various symmetries [56] and even for Floquet systems [57].

3.1 Berry phase

The postulates of quantum mechanics dictate that any quantum system is entirely described by its state vector $|\psi\rangle \in \mathscr{H}$ where $\mathscr{H}$ denotes a complex Hilbert space composed of all physical states. Moreover, any state $|\psi\rangle$ is defined only up to a global complex phase factor, up to the so-called global gauge transformation, and all observable properties must be gauge invariant. The gauge invariance is a redundancy in the mathematical representation of quantum mechanics and it cannot be explicitly broken or violated under any circumstances, at least in conventional formulations of quantum theory. Even though the global phase factors alone must not have any physical significance, they may recombine into gauge invariant topologically robust

quantities after integrating them over a curve or a surface in some parameter space.

To see how this comes about, let us take a quantum system described by a smooth parametric Hamiltonian $H(\boldsymbol{r})$ with $\boldsymbol{r}=(r_1,...,r_N)$ a set of parameters living in some N-dimensional parameter space $\mathscr{P}$, which is assumed to be a smooth manifold. We also consider the Hilbert space $\mathscr{H}$ of all physical states to be independent of $\boldsymbol{r}$. It follows that the orthonormal eigenstates $|n;\boldsymbol{r}\rangle$ of the Hamiltonian $H(\boldsymbol{r})$, defined via

$$H(\boldsymbol{r})|n;\boldsymbol{r}\rangle=\varepsilon_n(\boldsymbol{r})|n;\boldsymbol{r}\rangle, \tag{3.1}$$

span the same Hilbert space $\mathscr{H}$ for any choice of $\boldsymbol{r}\in\mathscr{P}$. Also we consider the projectors $|n;\boldsymbol{r}\rangle\langle n;\boldsymbol{r}|$ to be single-valued smooth functions of $\boldsymbol{r}$. This condition is always satisfied for non-degenerate Hamiltonians $H(\boldsymbol{r})$; the generalization to the degenerate case is given in Appendix A.

Take now any directed curve $p:[0,1)\to\mathscr{P}$ and assume that the states $|n;\boldsymbol{r}\rangle$ are smooth and single-valued everywhere in a neighborhood of this curve. By continuously varying the Hamiltonian $H(\boldsymbol{r})$ to $H(\boldsymbol{r}+d\boldsymbol{r})$ the eigenstate $|n;\boldsymbol{r}\rangle$ will transform into $|n;\boldsymbol{r}+d\boldsymbol{r}\rangle$ and pick up a relative phase $\delta\gamma_n$ given by

$$e^{-i\delta\gamma_n}=\frac{\langle n;\boldsymbol{r}|n;\boldsymbol{r}+d\boldsymbol{r}\rangle}{|\langle n;\boldsymbol{r}|n;\boldsymbol{r}+d\boldsymbol{r}\rangle|}. \tag{3.2}$$

By letting $d\boldsymbol{r}\to 0$ we obtain

$$\delta\gamma_n=\boldsymbol{A}^n(\boldsymbol{r})\cdot d\boldsymbol{r};\quad \boldsymbol{A}^n(\boldsymbol{r})=i\langle n;\boldsymbol{r}|\nabla_{\boldsymbol{r}}|n;\boldsymbol{r}\rangle. \tag{3.3}$$

The quantity $\boldsymbol{A}^n(\boldsymbol{r})$, generally known as a Berry connection, is not gauge invariant. This can be seen via a direct substitution of $|n;\boldsymbol{r}\rangle'=e^{i\phi(\boldsymbol{r})}|n;\boldsymbol{r}\rangle$ into Eq. (3.3) resulting in $\boldsymbol{A}'^n(\boldsymbol{r})=\boldsymbol{A}^n(\boldsymbol{r})-\nabla_{\boldsymbol{r}}\phi(\boldsymbol{r})$. Note that $A^n(\boldsymbol{r})$ with $\boldsymbol{r}\in\mathscr{P}$ transforms similarly to the conventional electromagnetic vector potential $\vec{A}(\boldsymbol{x})$ with position $\boldsymbol{x}\in\mathscr{R}^3$: The Berry connection $A^n(\boldsymbol{r})$ is, thus, a generalization of the vector potential $\vec{A}(\boldsymbol{x})$ to an arbitrary parameter space $\boldsymbol{r}\in\mathscr{P}$.

To obtain a gauge invariant quantity one will need to integrate the acquired phase $\delta\gamma_n$ over the curve p introduced above, now assumed to be closed:

$$\gamma_n(p)=-\arg\exp\left[-i\oint_p\boldsymbol{A}^n(\boldsymbol{r})\cdot d\boldsymbol{r}\right]. \tag{3.4}$$

The phase $\gamma_n(p)$, known as a Berry phase, is a gauge invariant phase acquired by the state $|n;\boldsymbol{r}\rangle$ while continuously deforming the Hamiltonian $H(\boldsymbol{r})$ in the parameter space $\mathscr{P}$. It is a very important quantity reflecting the geometric specifications of the state space. In analogy to the theory of electromagnetism it makes sense to look for a gauge field tensor $\Omega^n_{\mu,\nu}(\boldsymbol{r})$ playing a role similar to the convectional gauge invariant magnetic field $\vec{B}(\boldsymbol{x})=\nabla\times\vec{A}(\boldsymbol{x})$. Explicitly, the gauge invariant field tensor $\Omega^n_{\mu,\nu}(\boldsymbol{r})$, generally known as a Berry curvature tensor, is defined via

$$\Omega^n_{\mu,\nu}(\boldsymbol{r})=\Big[\partial_\mu A^n_\nu(\boldsymbol{r})-\partial_\nu A^n_\mu(\boldsymbol{r})\Big], \tag{3.5}$$

where the partial derivatives are defined with respect to the parameter variable $\boldsymbol{r}$, $\partial_\mu=\partial/\partial r_\mu$. Note that the Berry curvature tensor $\Omega^n_{\mu,\nu}(\boldsymbol{r})$ is a gauge invariant quantity and it reproduces the magnetic field $\vec{B}(\boldsymbol{x})=\nabla\times\vec{A}(\boldsymbol{x})$ if we choose $\boldsymbol{r}\in\mathscr{P}$ to represent the position $\boldsymbol{x}\in\mathscr{R}^3$.

By straightforward transformations of Eqs. (3.1-3.5) one can represent the Berry curvature tensor $\Omega^n_{\mu,\nu}(\boldsymbol{r})$ in the following form:

$$\Omega^n_{\mu,\nu}(\boldsymbol{r})=i\sum_{m\neq n}\frac{\langle n;\boldsymbol{r}|\partial_\mu H(\boldsymbol{r})|m;\boldsymbol{r}\rangle\langle m;\boldsymbol{r}|\partial_\nu H(\boldsymbol{r})|n;\boldsymbol{r}\rangle-(\mu\longleftrightarrow\nu)}{[\varepsilon_n(\boldsymbol{r})-\varepsilon_m(\boldsymbol{r})]^2}. \tag{3.6}$$

Different from the Berry connection, the Berry curvature represented by Eq. (3.6) does not require the states $|n;\boldsymbol{r}\rangle$ to be smooth or single-valued which allows one to painlessly define $\Omega^n_{\mu,\nu}(\boldsymbol{r})$ everywhere in $\mathscr{P}$. Moreover, by applying Stoke's theorem[1] to Eq. (3.4) we obtain a relation between the Berry

[1] By employing basic terminology of differential geometry, the Stoke's theorem reads as $\int_{\partial S}\omega=\int_S d\omega$, where ω is a differential form, $d\omega$ denotes its exterior derivative, and ∂S is the boundary of some manifold S. In our case we take $\omega=A^n_\nu d r_\nu$ and $\partial S=p$ resulting in the correspondence between Eqs. (3.4) and (3.7).

phase and the Berry curvature[2]

$$\gamma_n(p) = -\arg\exp\left[-\frac{i}{2}\int_{S_p}\Omega^n_{\mu,\nu}(\boldsymbol{r})\,d\,r_\mu\wedge d\,r_\nu\right], \tag{3.7}$$

where $S_p \subset \mathscr{P}$ denotes a surface enclosed by the curve p and the indices μ, ν are contracted according to the Einstein's convention of summation on repeated indices. Here $\wedge$ is the usual wedge product, and $d\,r_\mu \wedge d\,r_\nu$ represents the area of an infinitesimal oriented parallelogram defined by $d\,r_\mu$ and $d\,r_\nu$. Note that for $\dim(\mathscr{P}) = 3$ the formula reduces to

$$\gamma_n(p) = -\arg\exp\left[-i\int_{S_p}\boldsymbol{B}^n\cdot d\,\boldsymbol{S}\right], \tag{3.8}$$

where $\boldsymbol{B}^n(\boldsymbol{r}) = \nabla\times\boldsymbol{A}^n(\boldsymbol{r})$ and the Berry phase $\gamma_n(p)$ reproduces the well-known Aharonov–Bohm phase [58].

The Berry phase can formally be identified as a *holonomy* of the principle fiber bundle defined by the states $\lambda^n_{\boldsymbol{r}} = \{|\psi\rangle\in\mathscr{H} : |\psi\rangle = e^{i\phi}|n;\boldsymbol{r}\rangle, \phi\in[0,2\pi]\}$ over the parameter space $\boldsymbol{r}\in\mathscr{P}$ [59]. In short, a fiber bundle over a base manifold M is a collection of manifolds F_x, so-called fibers, assigned to each $x\in M$ that are 'glued' together with a requirement to be locally diffeomorphic to a Cartesian product $U_y\times F$ where U_y is an open cover of M. All fibers are required to be diffeomorphic to some manifold F, called the typical fiber. A principle fiber bundle is a fiber bundle with a property $F\simeq G$ where G denotes the structure group of F, i.e. a Lie group of diffeomorphisms of F. The sets of states $\lambda^n_{\boldsymbol{r}}$ over $\mathscr{P}$ define a fiber bundle that is a principle fiber bundle because $\lambda^n_{\boldsymbol{r}}\simeq U(1)$ and $G\simeq U(1)$. The holonomy quantifies parallel transport around any closed loop: In this particular case it can be shown that after completing a loop on the bundle the starting and final states will differ by an element of the structure group $G\simeq U(1)$, i.e. by a phase. This phase

[2] For the Stoke's theorem to be applicable we have to request the states $|n;\boldsymbol{r}\rangle$ to be smooth and single-valued everywhere in S_p, which is not always the case! This problem can be resolved by partitioning the surface S_p into small regions and fixing a gauge independently for each region to make the states $|n;\boldsymbol{r}\rangle$ smooth and single-valued within each distinct region. The Stoke's theorem can then be applied separately for each of the regions and the outcomes summed up. The Berry curvature tensor $\Omega^n_{\mu,\nu}(\boldsymbol{r})$ is gauge invariant and therefore Eq. 3.7 will not be affected.

is precisely the Berry phase $\gamma_n(p)$ defined above. Note that each parallel transport along a curve is determined by some local *connection* that formally defines the 'bending' of the fibers. In the considered case the connection turns out to be exactly the Berry connection $\boldsymbol{A}^n(\boldsymbol{r})$. Also, for any parallel transport one may define a so-called *curvature* that in general represents a holonomy associated with an infinitesimal loop: The curvature is a tensor with elements describing parallel transport over infinitesimal parallelograms. The Berry curvature defines a curvature for the parallel transport associated with the Berry phase $\gamma_n(p)$. Thus, the Berry phase, Berry connection, and Berry curvature can be formalized within the theory of parallel transport on fiber bundles.

To sum up, geometric phases in quantum physics pop up naturally as a result of continuous transport of the Hamiltonian eigenstates controlled by some parameters $\boldsymbol{r} \in \mathscr{P}$. These phases specify geometric features of the state space, and, to be shown below, define topologically robust quantities like a Chern number C.

3.2 Homotopy-based classification

Recall that within non-interacting theories the description of fermions in a periodic potential reduces to solving the effective one-body Schrödinger equation for each point inside the first Brillouin zone (BZ). This means that for every momentum $\boldsymbol{k}$ inside the BZ we construct a so-called Bloch Hamiltonian $H(\boldsymbol{k})$ which completely determines the eigenstate solutions. From a mathematical point of view we formulate this as a map that sends every $\boldsymbol{k} \in BZ \to H(\boldsymbol{k})$. Thus, for every momentum $\boldsymbol{k} \in BZ$ we obtain the eigenvalue problem

$$H(\boldsymbol{k})|u_n(\boldsymbol{k})\rangle = \varepsilon_n|u_n(\boldsymbol{k})\rangle, \tag{3.9}$$

where the momentum $\boldsymbol{k}$ acts as a parameter and the index n corresponds to the n-th energy band solution.

Let us assume that there are only non-degenerate bands. By treating the momentum space as a parameter space $\mathscr{P}$, every Bloch state $|u_n(\boldsymbol{k})\rangle$ acquires a Berry phase after parallel transporting the state over any closed curve in the momentum space. Note that in two dimensions the BZ is a closed surface, homeomorphic to the two-dimensional torus $\mathbb{T}^2$. Thus, in two dimensions the surface integral of the Berry curvature over the entire BZ must be an

integer multiple of 2π because the Berry phase over an infinitesimal loop must be zero. Explicitly,

$$C_n = \frac{1}{2\pi}\int_{BZ} dk_x dk_y\, \Omega^n_{x,y}(\boldsymbol{k}); \quad \Omega^n_{x,y}(\boldsymbol{k}) = \frac{\partial A^n_y(\boldsymbol{k})}{\partial k_x} - \frac{\partial A^n_x(\boldsymbol{k})}{\partial k_y}, \tag{3.10}$$

where $\boldsymbol{A}^n(\boldsymbol{k}) = i\langle u_n(\boldsymbol{k})|\nabla|u_n(\boldsymbol{k})\rangle$. The integer number C_n is called the (first) Chern number [60]. The generalization to degenerate bands is simply done by replacing the Abelian Berry curvature $\Omega^n_{x,y}(\boldsymbol{k})$ in Eq. (3.10) by the trace of its non-Abelian analogue defined in Appendix A [36]. Note that if the bands do not cross then the multiband Chern number reduces to a sum of Chern numbers corresponding to individual bands. The Chern number corresponding to the occupied space of a 2d band insulator is clearly a topologically robust quantity, it cannot change unless a gap-closing deformation is applied. One should here note that in the case of a two-band system, a simple transformation of the eigenstates allows the expression in (3.10) to be rewritten in the form of Eq. (2.8).

The Chern number is a topologically robust quantity, i.e. a topological index, but how general is its construction? Are there other quantities that are topologically protected? If so, how many different topological phases do they define? Those are the key questions addressed within the classification theory of topological phases. Recall that many-body ground states are said to be *topologically equivalent*, i.e. represent the same SPT phase, if they can be continuously transformed into each other without breaking the underlying symmetries. The topological equivalences then establish equivalence classes of many-body states and the classification schemes quantify how many different equivalence classes, i.e. phases, exist within a particular symmetry class. Over the years there have been proposed multiple analytic classification protocols applying various state-of-the-art theories such as twisted equivariant topological K-theory [13, 14, 15, 16], cobordism cohomology [17, 18], GSO projection [19], and others. The classification of topological quantum phases within non-interacting theories can be efficiently addressed using standard homotopy theory [60] supplemented by K-theory [36]. Here we briefly outline the key ideas and results of this approach.

The topological classification of topological quantum phases of noninteracting fermions on a lattice is defined by classifying many-body ground states – built from the single-particle Bloch states corresponding to the occupied energy bands – and therefore shall not explicitly depend on the single-particle

energies. It is convenient then to flatten the energy spectrum and by this define the topologically equivalent Q-matrix representation. The Q-matrix is defined as follows

$$Q(\boldsymbol{k}) = 2P(\boldsymbol{k}) - I; \quad P(\boldsymbol{k}) = \sum_{n=1}^{\mathrm{m}} |u_n(\boldsymbol{k})\rangle\langle u_n(\boldsymbol{k})|. \tag{3.11}$$

where $P(\boldsymbol{k})$ is the projection operator on the filled bands, and with m indexing the highest filled band. Clearly, by continuously varying the Q-matrix we continuously vary the many-body ground state and the reverse statement is also true. It follows that the ground states share the same topological properties as the Q-matrices.

It is important to stress that the Q-matrix has a direct connection to the Hamiltonian and as a consequence it must respect all symmetries of the system. As was mentioned above, we may use the structure of the Q-matrices as a characteristic property of the topological classification of phases and this is done using the following idea. Let us first consider generic quantum systems without putting any symmetry restrictions on them. Within the classification just outlined we separate all maps $Q(\boldsymbol{k}) : \mathbb{T}^d \to \mathbb{Y}$ into equivalence classes, where $\mathbb{T}^d$ is a d-dimensional torus and $\mathbb{Y}$ denotes the space of all Q-matrices, and label such classes by distinct indices (or collection of indices). We place two different maps into the same class if and only if they can be continuously transformed into each other. Thus, the states associated with inequivalent maps $Q(\boldsymbol{k})$ can be continuously deformed into each other only by moving outside of the initially assumed topological space, i.e. via a gap-closing phase transition.

In topology, the continuous maps $Q(\boldsymbol{k})$ are called homotopies, and the corresponding sets of equivalent objects are called homotopy classes denoted by $[\mathbb{T}^d, \mathbb{Y}]$. The set $[\mathbb{T}^d, \mathbb{Y}]$ is then usually broken up into $[\mathbb{S}^d, \mathbb{Y}] \subset [\mathbb{T}^d, \mathbb{Y}]$, where $\mathbb{S}^d$ represents a d-dimensional sphere, and some remaining piece [61]. The term $[\mathbb{S}^d, \mathbb{Y}]$ is generally known as a homotopy group, commonly denoted by $\pi_d(\mathbb{Y})$, and it defines so-called strong topological indices. The remaining piece, generally non-trivial, reflects the fact that the base torus $\mathbb{T}^d$ is by itself topologically nontrivial, in other words, it is not equivalent to a d-dimensional sphere $\mathbb{S}^d$. This remaining piece of $[\mathbb{T}^d, \mathbb{Y}]$ is usually composed of low-dimensional homotopy groups $\pi_c(\mathbb{Y})$ with $c < d$ and defines so-called weak topological indices [61]. The strong topological indices $[\mathbb{S}^d, \mathbb{Y}]$ are of

more fundamental importance: They highlight the intrinsic nontriviality of the physical space $\mathbb{Y}$. For both types of indices one may establish the bulk-edge correspondence, however, only in strong topological phases the boundary states are generically robust against local perturbations, and by this are probably of higher importance for applications. It must be noted though that the boundary states in weak topological insulators may also exhibit robustness to some extent, however, this feature is not generic [63, 64]. The space of all Q-matrices corresponding to band insulators with a total of M bands and m occupied bands is isomorphic to the complex Grassmannian $G_{m,M}(\mathscr{C}) \simeq U(M)/U(m)\times U(M-m)$, where $U(x)$ denotes the unitary group, and for large M and m the homotopy groups $\pi_d(\mathbb{G}_{m,N}(\mathscr{C}))$ are known to be trivial for odd dimensions, and isomorphic to $\mathbb{Z}$ for even dimensions[36]. The integer topological number in d = 2 is exactly the Chern number discussed earlier.

By restricting the maps $Q(\boldsymbol{k})$ to respect some symmetries, the topological equivalence classes can change: The symmetries here come as additional constraints on the allowed continuous deformations of $Q(\boldsymbol{k})$ and two initially topologically equivalent states can become inequivalent under such symmetry restrictions and therefore lie in different equivalence sectors within the particular symmetry class. The symmetries, thus, enrich the formalism of topologically nontrivial matter and play a crucial role in establishing the so-called SPT phases. Generally speaking, a symmetry transformation can be thought of as a change of coordinates in the Hilbert space that leaves the system invariant. Wigner's theorem [65] then implies that any symmetry transformation may be associated with a unitary or an antiunitary operator defined up to a phase factor. Before proceeding with the standard classification scheme of SPT phases, we shall discuss the symmetries in a more structural way and this is done below.

3.3 Symmetries of quadratic Hamiltonians

Given a lattice model, the relevant symmetries are usually differentiated as either on-site or crystalline depending on their action on the spatial degrees of freedom [36]. The on-site symmetries can be represented as a product of transformations acting on individual lattice sites. These symmetries compose ten irreducible symmetry classes that are commonly considered within

the standard classification schemes. The crystalline symmetries extend the classification to a broader scheme [66], that is less frequently discussed because the corresponding many-body states are generically more fragile to perturbations. In the following we only consider on-site symmetries and leave out the generalization to crystalline symmetries.

It is convenient to introduce different symmetries into the theory by their action on the creation $\hat{c}_i^\dagger$ and annihilation $\hat{c}_i$ operators, where the index i sweeps through the whole Hilbert space. To preserve the fermionic anticommutation relations the symmetry transformations are required to be unitary or antiunitary.

3.3.1 Conventional unitary symmetries

Wigner's theorem[65] allows to have unitary symmetries $\hat{\mathscr{U}}$ defined by the following actions on the creation and annihilation operators

$$\hat{\mathscr{U}}\,\hat{c}_i\,\hat{\mathscr{U}}^{-1} = (U)_{ij}\,\hat{c}_j \qquad\qquad \hat{\mathscr{U}}\,i\,\hat{\mathscr{U}}^{-1} = i \tag{3.12}$$

for some unitary matrix U. The system is then said to be symmetric under the transformation $\hat{\mathscr{U}}$ if this operator commutes with the Hamiltonian. Here we focus only on non-interacting fermionic systems. Any Hamiltonian in absence of interactions can be represented as

$$\hat{H} = \sum_{ij} \hat{c}_i^\dagger\, H_{ij}\, \hat{c}_j, \tag{3.13}$$

where H is the first-quantized Hamiltonian corresponding to $\hat{H}$. Clearly, $\hat{\mathscr{U}}$ commutes with $\hat{H}$ if only if U commutes with H, which leads to

$$H = U^\dagger H U. \tag{3.14}$$

It directly follows that the matrices H and U can be simultaneously diagonalized, meaning that any Hamiltonian satisfying Eq. (3.14) can be transformed into a block-diagonal form by employing conveniently chosen basis states. The classification of systems respecting such unitary symmetries is then not fundamentally interesting because it can be reconstructed from the classifications of simpler symmetry classes. Thus, the symmetry classes associated with at least one such symmetry are considered to be *reducible* and are left our from the classification scheme intended to capture only fundamentally different scenarios.

3.3.2 Time - reversal symmetry

Time-reversal $\hat{\mathscr{T}}$ is an antiunitary operator defined by the following transformations

$$\hat{\mathscr{T}}\,\hat{c}_i\,\hat{\mathscr{T}}^{-1} = (U_T)_{ij}\,\hat{c}_j \qquad\qquad \hat{\mathscr{T}}\,i\,\hat{\mathscr{T}}^{-1} = -i \tag{3.15}$$

for some unitary matrix U_T. The time-reversal symmetry then implies that $\hat{H}$ and $\hat{\mathscr{T}}$ commute, explicitly leading to

$$H = (U_T)^\dagger H^* U_T \tag{3.16}$$

By applying the time-reversal operation $\hat{\mathscr{T}}$ twice we obtain the relation $(U_T^* U_T)^\dagger H (U_T^* U_T) = H$. Now, we consider that the first-quantized Hamiltonians H with time-reversal symmetry (TRS) form an irreducible representation of the symmetry group generated by TRS. Otherwise the corresponding matrices would all be block diagonal in some basis and we can then look at each block separately, by this switching to the fundamental case from which all other ones can be reconstructed. It follows that in such an irreducible case $U_T^* U_T$ is equal to the identity matrix up to a multiplicative phase according to Schur's first lemma [67], i.e. $U_T^* U_T = \alpha \cdot I$ with $|\alpha| = 1$ due to unitarity. From a physical point of view this statement simply reflects the fact that the time-reversal operation applied twice should do nothing to the system. To be more precise by this we mean that under $(\hat{\mathscr{T}})^2$ a physical state can only pick up some non-measurable phase, which directly leads to the desired relation $U_T^* U_T = \alpha \cdot I$ where $|\alpha| = 1$. It immediately follows that $(U_T)^\dagger = \alpha U_T$ and therefore $\alpha^2 = 1$. As a consequence, there are only two possible options for a phase factor, $\alpha = \pm 1$. In the case of $\alpha = +1$ the time-reversal transformation applied twice does nothing to the state. On the other hand, in the case of $\alpha = -1$ it adds an extra minus sign. These two cases bring completely different physics to the systems. For example, the Kramer degeneracy theorem can be applied only for systems symmetric under $\mathscr{T}$ with $\alpha = -1$. We denote these two different time-reversal symmetries by TRS-1 and TRS+1 respectively.

3.3.3 Particle - hole symmetry

Particle-hole symmetry (PHS), sometimes called charge-conjugation symmetry, may be defined within the same framework as was done for TRS. A

transformation $\hat{\mathscr{C}}$ is said to be a particle-hole transformation if the following relations are satisfied

$$\hat{\mathscr{C}}\,\hat{c}_i\,\hat{\mathscr{C}}^{-1} = (U_C)_{ij}\,\hat{c}_j^\dagger \qquad\qquad \hat{\mathscr{C}}\,i\,\hat{\mathscr{C}}^{-1} = -i \tag{3.17}$$

for some unitary matrix U_C. The system is then invariant under PHS if the Hamiltonian commutes with $\hat{\mathscr{C}}$ leading to the following relation between the first-quantized Hamiltonian H and the unitary matrix U_C:

$$H = -(U_C)^\dagger H^* U_C \tag{3.18}$$

In other words, we say that system is symmetric under a particle-hole transformation if and only if Eq. (3.18) is satisfied. By applying exactly the same arguments as in the case of TRS, we get $U_C^* U_C = \pm I$. Depending on the phase, we refer to these two different particle-hole symmetries as PHS+1 and PHS-1.

One may here mention that PHS is a key element of the BCS theory of superconductivity. The ground state of a BCS superconductor is gapped by the superconducting gap and in this sense similar to that of an insulator. To describe a state within the BCS formalism we usually use the so-called Bogoliubov-de Gennes equations, an analogue of the single-particle Schrödinger equation acting on Nambu spinors (states given in a particle-hole vector form). It follows that the corresponding Bogoliubov-de Gennes Hamiltonian H_{BdG} may be treated on the same footing as the first-quantized Hamiltonian H. It is crucial to point out that the matrix H_{BdG} of any superconductor by construction respects PHS. Without going into much detail this can be proven by showing that each solution to the BdG equations has a particle-hole counterpart. This symmetry is PHS+1 because the corresponding transformation applied twice does not produce any extra minus sign. It is important to notice that the symmetry of H_{BdG} under a particle-hole transformation is a consequence of the BCS model and cannot be broken under any circumstances.

3.3.4 Chiral symmetry

The idea behind chiral symmetry (CS) is straightforward: The chiral transformation $\hat{\mathscr{S}}$ is defined to be a combination of the time-reversal transformation $\hat{\mathscr{T}}$ and particle-hole transformation $\hat{\mathscr{C}}$. More precisely, if a Hamiltonian $\hat{H}$ is symmetric under $\hat{\mathscr{T}}$ and $\hat{\mathscr{C}}$ simultaneously then it is also symmetric under

the unitary transformation $\hat{\mathscr{S}} = \hat{\mathscr{T}}\hat{\mathscr{C}}$. The reverse statement is not true: a Hamiltonian may not separately commute with both $\hat{\mathscr{T}}$ and $\hat{\mathscr{C}}$ but at the same time might commute with $\hat{\mathscr{S}}$. It is also worth to mention that if the Hamiltonian $\hat{H}$ is symmetric under a chiral transformation and one of the TRS or PHS transformations, then it is symmetric under all of these three operations.

The first-quantized Hamiltonian H is then required to fulfill the following condition

$$H = -(U_S)^\dagger H U_S, \tag{3.19}$$

where $U_S = U_T^* U_C$. Applying the same argumentation as before we deduce that $(U_S)^2 = \alpha \cdot I$ with $|\alpha| = 1$. The symmetry operations U_S are defined up to a phase factor that has no physical meaning. Therefore, without loss of generality we may always shift the global phase of U_S by $\sqrt{\alpha}$ and simply obtain that $(U_S)^2 = I$. Note that this trick would not work in case of TRS or PHS because even though the analogous relations are very similar, they involve complex conjugation that spoils the argument.

For every U_S we may define projection operators $P_A = (I + U_S)/2$ and $P_B = (I - U_S)/2$. It follows that in the basis that diagonalizes U_S (and therefore P_A and P_B), the first-quantized Hamiltonian H becomes off-block-diagonal

$$H = \begin{bmatrix} 0 & H_{AB} \\ H_{BA} & 0 \end{bmatrix} \tag{3.20}$$

with blocks $P_X H P_Y = H_{XY}$ with $X \in \{A, B\}$ and $Y \in \{A, B\}$. The CS then results in $H_{AA} = H_{BB} = 0$. The projectors P_A and P_B can be seen as projectors onto some abstract sublattices A and B and the chiral symmetry then restricts the Hamiltonian H to contain only elements that couple these two sublattices. That is why the chiral symmetry is also sometimes called a sublattice symmetry. In fact, the sublattice symmetry of the SSH model, discussed in Chapter 2, can also be represented in the form of Eq. (3.19) by using the sublattice symmetry operator $U_S = \sum_j \big(|A, j\rangle\langle A, j| - |B, j\rangle\langle B, j|\big)$.

3.4 The ten-fold way of topological classification

Recall that the symmetry classes associated with at least one symmetry U of type $H = U^\dagger H U$ are not fundamental and can be described by combining classifications of simpler symmetry classes. Moreover, within the irreducible

picture one can have only one type of TRS, PHS, and CS. This is simply because a composition of two symmetries of the same type results in a reducible unitary symmetry. It follows that there are in total only ten irreducible symmetry classes, the so-called Altland-Zirnbauer (AZ) classes [68], and all of them are collected in Table 3.1. It is also important to note that the cases with only two of PHS, TRS and CS are missing. This is because the presence of two of these symmetries automatically requires the presence of the third one. In passing, the names used in the table is the standard way of labeling the symmetry classes in mathematics introduced by Cartan [69, 70].

All gapped Bloch Hamiltonians $H(\boldsymbol{k})$, or equivalently their flattened representations $Q(\boldsymbol{k})$, are then classified via homotopy equivalences defined within a particular symmetry class: We say that two systems are topolog-

Name	AZ	TRS	PHS	CS
Standard	A (unitary)	0	0	0
(Wigner-Dyson)	AI (orthogonal)	+1	0	0
	AII (symplectic)	-1	0	0
Chiral	AIII (chiral unitary)	0	0	1
(sublattice)	BDI (chiral orthogonal)	+1	+1	1
	CII (chiral symplectic)	-1	-1	1
BdG	D	0	+1	0
	C	0	-1	0
	DIII	-1	+1	1
	CI	+1	-1	1

Table 3.1: In this table all possible symmetry classes associated with the presence or absence of TRS, PHS and CS are collected. The value "0" represents the absence of the corresponding symmetry, the number "1" indicates its presence. Moreover, we distinguish the symmetries TRS+1, TRS-1, PHS+1 and PHS-1 by assigning a plus or a minus sign to the number "1" in the corresponding TRS and PHS columns.

ically equivalent if they can be continuously transformed into each other without breaking the symmetries of the corresponding symmetry class. The formal classification is usually conducted employing the language of K-theory [13, 14, 36, 61]. The essential idea here is to notice that Bloch Hamiltonians H within the same symmetry class form a commutative monoid[3] under a direct sum operation $\oplus$ defined via

$$H_1 \oplus H_2 = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} \tag{3.21}$$

where the explicit momentum dependence was dropped for the clarity of notation. The direct sum operation $\oplus$ can then be transferred to the homotopy equivalence classes by introducing a notion of so-called *stable* topological equivalence. The stable equivalence is basically a homotopy equivalence up to the addition of a number of trivial atomic bands and it allows a topological comparison of H_1 and H_2 even if their matrix representations have different dimensions. This definition is often physically motivated by the fact that within an effective model we often neglect the inner atomic shells but in practice they are always present. Let us denote the corresponding stable equivalence classes by $[H]$ with the inherited operation $[H_1]\oplus[H_2]=[H_1\oplus H_2]$. Note that $[H]\oplus[-H]=[\emptyset]$ (trivial atomic class) because the eigenspace below the gap of $H\oplus -H$ consist of all eigenstates of H, and therefore it can be continuously trivialized. Thus, the stable equivalence classes of Bloch Hamiltonians form a group under the direct sum operation $\oplus$. The group structure of the equivalence classes $[H]$ is essential here: It can be shown that the group of stable equivalence classes within a particular symmetry class is isomorphic to a certain K-group [13, 14, 36, 61]. Here we skip the details of the derivations of these K-groups and present only the result.

In Table 3.2 all homotopy (stable) groups corresponding to strong topological indices, distinct symmetry classes, and different dimensions of the BZ are presented [36]. Analyzing the table we notice that all non-trivial groups are isomorphic[4] either to the space of integer numbers $\mathbb{Z}$, or to the cyclic group of two elements $\mathbb{Z}_2$. Thus, in the case when the homotopy group is

[3] A monoid is a pair $(S,\oplus)$ satisfying the associativity and identity element axioms.

[4] In the literature, in addition to two nontrivial equivalence groups $\mathbb{Z}_2$ and $\mathbb{Z}$, one sometimes may encounter the group $2\mathbb{Z}$, see for example Ref. [36]. However, the groups $\mathbb{Z}$ and $2\mathbb{Z}$ are isomorphic and here we denote both of them simply by $\mathbb{Z}$.

isomorphic to $\mathbb{Z}$ we may assign an integer number to every system depending on the equivalence class it belongs to, a so-called integer topological invariant. Similar considerations can also be applied to the case when the group is isomorphic to $\mathbb{Z}_2$. The only difference is that in this case the topological invariant is given by 0 or 1 instead of $n \in \mathbb{Z}$, a so-called $\mathbb{Z}_2$-invariant, introduced first in 2005 by Kane and Mele [71]. It turns out that a very similar topological classification of bands can be constructed also for time-periodic systems. The weak topological indices are not presented here because they can be constructed by looking at strong topological invariants in lower dimensions. In this context, the weak topological indices are not fundamental and in the following we focus on the strong invariants only, from now on called simply topological invariants.

Before we conclude this chapter, let us briefly outline a less formal but still effective strategy to derive essentially the same classification as in Ta-

Name	Cartan	d=1	d=2	d=3
Standard	A (unitary)	-	$\mathbb{Z}$	-
(Wigner-Dyson)	AI (orthogonal)	-	-	-
	AII (symplectic)	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$
Chiral	AIII (chiral unitary)	$\mathbb{Z}$	-	$\mathbb{Z}$
(sublattice)	BDI (chiral orthogonal)	$\mathbb{Z}$	-	-
	CII (chiral symplectic)	$\mathbb{Z}$	-	$\mathbb{Z}_2$
BdG	D	$\mathbb{Z}_2$	$\mathbb{Z}$	-
	C	-	$\mathbb{Z}$	-
	DIII	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$
	CI	-	-	$\mathbb{Z}$

Table 3.2: The table of homotopy groups associated with the symmetry classes and dimensions of the BZ. The sign '-' denotes the trivial group. The symbols $\mathbb{Z}$ and $\mathbb{Z}_2$ represent the groups isomorphic to the group of integer numbers and the cyclic group of two elements respectively.

ble 3.2. The idea is to focus only on gapped Dirac Hamiltonians and study their behavior under various symmetry-preserving continuous transformations. Within this perspective the classification basically reduces to finding homotopically inequivalent mass terms corresponding to the gapped Dirac Hamiltonians. Here, the mass term is defined as any matrix that anticommutes with all Dirac kinetic matrices present in the Dirac Hamiltonian and therefore it generically gaps out its eigenspectrum. It turns out that the Dirac kinetic matrices and irreducible AZ symmetries generate either a complex or real Clifford algebra.[5] The mass term then extends this algebra to a higher dimension. Thus, the classification problem then translates into finding topologically inequivalent ways to extend the Clifford algebras by one more generator. One then obtains the same homotopy indices as presented in Table 3.2 [36].

[5] A complex Clifford algebra Cl_n is generated by complex matrices $\{e_1, ..., e_n\}$ which satisfy $\{e_i, e_j\} = 2\delta_{ij}$. A real Clifford algebra $Cl_{p,q}$ is generated by real matrices $\{e_1, ..., e_{p+q},\}$ fulfilling $\{e_i, e_j\} = 0$ for $i \neq j$, $e_i^2 = -I$ for $i \leq p$, and $e_i^2 = I$ for $i > p$.

4 Floquet theory

Here we review some basics of the Floquet formalism[1] laying the ground for Chapters 5-6 where the Floquet theory is explicitly applied for studying periodically-driven SPT phases and their properties.

Our goal here is twofold: 1) To present an easily accessible introduction to Floquet formalism and 2) Introduce two equivalent techniques for doing calculations, namely by computing time evolution operators and by employing the so-called Floquet-Sambe construction. Both these calculation methods are very important and commonly used for different classes of periodic drives. For example, time-periodic systems obtained by quenching parameters periodically in time are usually described using the evolution operators. On the other hand, harmonically driven systems are easier to represent through the Floquet-Sambe construction. We present both techniques below. To begin with, let us take a look at the time-periodic Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle, \tag{4.1}$$

where $\hat{H}(t) = \hat{H}(t+T)$ for some period T and all times t. Eq. (4.1) has the form of a Floquet differential equation and therefore the Floquet theorem must be satisfied [72]. The theorem states that there exists a complete set of solutions $|\psi_n(t)\rangle$ to Eq. (4.1) of the form

$$|\psi_n(t)\rangle = |u_n(t)\rangle e^{-i\varepsilon_n t/\hbar}, \tag{4.2}$$

where ε_n is time-independent and $|u_n(t)\rangle = |u_n(t+T)\rangle$. However, in its general form the theorem does not imply ε_n to be real and does not provide any information on orthogonality of the basis states. We solve this issue in an elegant way by noticing the following property:

$$\hat{U}(t+T,t)|u_n(t)\rangle = e^{-i\varepsilon_n T/\hbar}|u_n(t)\rangle, \tag{4.3}$$

with the time-evolution operator $\hat{U}(t',t)$ defined by

$$|\psi_n(t')\rangle = \hat{U}(t',t)|\psi_n(t)\rangle. \tag{4.4}$$

[1] Chapters 4 and 5 are mainly based on material from Ref. [62].

The evolution operator is unitary and eigenvalues of any unitary operator have modulus 1. Analyzing Eq. (4.3) we conclude that ε_n are real. In the literature they are called quasienergies. Moreover, $|u_n(t)\rangle$ are eigenmodes of $\hat{U}(t+T,t)$ and therefore they are orthogonal to each other for all t. We will call $|u_n(t)\rangle$ Floquet (steady) modes and $|\psi_n(t)\rangle$ will be called Floquet (steady) states [2].

It is important to note that quasienergies ε_n and Floquet modes $|u_n(t)\rangle$ are not uniquely defined for distinct $|\psi_n(t)\rangle = |u_n(t)\rangle e^{-i\varepsilon_n t/\hbar}$. Indeed, if we define

$$|u'_n(t)\rangle = |u_n(t)\rangle e^{im\Omega t}, \qquad \varepsilon'_n = \varepsilon_n + m\hbar\Omega \tag{4.5}$$

for some integer m then $|\psi_n(t)\rangle = |u'_n(t)\rangle e^{-i\varepsilon'_n t/\hbar}$ with unchanged property $|u'_n(t)\rangle = |u'_n(t+T)\rangle$. This implies that there are different Floquet modes representing the same physical Floquet state. To avoid this redundancy we may restrict our attention to a particular zone of quasienergies associated with an arbitrary real constant E_0,

$$E_0 - \hbar\Omega/2 < \varepsilon_n \leq E_0 + \hbar\Omega/2. \tag{4.6}$$

In this way we can characterize the physically different states $|\psi_n(t)\rangle$ by the uniquely defined quasienergies and corresponding Floquet modes.

4.1 Time-evolution operators within Floquet formalism

Time-evolution operators are key objects in the Floquet formalism. Indeed, we have already seen the relation between the time-evolution operators $\hat{U}(t+T,t)$ and Floquet steady modes $|u_n(t)\rangle$, cf. Eq. (4.3). In general, any time-evolution operator $\hat{U}(t_2,t_1)$ of a time-dependent system is connected to the corresponding Hamiltonian $\hat{H}(t)$ through the relation

$$\hat{U}(t_2,t_1) = \mathscr{T}\exp\left[-i\int_{t_1}^{t_2} dt\, \hat{H}(t)\right], \tag{4.7}$$

[2] Both $|\psi_n(t)\rangle$ and $|u_n(t)\rangle$ are often called by the same name in literature, see for example [48, 73, 74, 75, 76, 77]. Here we give them different names for avoiding confusion.

where $\mathscr{T}$ denotes the conventional time-ordering operator. Written out explicitly, the evolution operator reads as

$$\hat{U}(t_2,t_1)=\sum_n(-i)^n\int_{t_1}^{t_2}d\tau_1\ldots\int_{t_1}^{\tau_{n-1}}d\tau_n\,\hat{H}(\tau_1)\ldots\hat{H}(\tau_n). \tag{4.8}$$

It follows that for periodically driven systems, $\hat{H}(t)=\hat{H}(t+T)$, the time-evolution operators satisfy $\hat{U}(t+T,t)=\hat{U}(t+2T,t+T)$ and therefore they obey a more general relation

$$\hat{U}(t_0+nT,t_0)=\hat{U}(t_0+T,t_0)^n \tag{4.9}$$

for every integer n and any t_0, retrieved using the factorization property of time-evolution operators, $\hat{U}(t_2,t_1)=\hat{U}(t_2,t_3)\hat{U}(t_3,t_1)$. It is then useful to separate the entire time evolution into so-called stroboscopic and non-stroboscopic parts [78]. The stroboscopic evolution takes a state from time t_0 to time t_0+nT with integer n. It addresses evolution of a time-periodic system in discrete frames of time separated by the period T of the driving and specified by the starting time t_0. The non-stroboscopic part is the remaining piece of the time evolution: it corresponds to the finer details skipped by the stroboscopic evolution. The stroboscopic evolution can by itself be very helpful, for example, it can evolve a state to infinite (large) time. Also, one may retrieve the quasienergy band structure from it by simply exploiting the fact that Floquet steady modes are eigenstates of the stroboscopic evolution operators, cf. Eq. (4.3). In many cases the quasienergies are of central importance, for example, when looking for the existence of the midgap boundary states in Floquet topological band insulators and verifying that they stay robust against perturbations. This is one of the ways to conclude that the system is topologically nontrivial. However, for a complete topological classification we are usually interested in bulk properties, and for this, monitoring of the full time dependence of states is necessary. Indeed, as we have already discussed, the topological characterization is done on the many-body states, on their ability to be continuously transformed into each other, not on energy or quasienergy spectra alone. If we would look at just stroboscopic evolution then we would lose a large amount of information about the time-periodic states. On the other hand, the set of non-stroboscopic evolution operators $\hat{U}(t_2,t_1)$ with $t_2-t_1\in[0,T)$ uniquely define Floquet modes $|u_n(t)\rangle$ at all evolution times t and therefore they should be enough for retrieving topological properties.

Let us focus on the stroboscopic evolution for a moment. It is then useful to introduce the so-called stroboscopic Floquet Hamiltonian $\hat{H}_F[t_0]$ for every fixed time t_0 [78]:

$$\hat{U}(t_0+T,t_0)=\exp\left[-i\hat{H}_F[t_0]T/\hbar\right], \tag{4.10}$$

The Hamiltonian $\hat{H}_F[t_0]$ defines an effective time-independent model with the same stroboscopic time evolution as the original time-periodic system described by $\hat{H}(t)$. It follows that

$$\hat{H}_F[t_0]=\sum_n \varepsilon_n |u_n(t_0)\rangle\langle u_n(t_0)|, \tag{4.11}$$

with the sum over Floquet steady modes $|u_n(t_0)\rangle$ representing distinct physical states. Clearly, the Floquet Hamiltonian $\hat{H}_F[t_0]$ is not uniquely defined because of the ambiguity in the definition of the quasienergies. However, if we restrict ourselves to a certain Floquet quasienergy zone then this ambiguity in the definition disappears. There is another detail to notice: The stroboscopic Floquet Hamiltonian takes the starting time of the stroboscopic evolution t_0 as argument. The particular choice of t_0, defining the so-called Floquet gauge [78], must not have any effect on the system's dynamics. To see how this gauge actually works, let us consider the time-evolution operator $\hat{U}(t_2,t_1)$ for arbitrary times t_1 and t_2 and rewrite it in the following form:

$$\begin{aligned}\hat{U}(t_2,t_1)&=\hat{U}(t_2,t_0+nT)\exp\left[-i\hat{H}_F[t_0]nT/\hbar\right]\hat{U}(t_0,t_1)\\&=\hat{P}(t_2,t_0)\exp\left[-i\hat{H}_F[t_0](t_2-t_1)/\hbar\right]\hat{P}^\dagger(t_1,t_0)\\&=\exp\left[-i\hat{K}_F[t_0](t_2)/\hbar\right]\exp\left[-i\hat{H}_F[t_0](t_2-t_1)/\hbar\right]\exp\left[i\hat{K}_F[t_0](t_1)/\hbar\right],\end{aligned} \tag{4.12}$$

where we have introduced two new operators $\hat{P}(t,t_0)$ and $\hat{K}_F[t_0](t)$,

$$\hat{P}(t_2,t_0)\equiv\hat{U}(t_2,t_0)\exp\left[-i\hat{H}_F[t_0](t_2-t_0)/\hbar\right]\equiv\exp\left[-i\hat{K}_F[t_0](t_2)/\hbar\right]. \tag{4.13}$$

The family of operators $\hat{K}_F[t_0](t)$ are called stroboscopic kick operators. We also note that $\hat{P}(t,t_0)=\hat{P}(t+nT,t_0)=\hat{P}(t,t_0+nT)$ because the drive is time-periodic. Now, within this formulation the choice for t_0 is completely arbitrary and the evolution operator does not depend on it at all. It follows that one must not be able to see any difference between two different choices

for t_0; the systems must look identical. This means that all stroboscopic Floquet Hamiltonians $\hat{H}_F[t_0]$ are gauge equivalent, and all are equivalent to some (not unique) effective Hamiltonian $\hat{H}_{\text{eff}}$. The equivalence means that they should all be equal up to a time-independent unitary transformation. In other words, there has to exist a family of Hermitian operators $\hat{K}(t_0)$ such that

$$\hat{H}_{\text{eff}} = e^{-i\hat{K}(t_0)/\hbar}\hat{H}_F[t_0]e^{i\hat{K}(t_0)/\hbar}. \tag{4.14}$$

The operators $\hat{K}(t_0)$ are called kick operators, not to be confused with the stroboscopic kick operators $\hat{K}_F[t_0](t)$. Note that the gauge equivalence of all stroboscopic Floquet Hamiltonians $\hat{H}_F[t_0]$ can be explicitly demonstrated from Eq. (4.13), $\hat{H}_F[t_0+\delta t_0] = \hat{P}(t_0+\delta t_0, \delta t_0)\hat{H}_F[t_0]P^{\dagger}(t_0+\delta t_0, \delta t_0)$ for all t_0 and δt_0, and therefore they are indeed all equal up to a unitary transformation. However, one may use some other choice for defining the kick operators $\hat{K}(t_0)$, and usually one chooses them to simplify $\hat{H}_{\text{eff}}$ as much as possible [78, 79, 80]. We stress one more time that every stroboscopic Floquet Hamiltonian $\hat{H}_F[t_0]$ and effective Hamiltonian $\hat{H}_{\text{eff}}$ have the same eigenvalues and therefore any of them can be used for extracting the quasienergy spectrum.

4.2 Floquet - Magnus expansion

In general, time-evolution operators are very complex objects and can rarely be addressed analytically. One may therefore look for approximations to find the stroboscopic Floquet Hamiltonian $\hat{H}_F[t_0]$ and kick operators $\hat{K}[t_0](t)$. One such approximate technique is the so-called Floquet-Magnus expansion.

In its most general form, the Magnus expansion does not require any time periodicity and can be applied basically to any time-dependent system [81, 82]. The idea is to bring the time-dependent evolution operator $\hat{U}(t_0+t, t_0)$ into the exponential form $\exp(\hat{\Omega}[t_0](t))$. Here $\hat{\Omega}[t_0](t)$ is an effective time-dependent operator that also depends on the choice of the initial time t_0 and it should not be confused with the frequency $\Omega = 2\pi/T$. The operator $\hat{\Omega}[t_0](t)$ is then found perturbatively, i.e. one represents $\hat{\Omega}[t_0](t)$ by some series $\sum_{n=1}^{\infty}\hat{\Omega}_n[t_0](t)$ that hopefully converges in the case of interest. The

series is extracted from the Magnus theorem[81] (the complete proof can be found in Ref. [82]):

Theorem (Magnus). *The evolution operator* $\hat{U}(t_0+t,t_0)$ *can be represented by* $\exp(\hat{\Omega}[t_0](t))$*, where*

$$\frac{d\hat{\Omega}[t_0](t)}{dt}=\sum_{n=0}^{\infty}\frac{B_n}{n!}ad^n_{\hat{\Omega}[t_0](t)}\frac{\hat{H}(t_0+t)}{i\hbar}, \tag{4.15}$$

with $ad^j_A B=[A,ad^{j-1}_A B]$, $ad^0_A B=B$, $[A,B]=AB-BA$. *Here* $\hat{H}(t_0+t)$ *is the time-dependent Hamiltonian and* B_n *are the Bernoulli numbers.*

The infinite series $\sum_{k=1}^{\infty}\hat{\Omega}_k(t)$ is then obtained by iteratively integrating Eq. (4.15). The procedure is similar to the way how the time-evolution operators, cf. Eq. (4.8), are usually obtained from the time-dependent Schrödinger equation. Explicitly, the first few terms are given by [82]

$$\begin{aligned}
\hat{\Omega}_1[t_0](t)&=\frac{1}{i\hbar}\int_{t_0}^{t+t_0}dt_1\,\hat{H}(t_1),\\
\hat{\Omega}_2[t_0](t)&=\frac{1}{2!(i\hbar)^2}\int_{t_0}^{t+t_0}dt_1\int_{t_0}^{t_1}dt_2\,[\hat{H}(t_1),\hat{H}(t_2)],\\
\hat{\Omega}_3[t_0](t)&=\\
&=\frac{1}{3!(i\hbar)^3}\int_{t_0}^{t+t_0}dt_1\int_{t_0}^{t_1}dt_2\int_{t_0}^{t_2}dt_3\Big([\hat{H}(t_1),[\hat{H}(t_2),\hat{H}(t_3)]]+[\hat{H}(t_3),[\hat{H}(t_2),\hat{H}(t_1)]]\Big).
\end{aligned} \tag{4.16}$$

The time-periodic variety of the expansion, dubbed by Floquet-Magnus expansion [82, 83], may be derived in a similar fashion: We look for series representations of the stroboscopic Floquet Hamiltonian $\hat{H}_F[t_0]$ and stroboscopic kick operators $\hat{K}_F[t_0](t)$. The most direct way to get the desired series for $\hat{H}_F[t_0]$ is to decompose the time-periodic Hamiltonian $\hat{H}(t)$ into Fourier components and then perform the corresponding integrals in Eq. (4.16) [82]. The series for $\hat{K}_F[t_0](t)$ is then obtained iteratively from Eq. (4.13). The first

two components of each expansion are given by [82]

$$\begin{aligned}
\hat{H}_F^{(0)}[t_0] &= \frac{1}{T}\int_{t_0}^{T+t_0} d\,t\,\hat{H}(t) = \hat{H}_0, \\
\hat{H}_F^{(1)}[t_0] &= \frac{1}{2!T(i\hbar)}\int_{t_0}^{T+t_0} d\,t_1 \int_{t_0}^{t_1} d\,t_2\,[\hat{H}(t_1),\hat{H}(t_2)] \\
&= \frac{1}{\hbar\Omega}\sum_{l=1}^{\infty}\frac{1}{l_1}\left([H_l,H_{-l}] - e^{il\Omega t_0}[H_l,H_0] + e^{-il\Omega t_0}[H_{-l},H_0]\right)
\end{aligned} \tag{4.17}$$

and

$$\begin{aligned}
\hat{K}_F^{(0)}[t_0](t) &= 0, \\
\hat{K}_F^{(1)}[t_0](t) &= \frac{1}{\hbar}\int_{t_0}^{t} d\,t'\left(\hat{H}(t') - \hat{H}_F^{(1)}[t_0]\right) \\
&= \frac{1}{i\hbar\Omega}\sum_{l\neq 0} H_l\,\frac{e^{il\Omega t} - e^{il\Omega t_0}}{l},
\end{aligned} \tag{4.18}$$

where $\hat{H}_n = \int_0^T \exp(-in\Omega t)\hat{H}(t)$ are Fourier components of the time-periodic Hamiltonian $\hat{H}(t)$. The terms $\hat{H}_F^{(n)}[t_0]$ are of order Ω^{-n}, because the terms proportional to Ω^{-k} with $k < n$ vanish[3] for each $\hat{\Omega}_n[t_0](T)$ in Eq. (4.16). It follows that Floquet-Magnus expansion is very handy for approximating stroboscopic Floquet Hamiltonians $\hat{H}_F[t_0]$ in cases with large Ω. Recall that one may retrieve the full quasienergy spectrum from any of the stroboscopic Floquet Hamiltonians $\hat{H}_F[t_0]$, however, alone they are not enough to describe the system's time evolution.

4.3 Floquet - Sambe formalism

The time-evolution operators discussed so far provide a nice platform for performing calculations within the Floquet formalism, however, often they are very complicated and cannot be efficiently addressed. It is then useful to

[3] Here we shall assume that the parameters in the Hamiltonian do not scale with the driving frequency. The scenarios where this happens are certainly very interesting, however, beyond the scope of this thesis.

try out a completely different way for finding the quasienergy spectrum and Floquet states [73, 74]. Let us consider a linear space consisting of all states $|x(t)\rangle$ that at each point of time live in Hilbert space $\mathscr{H}$ and are periodic in time with period T. This vector space is commonly denoted by $\mathscr{H}\otimes\mathscr{T}$ and each time-periodic state $|x(t)\rangle$ is usually dubbed $|x\rangle\!\rangle$ when referring to this space. Note that one may straightforwardly find a basis spanning the entire $\mathscr{H}\otimes\mathscr{T}$. Usually the easiest choice is to use $|i,m\rangle\!\rangle = e^{im\Omega t}|i\rangle$, where $|i\rangle$ is any set spanning the Hilbert space $\mathscr{H}$.

Consider now the following relation defined for any pair of time-periodic states $|x\rangle\!\rangle$ and $|y\rangle\!\rangle$:

$$\langle\!\langle x(t)|y(t)\rangle\!\rangle = \frac{1}{T}\int_0^T dt\,\langle x(t)|y(t)\rangle. \tag{4.19}$$

The relation $\langle\!\langle x(t)|y(t)\rangle\!\rangle$ satisfies all rules of an inner product in the vector space $\mathscr{H}\otimes\mathscr{T}$ and we may promote $\mathscr{H}\otimes\mathscr{T}$ to become an extended Hilbert space associated with the inner product $\langle\!\langle x(t)|y(t)\rangle\!\rangle$, called Floquet-Sambe space or frequency domain. This implies that all the tools we are used to apply within the convectional time-independent formalism should be directly adaptable here. For example, we may use the resolution of the identity $\sum_{i,m}|i,m\rangle\!\rangle\langle\!\langle i,m| = I$, where I is the identity operator[4] within $\mathscr{H}\otimes\mathscr{T}$.

The idea now is to view the time-periodic Schrödinger equation as an equation given within $\mathscr{H}\otimes\mathscr{T}$ [73, 74]. To do so we first directly plug the Floquet solutions into the Schrödinger equation to obtain

$$\left[\hat{H}(t) - i\hbar\frac{\partial}{\partial t}\right]|u_n(t)\rangle = \epsilon_n|u_n(t)\rangle, \tag{4.20}$$

where $\hat{H}(t)=\hat{H}(t+T)$, $|u_n(t)\rangle = |u_n(t+T)\rangle$, and ϵ_n is real for all integer n and times t. We define the so-called Floquet-Sambe Hamiltonian $\hat{\mathscr{H}}$ to be the operator $(\hat{H}(t) - i\hbar I\frac{\partial}{\partial t})$ with identity operator I. Note that the Hilbert space $\mathscr{H}\otimes\mathscr{T}$ is closed under operation of the Floquet-Sambe Hamiltonian, i.e. for any time-periodic state $|x(t)\rangle$, the state $\hat{\mathscr{H}}|x(t)\rangle$ is also periodic in time with the same period T. Thus, using completeness of the Floquet-Sambe space

[4] In this thesis we will be denoting by I both conventional time-independent and Floquet-Sambe identity operators. One has to distinguish between them from the context.

and closeness of $\hat{\mathscr{H}}$ we transform Eq. (4.20) in the following way:

$$\sum_{i,m,j,m'} \langle\!\langle i,m|\left[\hat{H}(t)-i\hbar I\frac{\partial}{\partial t}\right]|j,m'\rangle\!\rangle\langle\!\langle j,m'|u_n\rangle\!\rangle = \epsilon_n\langle\!\langle i,m|u_n\rangle\!\rangle, \qquad (4.21)$$

where the states $|i,m\rangle\!\rangle = e^{im\Omega t}|i\rangle$ form a complete basis of $\mathscr{H}\otimes\mathscr{T}$. Eq. (4.21) is simply an eigenvalue problem for a matrix with entries $\langle\!\langle i,m|\hat{\mathscr{H}}|j,m'\rangle\!\rangle = \langle\!\langle i,m|(\hat{H}(t)-i\hbar\frac{\partial}{\partial t})|j,m'\rangle\!\rangle$. The transformation (4.21) can be viewed as a decomposition of the Floquet-Sambe Hamiltonian and corresponding eigenmodes in Fourier basis functions of $\mathscr{H}\otimes\mathscr{T}$ [73]. However, in general any basis of the Floquet-Sambe space could be used for representing $\hat{\mathscr{H}}$ as a matrix and in some cases it is more practical to use slightly different bases, see for example [75].

The Floquet-Sambe Hamiltonian $\hat{\mathscr{H}}$ represented using the basis $|i,m\rangle\!\rangle = e^{im\Omega t}|i\rangle$ is explicitly given by the following infinite-dimensional matrix:

$$\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \dots & H^{(0)}-\hbar\Omega\mathrm{I} & H^{(1)} & H^{(2)} & \dots \\ \dots & H^{(-1)} & H^{(0)} & H^{(1)} & \dots \\ \dots & H^{(-2)} & H^{(-1)} & H^{(0)}+\hbar\Omega\mathrm{I} & \dots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad (4.22)$$

with matrix elements

$$\langle\!\langle i,m|\mathscr{H}|j,m'\rangle\!\rangle = \langle i|H^{(m'-m)}+m\hbar\Omega I\delta_{m,m'}|j\rangle, \qquad (4.23)$$

where $\langle i|H^{(n)}|j\rangle = 1/T\int_0^T dt\, e^{in\Omega t}\langle i|H(t)|j\rangle$ and I is the identity. This directly follows from the definitions of the inner product $\langle\!\langle x(t)|y(t)\rangle\!\rangle$ and basis states $|i,m\rangle\!\rangle$. Note that the term $m\hbar\Omega I\delta_{m,m'}$ has emerged as a result of differentiating the basis states $|i,m\rangle\!\rangle$. For the sake of clearness in notations we now introduce a few more concepts: We will be referring to the indices m and m' of any matrix given within Floquet-Sambe space as Floquet-Sambe row and column indices respectively. These indices divide any Floquet-Sambe matrix into an infinite number of blocks of the same dimension as the conventional Hilbert space $\mathscr{H}$, we will be calling them Floquet-Sambe blocks. These blocks are not all independent: for any integer n the blocks $(m,\, m')$ and $(m+n,\, m'+n)$ are equal up to a diagonal matrix $n\hbar\Omega I$ in case $m = m'$.

Thus, Floquet-Sambe matrices possess very unique structures and are in some sense redundant in their representation, i.e. they contain a lot of repetitive information. This simply reflects the fact that the Floquet modes $|u_n(t)\rangle$ and quasienergies ε_n are not uniquely defined.

It should be realized that the diagonalization of the Floquet-Sambe Hamiltonian $\hat{\mathscr{H}}$ implies solving an infinite-dimensional eigenvalue problem and therefore it can rarely be done analytically. However, the eigenstate solutions $|u_n\rangle\!\rangle$ are typically narrow[5] in the Floquet index [48]. This allows for very efficient numerical calculations by simply truncating the Floquet-Sambe matrices at sufficiently large Floquet-Sambe row and column indices. In Papers A and B it was sufficient to consider up to $(m,m') = (-3,...,3)$ for reaching good convergence of the results.

To sum up, the vector space $\mathscr{H}\otimes\mathscr{T}$ along with the inner product $\langle\!\langle x|y\rangle\!\rangle$ form a well defined Hilbert space and the Floquet modes $|u_n(t)\rangle = |u_n\rangle\!\rangle$ are elements of this space. Moreover, they are eigenstates of the Floquet-Sambe Hamiltonian $\hat{\mathscr{H}}$ defined in Eq. (4.22). It is important to note that eigenstates of the Floquet-Sambe Hamiltonian contain complete information about the system, i.e. we can reconstruct the entire time evolution of the modes, and we can also access the quesienergies. Thus, by switching to the Floquet-Sambe formalism we do not lose any information about the time evolution of the system, in contrast to the evolution operator approach discussed earlier.

[5] By this we mean that a typical Floquet-Sambe eigenmode $|u_n\rangle\!\rangle$ has substantial nonzero elements only close to some Floquet index m_0 and all other elements decay rapidly as we go away from it.

5 Time-periodic (Floquet) topological band insulators

Recall that periodically driven systems can be treated on a very similar footing as time-independent ones. Any state of a time-periodic system can be represented in terms of so-called Floquet states, cf. Chapter 4. These states are very similar to the stationary ones: they also have time-dependent phase factors. Analogously to the concept of energy, within the Floquet formalism the proportionality constant in the time-dependent phase is called quasienergy. In contrast to energies though, the quasienergies are uniquely defined only modulo $\hbar\Omega$ where the frequency of the drive is $\Omega = 2\pi/T$. This difference has a large impact on the structure of a generic time-periodic state and in what follows it will be shown how it affects the (out-of-equilibrium) quantum phases in a nontrivial way [47].

5.1 Floquet topological band insulators: A first look

In Chapter 4 our discussion was quite general and no limitations on the Hamiltonian $\hat{\mathscr{H}}$ were assumed except that it had to be periodic in time. From now on we will focus on noninteracting systems only. In analogy to time-independent theory, in this case a general many-body solution to the time-periodic Schrödinger equation can be represented by populating single-particle Floquet steady states. These states are solutions to the single-particle Schrödinger equation, with the corresponding first-quantized Hamiltonian $H(t)$. All techniques discussed in Chapter 4 can be straightforwardly adapted here by replacing many-body operators with analogous single-body ones and then follow exactly the same analysis and derivations.

Noninteracting fermions in time- and space-periodic potentials may be described by a Floquet extension of the Bloch theorem. The Floquet-Bloch quasienergy bands may then be topologically classified in a conceptually similar fashion as the energy bands in the time-independent formalism. The machinery for doing so, however, is very different. The crucial difference between those two topological classifications is the intrinsic periodicity in quasienergy that results in additional topological features of the corresponding spaces. In short, the driving may open a gap at the edge of the

Floquet zone (the region where quasienergies start to repeat), by this potentially changing the topological properties of the bands [45, 48, 47]. It has also recently been shown that certain periodic drives may transform the otherwise topologically trivial systems into topologically nontrivial ones: they may renormalize parameters in the models and facilitate topological transitions [46, 75]. Moreover, external periodic fields are usually relatively easy to handle in practice and therefore the corresponding topological phase transitions can be controlled in a rather straightforward way. This opens up new possibilities for engineering topologically non-trivial systems, dubbed Floquet band structure engineering [47].

Let us first think about a possibility of having inverted bands in periodically driven systems, important for spotting topological phase transitions, cf. Chapter 1. The easiest way to address the notion of band inversion here is to employ the Floquet-Sambe construction and look at generic cases in which the bands may invert under continuous transformations of the Hamiltonian. In analogy to the time-independent theory, the Floquet-Sambe quasienergy bands are continuous in momentum, however, due to the redundancy in the Floquet formalism they are also repeated with periodicity $\hbar\Omega$. In the case of no driving, only the zeroth Fourier component of the Hamiltonian $H(t)$ is non-zero and therefore the Floquet-Sambe Hamiltonian $\mathscr{H}$ has a block-diagonal form. Each block has the same eigenvalues but shifted by $m\hbar\Omega$ where $m \in \mathbb{Z}$ denotes the block's Floquet-Sambe row and column indices. Thus, the Floquet-Sambe Hamiltonian $\mathscr{H}$ has the same eigenspectrum as $H(t) = H$ but artificially duplicated with the shift of an integer multiple of $\hbar\Omega$. In the limit $\Omega \to \infty$ these blocks are far from each other in eigenspectra and cannot be affected by each other even if we couple them through some off-diagonal elements, i.e. time-periodic driving. In other words, this scenario does not bring any new phenomena: all features are contained in the time-independent part of the Hamiltonian. In this way we establish the 'natural' order for the quasienergy bands: any two bands associated with different Floquet-Sambe row and column indices are ordered with respect to their indices, i.e. the larger the Floquet-Sambe index is, the higher is the band. Let us now come back to the system with arbitrary Ω and continuously add a time-periodic drive to it. The driving introduces some non-zero Fourier components in the expansion of $H(t)$ and therefore results in non-zero off-diagonal blocks in $\mathscr{H}$. Such off-diagonal elements may hybridize the bands belonging to different Floquet-Sambe diagonal blocks and open a gap near

the points where the bands intersect. To put it differently, the hybridization may invert the quasienergy bands from different Floquet-Sambe blocks. The necessary condition for this to happen is to have bands intersections and therefore the driving frequency has to be sufficiently small. Also, the driving has to open a gap and therefore the time-periodic part of the Hamiltonian must not commute with the time-independent part, otherwise, the gap will not open. To sum up, in addition to a conventional band inversion that can happen between conventional energy bands, in time-periodic band insulators we can have inversions between bands from different blocks of the Floquet-Sambe Hamiltonian. Such an extended scenario for having inverted bands suggests that the time-periodic band theory may contain (and most likely does contain) some additional topological features that have no analogues in the time-independent formalism. As we have repeatedly stressed, band inversions are very useful for catching topological phase transitions and this is explicitly demonstrated on the example of the periodically driven SSH model.

5.2 The periodically driven Su-Schrieffer-Heeger model

Here we illustrate the notion of a topological phase transition within the Floquet formalism on a simple case study: the periodically driven SSH model. The ideas developed below are then generalized in the succeeding section. Consider an SSH chain with the following time-periodic modulation of the hopping amplitudes:

$$H(t) = -\sum_{j=1}^{N} \Big((\gamma_{1,j} - v_j(t)) |A, j\rangle\langle B, j| + (\gamma_{2,j} + v_j(t)) |B, j\rangle\langle A, j+1| + \text{h.c.} \Big), \tag{5.1}$$

where as before A, B denote the sublattices, $\gamma_{1,j}$ and $\gamma_{2,j}$ are intercell and intracell static hopping amplitudes respectively, and $v_j(t) = 2V_j \cos(\Omega t)$ is the harmonically modulated component of the hopping. As before, we assume that the system is homogeneous in its interior, cf. Fig. 5.1. This model has been extensively studied in Ref. [84] and Papers A and B.

To begin with, let us take a look at homogeneous time-independent SSH chains under periodic boundary conditions and continuously add the time-

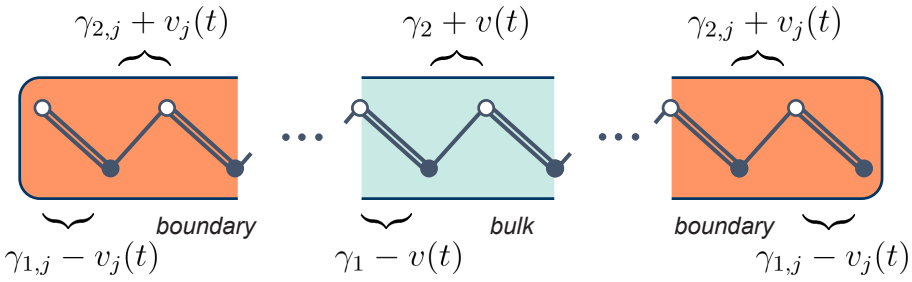

Figure 5.1: The periodically driven SSH model: An illustration of a driven SSH chain with time-periodic intercell $\gamma_{1,j} - v_j(t)$ and intracell $\gamma_{2,j} + v_j(t)$ hopping amplitudes. The chain is divided into interior region (bulk) and boundary regions (boundaries).

periodic modulation of the hopping amplitudes. Without any periodic drive[1], the Hamiltonian $H(t) = H$ defines two Bloch bands with the dispersion given by Eq. (2.3). Within the Floquet-Sambe formalism, however, these bands get duplicated with periodicity of $\hbar\Omega$. If the frequency is large enough, $\Omega > 2(|\gamma_1| + |\gamma_2|)$, there are no crossings between the bands from different Floquet-Sambe blocks. This regime is not interesting because the continuous addition of $v(t)$ will not invert any bands and therefore one does not expect any new phenomena coming from the time-periodic drive. However, for $\Omega < 2(|\gamma_1| + |\gamma_2|)$ there is at least one band crossing in the eigenspectrum of the Floquet-Sambe Hamiltonian and $v(t)$ induces band inversion(s). As a result one may expect, topological phase transitions in this case. Depending on Ω there can be multiple band crossings[2] at $v(t) = 0$ and therefore band inversions at $v(t) \neq 0$. Accordingly, one expects a very rich topological phase diagram for this model.

It is interesting and important to note the following: Before performing a continuous transformation that inverts bands, in a two-band time-independent Hamiltonian one may formally define valence (bottom) and conduction (top) bands. The band inversion then always happens with the conduction band on top of the valence band. This inversion signals a topological phase transition, with the topological index counting the number of

[1] It is probably better to say: with the infinitesimal homogeneous driving amplitude $V_j = V$.

[2] The band folding due to periodicity of quasienergies become more and more dense as Ω decreases.

symmetry-protected boundary states inside the gap. Analogously, in a Floquet two-band insulator one may also define top and bottom bands, for clearness we give them names Floquet-conduction and Floquet-valence bands respectively, by fixing the Floquet quasienergy zone to $-\hbar\Omega/2 \leq \varepsilon < \hbar\Omega/2$. It follows that under a continuous transformation the bands may invert in two different ways: with the Floquet-valence band or with the Floquet-conduction band on top of the other one. The former scenario may happen just because of the quasienergy band-structure periodicity and it is a pure artifact of the time-periodic formalism. It has no analogue in time-independent systems. The existence of such two flavors of band inversions suggest that there should be two kinds of topological phase transitions, described by different topological indices. It follows that the band inversion(s) with the Floquet-conduction band on top of the Floquet-valence band corresponds to a topological phase transition with the topological index counting the number of symmetry-protected boundary states in the conventional band gap. On the other hand, if the inversion is between the bands with the Floquet-valence band on top then the corresponding topological phase transition is anomalous, with the topological invariant counting the number of symmetry-protected boundary states in the anomalous band gap. The anomalous (sometimes also called induced) gap here corresponds to a gap in quasienergy that connects two bands through the Floquet zone boundary, i.e. at quasienergies $\pm\hbar\Omega/2$.

Analogously to the time-independent case, one should look for ways to establish robustness of the midgap edge states to local perturbations. In order to do so we first notice that the periodically driven SSH model, Eq. (5.1), is invariant under the sublattice symmetry adapted to the Floquet formalism. The sublattice symmetry in this context means that the relation $\Gamma H(t)\Gamma = -H(-t)$ is fulfilled, where $\Gamma = \sum_j \left(|A,j\rangle\langle A,j| - |B,j\rangle\langle B,j|\right)$ is the conventional time-independent sublattice operator. This relation simply manifests the fact that the hopping amplitudes are both even in time. The sublattice symmetry immediately leads to $\Gamma^\dagger U(T,0)\Gamma = U^\dagger(T,0)$ that imposes some restrictions on the quasienergy spectrum: the Floquet mode $|u(t=0)\rangle$ is an eigenstate of $U(T,0)$ with eigenvalue $\exp(-i\varepsilon T/\hbar)$ if and only if the Floquet mode $\Gamma|u(t=0)\rangle$ is also an eigenstate but with eigenvalue $\exp(i\varepsilon T/\hbar)$. It follows that under any continuous symmetry-preserving transformation the modes at quasienergies 0 and $\hbar\Omega/2$ can leave these levels only in pairs. Thus, as was anticipated, there are two flavors of protected edge states pos-

sible in this scenario, with quasienergy 0 or $\hbar\Omega/2$. Following exactly the same line of analysis as in the time-independent case, if these states exist in some large system with gapped interior[3], then they cannot disappear under any gap-preserving transformations of the bulk or any local symmetry-preserving deformations of the boundaries. Thus, both flavors of edge states are symmetry-protected in this sense. Note that this symmetry protection relies just on the fact that the (Floquet) sublattice symmetry is preserved, i.e. one may also consider perturbations that are time-periodic. To investigate such robustness is the main objective in Paper A.

5.3 Topological classification of Floquet insulators

In the past several years there were conducted quite a few works classifying periodically driven band insulators and superconductors, to name a few [48, 85, 86, 87, 88, 89], with comprehensive generalization to all ten Altland-Zirnbauer symmetry classes in Ref. [76]. In the previous section, 5.2, it was shown on a concrete case study that there may be two flavors of robust boundary states present in Floquet topological insulators. This suggests that the Floquet (non-equilibrium) phases do not follow the standard classification developed in Chapter 3. The Floquet classification is expected to be much richer and below we briefly sketch how it is developed. Here we focus on providing the overall picture and skip all unnecessary technicalities; see Ref. [76] for more details.

Without loss of generality, from now on we will always consider the quasienergies to fall into the Floquet zone $-\hbar\Omega/2 \leq \varepsilon < \hbar\Omega/2$ and fix the Floquet gauge to $t_0 = 0$. The idea behind the classification of periodically driven band insulators is to construct homotopy equivalence classes for Bloch time-evolution operators $U(\boldsymbol{k}, t)$ with $t \in [0, T)$, rather than for Floquet-Bloch Hamiltonians $H(\boldsymbol{k}, t)$. This is motivated by the fact that Floquet steady states are eigenstates of time-evolution operators, not of time-periodic Hamiltonians. Obviously, if we would not restrict the time-evolution operators in any way, then all time-periodic systems would be equivalent because all of them can be continuously transformed to the identity. This is similar to the time-independent case where the space of all Hamiltonians is trivial, but the

[3]The presence of quasienergy gaps anticipates that the states at quasienergies 0 or $\hbar\Omega/2$ are localized at the boundaries of the system.

space of Q-matrices (gapped Hamiltonians) is in general not. Following this analogy, we are then interested in the evolutions $U(\boldsymbol{k}, t)$ of time-periodic systems that are gapped around at least one value of quasienergy, i.e. the eigenspectrum of $U(\boldsymbol{k}, T)$ has at least one gap. In the time-independent classification, the gap was continuously shifted to zero energy, cf. Chapter 3, and then the corresponding Q-matrices were classified. In the Floquet case we do something similar: we continuously shift the gap of interest to zero but at the same time also leave the gap at $\hbar\Omega/2$ open[4]. As we shall discuss, both gaps can be described by separate topological invariants and may host topologically protected boundary states. We now say that two evolution operators are (homotopically) equivalent if they can be transformed into each other by local continuous deformations without closing the gaps. Note that this definition is very similar to the one from the time-independent theory where we define two time-independent Hamiltonians to be equivalent if they can be deformed into each other without gap closing. In analogy to the time-independent case we can further restrict the space of considered evolution operators by looking at different symmetry classes, to be briefly described below.

5.3.1 Symmetries in periodically driven systems

Recall that symmetries in time-independent systems mainly serve as certain restrictions on Hamiltonians and therefore on their eigenvalue spectra. There are in total ten fundamental (Altland-Zirnbauer) symmetry classes defined by the presence of TRS±1, PHS±1 and CS, cf. Chapter 3. The symmetry classes within the Floquet theory are directly inherited from the time-independent formalism: They are defined by fundamental symmetries of the stroboscopic evolution operator $U(T)$ or, equivalently, of the stroboscopic Floquet Hamiltonian H_F. As one may have noticed, here we dropped the momentum dependence $\boldsymbol{k}$ from the evolution operator $U(\boldsymbol{k}, T)$. This was done to highlight the fact that in general the translational invariance is not needed for defining symmetry invariance under TRS, PHS or CS. In the following we briefly look at each basic symmetry separately.

[4]We assume that the gap at $\hbar\Omega/2$ is also present. Otherwise, one of the topological invariants, to be presented, will lose its meaning, but the overall classification will still be valid.

Time - reversal symmetry. — Time-reversal symmetry (TRS) is an antiunitary symmetry that transforms the time-periodic Hamiltonian $H(t)$ as follows:

$$H(t) = U_T^\dagger H^*(T-t)U_T \tag{5.2}$$

for some unitary U_T, not to be confused with the time-evolution operator over one period $U(T)$. This unitary operator is required to fulfill $U_T^* U_T = \pm I$ that defines two fundamentally distinct symmetries TRS+1 and TRS−1. Clearly, one may immediately establish that $U(T) = U_T^\dagger [U(T)]^{\dagger *} U_T$ and therefore the stroboscopic Floquet Hamiltonian is also invariant under TRS, $H_F = U_T^\dagger H_F^* U_T$.

Particle - hole symmetry. — Particle - hole symmetry (PHS) is also an antiunitary symmetry but it acts on the time-periodic Hamiltonian $H(t)$ as

$$H(t) = -U_C^\dagger H^*(t)U_C, \tag{5.3}$$

where U_C is a unitary operator with $U_C^* U_C = \pm I$ representing two different symmetries PHS+1 and PHS−1. This symmetry implies that the evolution operator operator transforms as $U(T) = U_C^\dagger [U(T)]^* U_C$ and the Floquet Hamiltonian as $H_F = -U_C^\dagger H_F^* U_C$.

Chiral symmetry. — In analogy to the time-independent formalism, chiral symmetry (CS) is defined as the product of TRS and PHS:

$$H(t) = -U_S^\dagger H(T-t)U_S, \tag{5.4}$$

where U_S is a unitary operator. Note that the following relations hold: $U(T) = U_S^\dagger [U(T)]^\dagger U_S$ and $H_F = -(U_S)^\dagger H_F U_S$, meaning that the (time-independent) stroboscopic Floquet Hamiltonian H_F is chiral-symmetric.

As in the time-independent case, we then distinguish ten different symmetry classes by the presence (absence) of TRS±1, PHS±1 and CS, cf. Table 3.1. There is one important detail to notice here: Within the time-periodic formalism TRS and CH are intrinsically nonlocal in time, and in this respect they are fundamentally different from time-local PHS. One may play with the time-dependence of the drive to either maintain or break the corresponding symmetry. This creates a playground for very rich phenomena in time-periodic systems topologically protected by TRS or CS.

5.3.2 Classification of time-evolution operators

Let us now come back to the equivalence classes of gapped evolution operators discussed earlier. Recall that we define two evolutions to be equivalent if they can be transformed into each other by a continuous deformation keeping the gaps at 0 and $\hbar\Omega/2$ open. We now modify the equivalence relation by also requiring them to be within the same symmetry class, and the deformation to not break any of the underlying symmetries. We then look for topological invariant(s) corresponding to each of the ten symmetry classes.

The gap around $\hbar\Omega/2$ allows one to uniquely define the Floquet Hamiltonian $H_F(\boldsymbol{k}) = i\hbar \ln[U(\boldsymbol{k}, T)]/T$ with the branch cut of the logarithm placed at $\varepsilon = \hbar\Omega/2$. It is clear that if two Floquet Hamiltonians $H_F^{(1)}(\boldsymbol{k})$ and $H_F^{(2)}(\boldsymbol{k})$ are topologically inequivalent then so the corresponding time-evolution

Name	Cartan	d=1	d=2	d=3
Standard	A (unitary)	-	$\mathbb{Z}\times\mathbb{Z}$	-
(Wigner-Dyson)	AI (orthogonal)	-	-	-
	AII (symplectic)	-	$\mathbb{Z}_2\times\mathbb{Z}_2$	$\mathbb{Z}_2\times\mathbb{Z}_2$
Chiral	AIII (chiral unitary)	$\mathbb{Z}\times\mathbb{Z}$	-	$\mathbb{Z}\times\mathbb{Z}$
(sublattice)	BDI (chiral orthogonal)	$\mathbb{Z}\times\mathbb{Z}$	-	-
	CII (chiral symplectic)	$\mathbb{Z}\times\mathbb{Z}$	-	$\mathbb{Z}_2\times\mathbb{Z}_2$
BdG	D	$\mathbb{Z}_2\times\mathbb{Z}_2$	$\mathbb{Z}\times\mathbb{Z}$	-
	C	-	$\mathbb{Z}\times\mathbb{Z}$	-
	DIII	$\mathbb{Z}_2\times\mathbb{Z}_2$	$\mathbb{Z}_2\times\mathbb{Z}_2$	$\mathbb{Z}\times\mathbb{Z}$
	CI	-	-	$\mathbb{Z}\times\mathbb{Z}$

Table 5.1: The table of homotopy equivalence classes for unitaries $U(\boldsymbol{k}, t)$ for different Altland-Zirnbauer classes and dimensions of the BZ. The sign '-' marks the trivial group. The group $\mathbb{Z}\times\mathbb{Z}$ is isomorphic to the group of pairs of integers, $\mathbb{Z}_2\times\mathbb{Z}_2$ represents the group isomorphic to the pair of cyclic groups of two binary numbers.

operators $U^{(1)}(\boldsymbol{k},T)$ and $U^{(2)}(\boldsymbol{k},T)$. The topological nontriviality of $H_F(\boldsymbol{k})$ results in robust boundary states at zero quasienergy, however, completely ignores the properties at the anomalous gap, the gap at $\hbar\Omega/2$.

To capture the topological properties at the $\hbar\Omega/2$ gap we define the notion of a unitary loop, given by a time-evolution operator $U(\boldsymbol{k},t)$ with $U(\boldsymbol{k},0) = U(\boldsymbol{k},T) = I$. It can be shown that any time-evolution operator can be represented by a composition of a unitary loop and a constant (time-independent) Hamiltonian evolution [76]. Two time-evolution operators are then not equivalent if and only if either the corresponding loops or the time-independent Hamiltonians are not equivalent. One then distinguishes two topological invariants separately counting the equivalence classes of unitary loops and constant Hamiltonians. Clearly, the latter invariant may be described by the conventional topological classification, Table 3.2, and corresponds to the number of robust states at zero quasienergy. The K-theoretic classification of the unitary loops is extensively studied in Ref. [76]. Here we do not dig into the technical details and present only the result. Note that because the equivalence classes for unitary loops and time-independent Hamiltonians are constructed from the base space isomorphic to a d-dimensional torus, the notion of a weak and a strong topological invariant can be straightforwardly inherited from the time-independent classification, cf. Chapter 3.

In Table 5.1 we present (strong) topological invariants for all ten Altland-Zirnbauer symmetry classes. As expected, each nontrivial class contains a pair of topological indices corresponding to the robust boundary modes in each of the gaps, at zero and $\hbar\Omega/2$. In analogy to the time-independent classification, the weak topological invariants can be obtain by looking at strong topological indices in lower dimensions.

6 Protection of the boundary states against time-periodic perturbations

The classification schemes of SPT phases provide a platform for classifying fundamentally different classes of quantum matter. The topologically nontrivial phases within non-interacting theories are by now well understood and the formalism is well established. But what is their relevance to experiments? What concrete properties distinguish them from trivial insulators? In this chapter we describe one particular property of the noninteracting SPT phases, namely robustness of the symmetry-protected boundary states against time-periodic perturbations, and discuss its experimental relevance. All the results are described in details in Papers A, B, C in Part II, and in Secs. 6.1 and 6.2 we only briefly summarize the main findings of these papers. In Sec. 6.3 we propose another approach for studying the protection that might be within experimental reach.

6.1 The symmetry-protected states under time-periodic perturbations

Recall that any nontrivial SPT phase is often identified by the presence of boundary states, commonly known as the bulk-boundary correspondence. The correspondence between bulk and boundary implies that these states are robust to any gap-maintaining perturbation respecting the underlying symmetries. The protection of the boundary states in Floquet TIs can be established on a very similar footing as in conventional time-independent TIs, but now the Floquet formalism allows one to also include time-periodic perturbations into the picture. Besides providing wider stability of the boundary states, the time-dependence here brings an extra degree of control and many phenomena can be probed by simply tuning the time-dependent part of the perturbation, in this way opening up new opportunities for the experiments.

In Paper A we conduct a comprehensive study on the time-periodic robustness of 1d Floquet TIs protected by chiral symmetry and numerically verify all the results on the harmonically-driven SSH model, cf. Sec. 5.2. Note that the chiral symmetry within the Floquet formalism, $\Gamma H(t)\Gamma = -H(-t)$ for some unitary operator Γ, is nonlocal in time and therefore the time-

dependence here puts a non-trivial constraint on the Hamiltonian. This is in contrast to, for example, the particle-hole symmetry, $PH(t)P = -H^*(t)$ for some unitary P, that simply reduces to conventional time-independent particle-hole symmetry at each time frame. As expected, the edge states in 1d chiral Floquet TIs are found to exhibit protection against gap-preserving perturbations respecting the Floquet chiral relation: The edge states are required to stay put exactly at the middle of the band gap as long as the chiral symmetry is preserved. The formal proof of this fact is spelled out in Paper A and its simplified version has already been sketched in Sec. 5.2 considering the time-periodic SSH model: The protection here simply reflects the fact that the chiral relation $\Gamma H(t)\Gamma = -H(-t)$ obligates the steady modes to come in symmetry-bonded pairs, namely $|u(t)\rangle$ with quasienergy ε and $\Gamma|u(-t)\rangle$ with quasienergy $-\varepsilon$. Note that the symmetry preservation and therefore the robustness of the edge states in Floquet TIs strongly depends on the relative phase between the time-periodic drivings of the bulk and the disordering perturbation: The underlying protecting chiral symmetry is non-local in time and to maintain it the driving phases of the perturbation and bulk have to be fine-tuned.

Any time-independent system is trivially periodic in time, and therefore can be considered within Floquet theory as a special case, associated with an arbitrary period of the driving. This simple change of perspective allows one to painlessly embed the conventional TIs into the Floquet formalism and apply the same argument for establishing robustness of the boundary states against time-periodic perturbations. It follows that the boundary states in static chiral TIs in 1d are protected against time-periodic chiral-preserving perturbations on the same grounds as in Floquet TIs. Note, however, that time-independent chiral TIs fulfill the Floquet chiral relation $\Gamma H(t_0 + t)\Gamma = -H(t_0 - t)$ for any reference time t_0 and preservation of this relation for at least one reference time guarantees that the edge states stay protected under the perturbations. As we discuss in Paper A, it follows that the boundary states in time-independent TIs generically display robustness to a much broader class of time-periodic perturbations than in Floquet TIs.

The symmetry-protected boundary states are in general not expected to stay protected once the underlying symmetries are broken. In Paper A we reveal an unexpected resilience of the boundary states in time-independent TIs against a generic class of time-periodic perturbations that break the chiral symmetry. We analytically explain it within Floquet perturbation theory by

showing that the two leading-order quasienergy contributions vanish identically due to the chiral structure of the unperturbed spectrum. This is a very interesting feature because it suggests that the underlying symmetries can still have a residual effect on the TI's behavior even when they are explicitly broken.

To experimentally reveal the robustness of the edge states against time-periodic perturbations one must first identify viable and sensitive probes that might be feasible to design in a lab.

6.2 Transport signatures of the time-periodic protection

In Paper B we propose a mesoscopic transport setup consisting of a 1d finite-sized Floquet TI attached to two external leads, Fig. 1 in Paper B. We predict that this kind of setup will reveal unique signatures of the Floquet symmetry protection in linear conductance measurements. The basic idea here is simple, however, the actual calculation becomes rather technical and is carried out numerically employing a nonequilibrium Green's function formalism within Floquet-Landauer–Büttiker theory. The symmetry-protected midgap states in a large 1d Floquet TI are localized at the edges and therefore do not contribute to transport. However, they hybridize in a finite-sized Floquet TI and by this open-up a conduction channel. This happens for the symmetry-protected edge states at both band gaps, the conventional gap at zero quasienergy and the anomalous gap (induced by the driving) at the boundary of the Floquet zone. Thus, the presence of edge states shows up as midgap peaks in the transmission spectra, Fig. 4 in Paper B, with the peaks gradually disappearing when increasing the size of the TI. The transport set-up can then be perturbed by adding a time-periodic gate potential which effectively disorders the scattering region. The symmetry-breaking perturbations misalign the quasienergy levels of the edge states, by this destroying the channel for transport. On the other hand, the symmetry-preserving perturbations only affect the transport due to small finite-size effects. Interestingly, by adjusting the phase of a time-periodic perturbation one may switch between symmetry-preserving and symmetry-breaking cases and by this change the set-up from conducting to nearly insulating. The qualitative and numeric analysis is then supported by analytic predictions that are

developed by projecting the transport problem onto the space spanned by pairs of symmetry-protected states. Within this analytic description we could also explain the symmetry protection of the transmission midgap peaks, cf. Appendices C and D in Paper B.

All calculations are performed within a Floquet-Landauer-Büttiker formalism, employing a convenient representation of the nonequilibrium Green's functions presented in Paper C. The idea is to construct nonequilibrium Green's functions directly within Floquet-Sambe space. It is then shown that the transport expressions for currents and densities essentially replicate the well-known formulas from the time-independent formalism. We found this formalism particularly useful for doing analytic transmission predictions, detailed in Appendix D in Paper B. Note that the methodology in Paper C is developed in close analogy to the time-independent formalism and all the derived expressions are given in a form well-suited for performing self-consistent transport calculations.

6.3 Boundary density probes of the symmetry-protected edge states

In this section we present another blueprint for a study of the symmetry-protected edge states and their robustness against time-periodic perturbations. The idea is to track the time evolution of a state initialized at one edge of a 1d finite-sized TI and uncover oscillations of its density between the edges. We predict that these oscillations will be highly suppressed in presence of a local symmetry-breaking perturbation but remain if the symmetry is maintained.

Recall that according to the Floquet theorem there is a complete set of solutions to the time-periodic Schrödinger equation given in terms of steady states $|\psi_n(t)\rangle = e^{-i\varepsilon_n t/\hbar}|u_n(t)\rangle$. The time-periodic Floquet modes $|u_n(t)\rangle$ can explicitly be found by diagonalizing the evolution operator $U(t, t+T)$ or equivalently solving the Floquet-Sambe eigenvalue equation. Take now any time-periodic system and initialize it at some state $|\Psi(t=0)\rangle$. The evolution of $|\Psi(t)\rangle$ can then be rewritten in terms of the steady modes $|u_n(t)\rangle$. Explicitly,

$$|\Psi(t)\rangle = \sum_n a_n e^{-i\varepsilon_n t/\hbar}|u_n(t)\rangle, \tag{6.1}$$

with $a_n = \langle u_n(t=0)|\Psi(t=0)\rangle$ where we simply use the fact that $|u_n(t=0)\rangle$ constitute a complete basis because these states are eigenstates of $U(0,T)$ and therefore they all are linearly independent. Note that in this form the state $|\Psi(t)\rangle$ automatically satisfies the time-periodic Schrödinger equation because each steady state $|\psi_n(t)\rangle = e^{-i\varepsilon_n t/\hbar}|u_n(t)\rangle$ is by definition a solution.

6.3.1 The density evolution in an unperturbed Floquet TI

Now consider a 1d chiral Floquet TI in the thermodynamic limit and initialize a state $|\Psi(t)\rangle$ at, say, its left edge. This state will generically have a considerable overlap with the symmetry-protected steady modes at the left edge, and hence a large portion of the state will be trapped in the left boundary region. Differently, in finite-sized TIs the symmetry-protected edge states hybridize and the state $|\Psi(t)\rangle$ will no longer stay at just one of the boundary regions but instead it will oscillate between the edges with some characteristic frequency depending on the hybridization strength. Note that the bulk and other accidentally localized states can also have non-zero overlap with the initial state and they will in general contribute to the state's evolution as well. However, their behaviour is not expected to convey any generic character, generally leaving the oscillations of the symmetry-protected states practically visible.

In Fig. 6.1 we present numerical results obtained by considering the harmonically driven SSH model as a particular realization of a Floquet TI in 1d. Here we show the density evolution of $|\Psi(t)\rangle$ initialized at first site $j=1$ of a finite-sized SSH array. The numerics is implemented by evolving the state in small time intervals and the Hamiltonian is taken to be constant at each of the intervals. In different panels we consider different sets of hopping amplitudes γ_1 and γ_2 representing different parts of the topological phase diagram, characterized by the two topological numbers ν_0 and ν_π. These indices count the numbers of symmetry-protected states localized at each of the edges and associated with quasienergies 0 and $\hbar\Omega/2$. The topological phase diagram corresponding to this model is formally discussed in Paper A. As expected, in the topologically nontrivial cases we observe oscillations of the state's density between the two edges: We attribute these oscillations to the presence of symmetry-protected edge states. Note that in the case of $\nu_0 \neq 0$ and $\nu_\pi \neq 0$ the symmetry-protected edge states at 0 and $\hbar\Omega/2$ interfere and create a nontrivial pattern in the density.

Having neglected effects coming from the bulk and accidental edge states

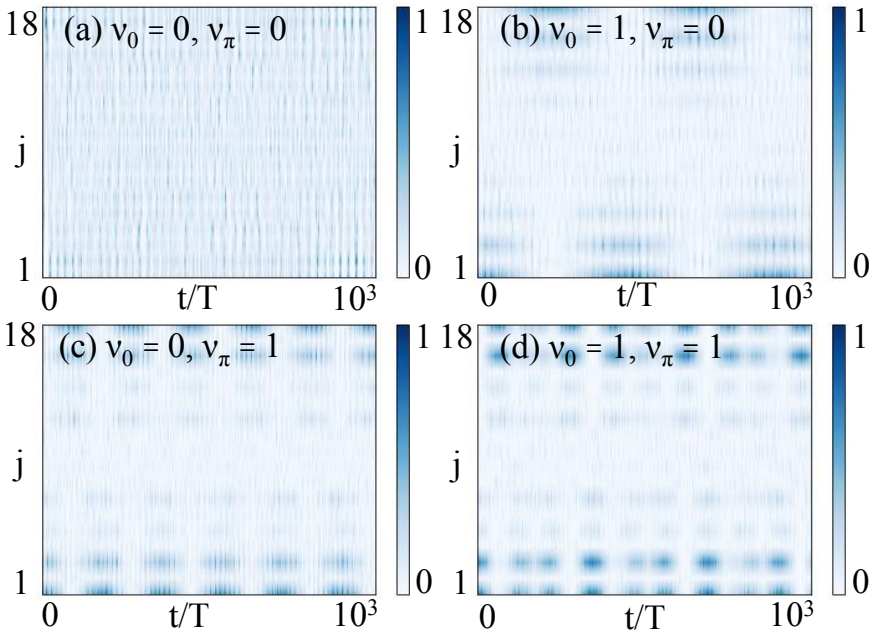

Figure 6.1: The density $|\Psi_j(t)|^2$ vs. time t and site number j: Here we consider the harmonically driven SSH chains of 18 sites and hopping constants (a) $\gamma_1 = 0.75$ and $\gamma_2 = 0.45$, (b) $\gamma_1 = 0.45$ and $\gamma_2 = 0.75$, (c) $\gamma_1 = 1.2$ and $\gamma_2 = 0.8$, and (d) $\gamma_1 = 0.8$ and $\gamma_2 = 1.2$. The driving was set to $v(t) = 0.4\cos(\Omega t)$ with $\Omega = 3$ and $T = 2\pi/\Omega$. The corresponding bulk topological indices ν_0 and ν_π are also shown. For doing numerics the time was discretized with interval $\delta t = 0.01T$.

we can analytically address the time evolution by projecting the Schrödinger equation onto the symmetry-protected states. For concreteness we here assume TIs with only one pair of the symmetry-protected edge states – one at the left edge, the other at the right edge – but the analysis to follow can be straightforwardly generalized to the case with multiple pairs. Recall that in time-periodic chiral systems the driving may open an additional gap, a so-called dynamical (or 'anomalous') gap, centered at the Floquet zone with quasienergy $\epsilon = \hbar\Omega/2$, which may also host symmetry-protected edge states. This is in addition to the 'normal' gap at zero quasienergy. Here we begin by considering this latter case, i.e. a Floquet TI with a pair of symmetry-protected states near zero quasienergy.

The pair of midgap edge states near zero quasienergy: According to the chiral symmetry condition $\Gamma H(t)\Gamma = -H(-t)$, with Γ the chiral operator, the midgap steady modes have to be of the form $|u_1^0(t)\rangle$ and $|u_2^0(t)\rangle = \Gamma|u_1^0(-t)\rangle$ with quasienergies ε_0 and $-\varepsilon_0$ respectively. It follows that the space spanned by these modes is also spanned by the two modes $|u_\pm^0(t)\rangle = (|u_1^0(t)\rangle \pm |u_2^0(t)\rangle)/\sqrt{2}$. These modes satisfy $|u_\pm^0(t)\rangle = \pm\Gamma|u_\pm^0(-t)\rangle$ implying that at $t = 0$ each of them has support on only one of the sublattices. Thus, in thermodynamic limit they correspond to two edge states localized at opposite edges of a TI.

The Floquet-Sambe Hamiltonian projected onto the space spanned by the midgap modes is given by

$$\mathscr{H}^0 = \begin{bmatrix} \langle\!\langle u_+^0|\mathscr{H}|u_+^0\rangle\!\rangle & \langle\!\langle u_+^0|\mathscr{H}|u_-^0\rangle\!\rangle \\ \langle\!\langle u_-^0|\mathscr{H}|u_+^0\rangle\!\rangle & \langle\!\langle u_-^0|\mathscr{H}|u_-^0\rangle\!\rangle \end{bmatrix} = \begin{bmatrix} 0 & \tau^0 \\ (\tau^0)^\dagger & 0 \end{bmatrix}, \tag{6.2}$$

where $\langle\!\langle u_\pm^0|\mathscr{H}|u_\pm^0\rangle\!\rangle = \int_0^T dt\, \langle u_\pm^0(t)|H(t) - i\hbar\partial_t|u_\pm^0(t)\rangle = 0$ because $|u_\pm^0(t)\rangle = \pm\Gamma|u_\pm^0(-t)\rangle$ and $\Gamma H(t)\Gamma = -H(-t)$, and we have defined $\tau^0 = \langle\!\langle u_+^0|\mathscr{H}|u_-^0\rangle\!\rangle$, implying that $|\tau^0|^2 = (\varepsilon_0)^2$. Here τ^0 must be real because $|u_1^0\rangle\!\rangle$ and $|u_2^0\rangle\!\rangle$ are by definition eigenstates of $\mathscr{H}^0$.

Thus, a general solution $|\Psi(t)\rangle$ to the time-periodic Schrödinger equation projected onto the pair of midgap states reads as

$$\begin{aligned} |\Psi^0(t)\rangle &= a_1|u_1^0(t)\rangle e^{-i\varepsilon_0 t/\hbar} + a_2|u_2^0(t)\rangle e^{i\varepsilon_0 t/\hbar} \\ &= \frac{1}{\sqrt{2}}\big[(a_1 e^{-i\varepsilon_0 t/\hbar} + a_2 e^{i\varepsilon_0 t/\hbar})|u_+^0(t)\rangle + (a_1 e^{-i\varepsilon_0 t/\hbar} - a_2 e^{i\varepsilon_0 t/\hbar})|u_-^0(t)\rangle\big], \end{aligned} \tag{6.3}$$

Here we assume that the state $|\Psi(t)\rangle$ is initialized at one of the edges and therefore it has an exponentially small overlap with the mode, say $|u^0_-(t)\rangle$, at the opposite edge at $t=0$. Note that in a finite-sized system this overlap is never exactly zero but it is generically much smaller than an overlap with the state at the same edge with $|\Psi(t=0)\rangle$. Hence, we may neglect the overlap with $|u^0_-(t)\rangle$ at $t=0$ which implies $a_1=a_2$ in Eq. (6.3). It follows that the component of state $|\Psi(t)\rangle$ projected onto the midgap states exhibits an oscillatory behaviour between the boundaries,

$$|\Psi^0(t)\rangle = |u^0_+(t)\rangle\cos(\varepsilon_0 t/\hbar) - |u^0_-(t)\rangle\sin(\varepsilon_0 t/\hbar), \tag{6.4}$$

given up to an overall prefactor representing the overlap of the full state $|\Psi(t)\rangle$ with the midgap mode $|u^0_+(t)\rangle$ at $t=0$. This is precisely what we observed in the numerical results depicted in Fig. 6.1, panels (b) and (d).

The pair of midgap edge states near $\hbar\Omega/2$ quasienergy: We now perform an analogous analysis considering the midgap states at $\hbar\Omega/2$ quasienergy. Take one of the Floquet modes near $\hbar\Omega/2$ quasienergy, say $|u^\pi_1(t)\rangle$ with quasienergy $\varepsilon_\pi+\hbar\Omega/2$. By the chiral constraint there is another Floquet mode $\Gamma|u^\pi_1(-t)\rangle$ with quasienergy $-\varepsilon_\pi-\hbar\Omega/2$. Recall that the Floquet quasienergy spectrum is intrinsically periodic. By shifting the quasienergy by one period we obtain the second mode of the pair near $\hbar\Omega/2$ quasienergy, namely $|u^\pi_2(t)\rangle = \exp(i\Omega t)\Gamma|u^\pi_1(-t)\rangle$ with quasienergy $-\varepsilon_\pi+\hbar\Omega/2$, where the phase factor pops up due to the cyclic shift in the quasienergy, cf. Chapter 4. The pair of midgap modes $|u^\pi_1(t)\rangle$ and $|u^\pi_2(t)\rangle$ are both near $\hbar\Omega/2$ quasienergy and we may now, similarly to the case with zero quasienergy, project the problem onto the space spanned by these modes, assuming that other Floquet states have negligible effect on them. The space spanned by $|u^\pi_1(t)\rangle$ and $|u^\pi_2(t)\rangle$ is also spanned by $|u^\pi_\pm(t)\rangle = (|u^\pi_1(t)\rangle \pm |u^\pi_2(t)\rangle)/\sqrt{2}$ and it is convenient to use these two modes as a basis for the projective space, again in analogy to the zero quasienergy case. These modes satisfy $|u^\pi_\pm(t)\rangle = \pm\exp(i\Omega t)\Gamma|u^\pi_\pm(-t)\rangle$. Again, this implies that each of these modes has support on only one of the sublattices at $t=0$ and therefore in thermodynamic limit they reproduce the edge states localized at opposite edges of a TI.

It follows that the Floquet-Sambe Hamiltonian projected onto the space

spanned by the modes at $\hbar\Omega/2$ quasienergy takes the form

$$\mathscr{H}^0 = \begin{bmatrix} \langle\langle u_+^0|\mathscr{H}|u_+^0\rangle\rangle & \langle\langle u_+^0|\mathscr{H}|u_-^0\rangle\rangle \\ \langle\langle u_-^0|\mathscr{H}|u_+^0\rangle\rangle & \langle\langle u_-^0|\mathscr{H}|u_-^0\rangle\rangle \end{bmatrix} = \begin{bmatrix} \hbar\Omega/2 & \tau^0 \\ (\tau^0)^\dagger & \hbar\Omega/2 \end{bmatrix}, \tag{6.5}$$

where $\langle\langle u_\pm^\pi|\mathscr{H}|u_\pm^\pi\rangle\rangle = \frac{1}{T}\int_0^T dt\,\langle u_\pm^\pi(t)|H(t) - i\hbar\partial_t|u_\pm^\pi(t)\rangle = \hbar\Omega/2$ because $|u_\pm^\pi(t)\rangle = \pm\exp(i\Omega t)\Gamma|u_\pm^\pi(-t)\rangle$ and $\Gamma H(t)\Gamma = -H(-t)$. Here τ_0 is defined analogously to the previous zero-quasienergy case.

In complete analogy to the case of the edge states at zero quasienergy, we obtain a very similar equation for a projected general solution to the time-periodic Schrödinger equation:

$$\begin{aligned} |\Psi^\pi(t)\rangle &= \left[a_1|u_1^\pi(t)\rangle e^{-i\varepsilon_\pi t/\hbar} + a_2|u_2^\pi(t)\rangle e^{i\varepsilon_\pi t/\hbar}\right] e^{-i\Omega t/2} \\ &= \frac{1}{\sqrt{2}}\left[(a_1 e^{-i\varepsilon_\pi t/\hbar} + a_2 e^{i\varepsilon_\pi t/\hbar})|u_+^\pi(t)\rangle + (a_1 e^{-i\varepsilon_\pi t/\hbar} - a_2 e^{i\varepsilon_\pi t/\hbar})|u_-^\pi(t)\rangle\right] e^{-i\Omega t/2}, \end{aligned} \tag{6.6}$$

By assuming the state $|\Psi(t)\rangle$ to be initialized at one of the edges we neglect the exponentially small overlap with the mode at the opposite edge, say $|u_-^\pi(t)\rangle$, at $t = 0$. This implies that $a_1 = a_2$ in Eq. (6.6). As expected, in this case the state's component projected onto the midgap states near $\hbar\Omega/2$ also exhibits an oscillation between the two boundary states, explicitly

$$|\Psi^\pi(t)\rangle = e^{-i\Omega t/2}\left[|u_+^\pi(t)\rangle\cos(\varepsilon_\pi t/\hbar) - |u_-^\pi(t)\rangle\sin(\varepsilon_\pi t/\hbar)\right] \tag{6.7}$$

given up to an overall coefficient representing the overlap of the full state $|\Psi(t)\rangle$ with the midgap mode $|u_+^\pi(t)\rangle$ at $t = 0$. This supports the numerical result displayed in Fig. 6.1, panels (c) and (d).

6.3.2 The evolution under time-periodic boundary perturbations

Now we investigate how various time-periodic boundary perturbations affect the evolution of the state $|\Psi(t)\rangle$ and pinpoint qualitative differences in behaviours under symmetry-breaking and symmetry-preserving time-periodic perturbations.

Let us start by taking a finite-sized Floquet TI perturbed by some arbitrary time-periodic boundary perturbation $p(t)$ localized at, say, left edge. As in Sec. 6.3.1 we initialize the state $|\Psi(t)\rangle$ at the left edge at $t = 0$ and analyze

its evolution. The perturbation $p(t)$ will generically have a considerable effect on the left edge modes while the effect on the right edge modes will be exponentially small. The states localized on the opposite sides will therefore split in quasienergy and any misalignment between the quasienergy levels generally decreases hybridization between the states. Thus, a generic time-periodic perturbation $p(t)$ will significantly reduce hybridization between the edge states, destroying the oscillatory behaviour of $|\Psi(t)\rangle$. However, if the perturbation $p(t)$ is not arbitrary but taken to be symmetry-preserving, it will not have any effect on the quasienergies of the edge states (up to finite-size corrections), and the oscillatory behaviour of $|\Psi(t)\rangle$ will remain present. This simple analysis results in a qualitative probe of the robustness of the edge states in Floquet TIs against time-periodic symmetry-preserving perturbations.

A numeric test confirming the analysis above is depicted in Fig. 6.2: We perturb a harmonically driven SSH chain in a nontrivial topological phase by various time-independent and time-periodic perturbations and track how the state $|\Psi(t)\rangle$ responds to the perturbations. In agreement with the analysis above, the density oscillation survives symmetry-preserving perturbations while being destroyed by symmetry-breaking perturbations: In the case of the periodically driven SSH model the chiral symmetry is maintained for any perturbation of the chemical potential that is odd in time, as well as for any perturbation of the hopping amplitudes that is even in time.

As before these results can be analytically explained by projecting the problem onto the midgap states. Here we consider a Floquet TI with only one pair of midgap modes and assume that they are near zero quasienergy. The case with multiple symmetry-protected states per edge corresponding to both quasienergy gaps can be considered analogously. The perturbed Floquet-Sambe Hamiltonian projected onto the states $|u^0_\pm(t)\rangle$ is given by the following expression

$$\mathscr{H}^0 = \begin{bmatrix} \delta h_L & \tau^0 + \delta\tau \\ (\tau^0 + \delta\tau)^\dagger & \delta h_R \end{bmatrix}, \tag{6.8}$$

with $\delta h_{L/R} = \frac{1}{T}\int_0^T dt\, \langle u^0_\pm(t)|p(t)|u^0_\pm(t)\rangle$ and $\delta\tau = \frac{1}{T}\int_0^T dt\, \langle u^0_+(t)|p(t)|u^0_-(t)\rangle$.

The corrections δh_R and $\delta\tau$ are exponentially small because they depend on the overlap between $p(t)$ and $|u^0_-(t)\rangle$ that are localized at different edges. The correction δh_L is in general not small and destroys the oscillation: This can be explicitly proved by solving Eq. (6.8) for the eigenstates

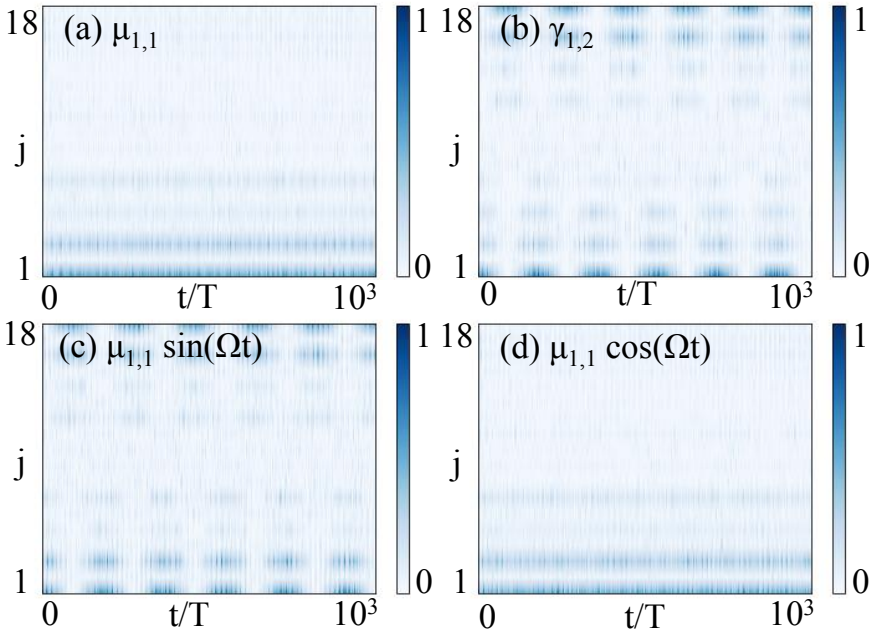

Figure 6.2: The density $|\Psi_j(t)|^2$ vs. time t and site number j: We perturb a driven SSH chain of 18 sites, hopping amplitudes $\gamma_1 = 1.2$ and $\gamma_2 = 0.8$ using boundary perturbations (a) $\mu_{1,1} = 0.2$, (b) $\gamma_{1,2} = 0.5i$, (c) - (d) $\mu_{1,1}\sin(\Omega t)$ and $\mu_{1,1}\cos(\Omega t)$ with $\mu_{1,1} = 0.5$. Here $\mu_{1,1}$ denotes the chemical potential localized at $j = 1$, and $\gamma_{1,2}$ is the hopping amplitude between first two sites. The driving is fixed to $v(t) = 0.4\cos(\Omega t)$ with $\Omega = 3$ and $T = 2\pi/\Omega$. The SSH chain is in a nontrivial phase associated with indices $\nu_0 = 0$ and $\nu_\pi = 1$. The numerical time interval was chosen to be $\delta t = 0.01T$.

and showing that the initialization of $|\Psi(t=0)\rangle$ at the left edge results in trapping the state $|\Psi(t)\rangle$ at the left edge at all times t. This calculation is straightforward but we skip it here because it brings unnecessary cumbersome expressions into the discussion. However, if the applied perturbation $p(t)$ respects the chiral symmetry, namely $\Gamma p(t)\Gamma = -p(-t)$, then the correction $\delta h_L = \frac{1}{T}\int_0^T dt\, \langle u_+^0(t)|p(t)|u_+^0(t)\rangle = -\delta h_L = 0$. Thus, the correction δh_L is identically zero if the time-periodic perturbation maintains the chiral symmetry, implying that the oscillatory behaviour of the state $|\Psi(t)\rangle$ is not affected by the perturbation. This simple analysis establishes how the symmetry protection extends to finite-sized Floquet TIs and provides a symmetry-based explanation of the results presented in Fig. 6.2.

6.3.3 Time-independent TIs under time-periodic perturbations

Recall that any time-independent system is trivially periodic in time and therefore the Floquet robustness in time-independent TIs can be established on the same grounds as in Floquet TIs.

Different from time-periodic chiral phases, the time-independent chiral TIs satisfy the chiral relation $\Gamma H(t_0+t)\Gamma = -H(t_0-t)$ for any reference time t_0. It is therefore sufficient to maintain the chiral relation for at least one reference time t_0 to ensure the symmetry protection of the edge states. Thus, in the case of the ordinary time-independent SSH model any harmonic boundary perturbation of the chemical potential respects the chiral relation above and therefore this perturbation will not affect the oscillatory evolution of $|\Psi(t)\rangle$ between the edges. On the other hand, by adding a second harmonic to the perturbation we may break the chiral relation for all reference times t_0 and the oscillatory behaviour is expected to be suppressed. The numerical results presented in Fig. 6.3 agree well with this symmetry analysis: The density oscillation between the edges disappears once the chiral symmetry is broken for all reference times t_0 but remains if the symmetry is maintained. Note that in Fig. 6.3 (d) we had to considerably increase the amplitudes of the applied perturbations to see the anticipated behaviour because the effect of the second harmonic is generically very weak, see Paper A.

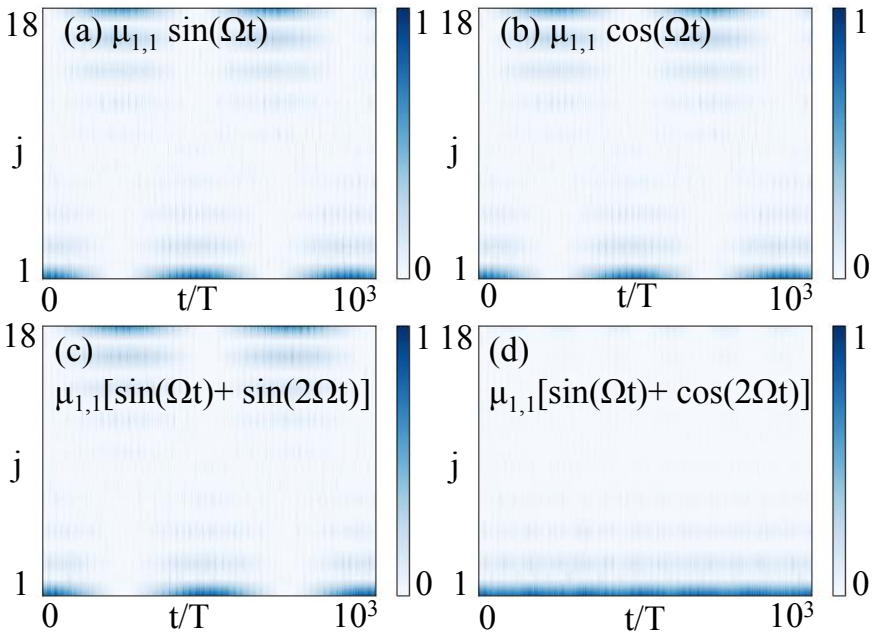

Figure 6.3: The density $|\Psi_j(t)|^2$ vs. time t and site number j: We take a static SSH chain of 18 sites, hopping amplitudes $\gamma_1 = 0.45$, $\gamma_2 = 0.75$ and perturb it with time-periodic boundary perturbations (a) $\mu_{1,1}\sin(\Omega t)$, (b) $\mu_{1,1}\cos(\Omega t)$, (c) $\mu_{1,1}[\sin(\Omega t)+\sin(2\Omega t)]$, and (d) $\mu_{1,1}[\sin(\Omega t)+\cos(2\Omega t)]$. Here $\mu_{1,1} = 4$ corresponds to the chemical potential localized at $j = 1$. The driving is set to $v(t) = 0.4\cos(\Omega t)$ with $\Omega = 3$ and $T = 2\pi/\Omega$. Here we employed the numerical discretization interval $\delta t = 0.01\,T$.

6.3.4 Experimental Test

To conclude this section let us briefly comment on the feasibility to experimentally observe the density oscillations discussed above. The recent experimental achievements in realizing topologically nontrivial matter within photonic setups [90, 91, 92] here appear very promising. In Ref. [91] a photonic system effectively identical to an SSH chain subject to various time-periodic local perturbations has been realized in a lab. The behavior of the corresponding edge states were then studied using leakage radiation microscopy and intensity distribution measurements. From a realization point of view, the density probes presented above are expected to be conceptually similar to the edge state stability study conducted in Ref. [91], making the photonic crystals a viable platform for observing the density oscillations in a carefully designed experiment. One should also mention here that the SSH model has recently been realized also in the solid state [93, 94] and with cold atoms [95, 96, 97]. The periodic drives have also been discussed within these settings [96], providing an alternative class of systems to look at for a potential experimental realization. Lastly, we note that the density probes are not just limited to SSH chains but can in principle be observed in any realization of a TI in 1d. Thus, in designing an experiment one may look for alternative models of 1d TIs that might be easier to implement in a lab.

7 Essentials of Artificial Neural Networks

An artificial neural network (NN) is a computational object constructed by interconnecting elementary nodes, called artificial neurons. The neurons are computational units designed to perform the following basic operation:

$$O_1 = g\left(\sum_i^N w_{1i}\, x_i - b_1\right), \tag{7.1}$$

where $\boldsymbol{x}$ is the input vector and O_1 is the neuron's output, Fig. 7.1a. The so-called weights w_{1i} and threshold b_1 constitute the neuron's variational parameters. The activation $g(x)$ is a predefined function introduced to make the neuron's action nonlinear and by this more nontrivial. It can be chosen to be any continuous function. Some examples of the most conventional activation functions are listed below:

$$g_1(x) = \tfrac{1}{1+e^{-x}}; \quad g_2(x) = \tanh(x); \quad g_3(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}. \tag{7.2}$$

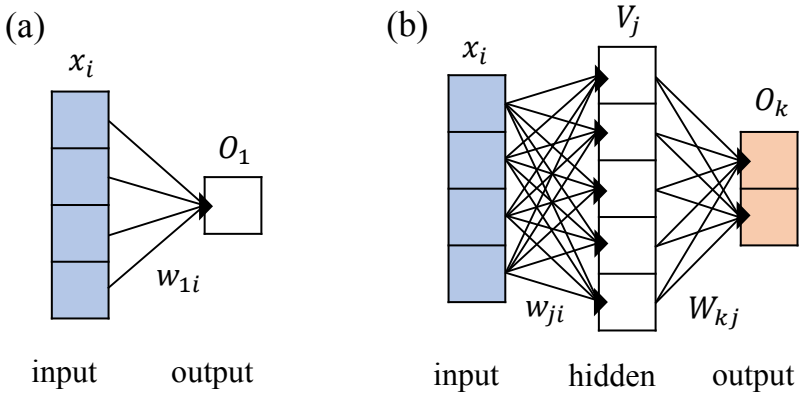

Figure 7.1: Schematic illustrations of a) one artificial neuron and b) a simple fully-connected neural network with one hidden layer.

7.1 Feed-forward neural networks

Artificial neurons are used as elementary building blocks for creating various NN types suitable for different practical tasks. In this thesis the focus is placed on so-called feed-forward NNs. The neurons in a feed-forward NN are strictly connected layer to layer: Connections within layers, back connections, connections skipping some layers are forbidden, Fig. 7.1b. The leftmost and rightmost layers are called input and output layers, respectively. The input layer consists of input terminals and does not perform any calculation. The output layer produces the final outcome of the network. The other layers, so-called hidden layers, consist of nodes that receive an input, perform a calculation, and feed the result to the next layer. A deep NN is a network having at least one hidden layer. To make the construction more versatile and nontrivial, it is common practice to have some hidden layers deviating from the conventional definition of neurons, Eq. (7.1). Usually these layers do not have any variational parameters and simply perform a predefined non-linear operation on the outputs of previous layer. Essentially any layer performing a piece-wise differentiable operation can be adapted to the network making the construction extremely flexible and powerful. The feed-forward network is basically a high-dimensional nonlinear ansatz $\boldsymbol{O}(\boldsymbol{x};\boldsymbol{\Theta})$, where $\boldsymbol{\Theta}$ consists of all weights and thresholds constituting variational parameters of the ansatz. The form of ansatz can be varied by changing types of the layers, their order, number of neurons per layer, activations, etc. There is an enormous flexibility in constructing the network and most choices are done via experience utilizing the trial and error approach. It is interesting to note here that the ansatz corresponding to a sufficiently large network with a single hidden layer can actually represent any continuous function to arbitrary precision, a result known as the universal approximation theorem [98].

The feed-forward NNs are commonly used for supervised data classification problems. In supervised learning one starts with a large set of manually labeled data $\{(\boldsymbol{x}^\mu, \boldsymbol{y}^\mu)\}$, with μ indexing all dataset elements, and aims to use this data for training the network: The objective is to find the net's variational parameters that will fit the outputs $\boldsymbol{O}(\boldsymbol{x}^\mu;\boldsymbol{\Theta})$ to the given labels $\boldsymbol{y}^\mu$ with high accuracy. The trained network may then be used for classifying new data with unknown labels. Alternatively, one may try to 'open up' the network aiming to interpret the network's computation. Usually, the networks are very complicated and their layout does not allow the user to interpret the

learned features. Remarkably, however, for some specific tasks this becomes possible. In Papers D and E we introduced a type of NNs designed to fit to particular representations of topological indices, cf. Sec. 8.2. Our protocol and results prove by example that the NNs can be used for finding some concrete representations of the features and then be interpreted by looking at the intermediate outputs.

7.2 Gradient-descent learning, Backpropagation

The training of the networks is commonly done via introducing the so-called loss (or cost) function $H(\boldsymbol{\Theta})$ pointing towards the deviation of the network outputs $\boldsymbol{O}(\boldsymbol{x}^\mu;\boldsymbol{\Theta})$ from true labels $\boldsymbol{y}^\mu$. As an example, consider the mean square error function, explicitly given by

$$H(\boldsymbol{\Theta}) = \frac{1}{2}\sum_{\mu,i}\left(y_i^\mu - O_i(\boldsymbol{x}^\mu;\boldsymbol{\Theta})\right)^2. \tag{7.3}$$

This function is always positive and has a global minimum where the outputs $\boldsymbol{O}(\boldsymbol{x}^\mu;\boldsymbol{\Theta}) = \boldsymbol{y}^\mu$ for every μ: the smaller the cost function, the better the outputs fit to the given labels. We thus train the network by minimizing the cost function via altering the variational parameters of the network. The minimization is done using gradient-descent learning. We numerically move in the parameter space of all possible $\boldsymbol{\Theta}$ in the direction opposite to the gradients:

$$\delta\Theta_k = -\eta\frac{\partial H(\boldsymbol{\Theta})}{\partial\Theta_k}. \tag{7.4}$$

Taking a small enough learning rate η, the cost function $H(\boldsymbol{\Theta})$ will always decrease or stay constant after each iteration step. In this formulation the gradient-descent algorithm can get stuck in a local minimum and it does not have any mechanism to get out. The local minima problem can be made less severe by adding some stochasticity to the training, performing the training in so-called minibatches. The minibatches partition the training set into small random subsets of some specified size. The iteration steps are then done by performing gradient-descent learning using distinct minibatches instead of the entire training set. After the NN has trained on each element $(\boldsymbol{x}^\mu, \boldsymbol{y}^\mu)$ exactly once we say that the training has completed one training epoch. After each epoch the partitioning of the training set into minibatches

is reset and the training is continued until the cost function is acceptably small.

In practice, the stochastic gradient-decent learning for feed-forward nets is programmed employing the so-called error backpropagation [20]. The concept is easiest explained via an example, using the network schematically depicted in Fig. 7.1b. It consists of an input layer, one hidden layer, and an output layer. The input, the hidden layer's outcome, and the output are denoted by $\boldsymbol{x}$, $\boldsymbol{V}$, and $\boldsymbol{O}$, respectively. Each pair of nodes from the input and hidden layers are connected, the same as each pair from the hidden and output layers. The weights and thresholds associated with the hidden layer, say w_{mn} and b_m, are enumerated using indices m and n running over all nodes in the hidden and input layers, respectively. Analogously, the weights and thresholds associated with the output layer are denoted by W_{mn} and B_m. The cost function is chosen to be the mean square error function, see Eq. (7.3). For obtaining the gradient-descent increments, Eq. (7.4), the partial derivatives are calculated using the conventional chain rule.

Consider first the increments δW_{mn} and δB_m for the weights and biases associated with the output layer:

$$\begin{aligned} \delta W_{mn} &= -\eta \frac{\partial H(\boldsymbol{\Theta})}{\partial W_{mn}} = \eta \sum_{\mu} \Delta_m^{\mu} V_n^{\mu}, \\ \delta B_m &= -\eta \frac{\partial H(\boldsymbol{\Theta})}{\partial B_m} = -\eta \sum_{\mu} \Delta_m^{\mu}, \end{aligned} \tag{7.5}$$

where the so-called error Δ_m^{μ} is given by

$$\Delta_m^{\mu} = \big(y_m^{\mu} - O_m(\boldsymbol{x}^{\mu}; \boldsymbol{\Theta})\big)\, g'(C_m^{\mu}), \tag{7.6}$$

with $C_i^{\mu} = \sum_j (W_{ij} V_j^{\mu} - B_i)$ and $g'(C_i^{\mu}) = d\,g(t)/d\,t$ evaluated at $t = C_i^{\mu}$. Note that the nodes of the hidden layer do not have incoming connections from the output layer and therefore the values V_n^{μ} do not explicitly depend on W_{mn} and B_m. This is an important consequence of the feed-forward layout of the network.

The learning increments δw_{mn} and δb_m associated with the hidden layer

can be calculated in an analogous way:

$$\begin{aligned}\delta w_{mn} &= -\eta \frac{\partial H(\mathbf{\Theta})}{\partial w_{mn}} = \eta \sum_{\mu} \delta_m^{\mu} x_n^{\mu}, \\ \delta b_m &= -\eta \frac{\partial H(\mathbf{\Theta})}{\partial b_m} = -\eta \sum_{\mu} \delta_m^{\mu},\end{aligned} \tag{7.7}$$

where the error δ_m^{μ} is given by

$$\delta_m^{\mu} = \sum_i \Delta_i^{\mu} W_{im} g'(c_m^{\mu}), \tag{7.8}$$

with $c_i^{\mu} = \sum_j (w_{ij} x_j^{\mu} - b_i)$ and $g'(c_i^{\mu}) = dg(t)/dt$ evaluated at $t = c_i^{\mu}$. The dependence of the error function δ_m^{μ} on the next layer's error function Δ_m^{μ} generically originates from the chain rule: The partial derivatives propagate backwards from the output layer to the input layer.

The example above gives a taste of a generic pattern of equations that one will get for training any feed-forward network: The increments will strictly depend only on the errors of next layers, a consequence of the chain rule applied to the feed-forward layout of the network. Thus, the numerical implementation of the training can be divided into three main steps, see Algorithm 1. These three steps lie at the core of neural nets: It follows that basically any computational operation through which one can backpropagate can be painlessly incorporated into the network. This makes the networks very flexible and efficient.

Concluding this section, it is worth mentioning that over the years there have been various proposals for upgrades of the stochastic gradient descent

Algorithm 1: Forward-Backward propagation

1. **Forward propagation**: Propagate the input forward through the network for obtaining output at every node.
2. **Backpropagation**: Propagate the errors backward through the network for calculating the learning increments.
3. Update the variational parameters and repeat the training.

method: AdaGrad [99], RMSProp [100], Adadelta [101], Ftrl [102], and others. One particular optimization algorithm, named Adam [103], has revolutionized the field. It is numerically efficient, has little memory requirements, and has only four hyperparameters to be externally specified. Without going into much detail, Adam assigns individual learning rates to distinct variational parameters (weights, thresholds) and optimizes the training by automatically decreasing the learning rates of noisy parameters with large derivative variations. Adam is surprisingly robust to the choice of hyperparameters. The parameters set by default are found to work extremely well in a wide range of applications. The only parameter one may consider to optimize in practice is the global learning rate, i.e. the learning rate assigned to each parameter before normalization by a 'variance'-dependent factor that depends on the remaining three hyperparameters [103].

7.3 Overfitting, Vanishing gradients

Nets in supervised learning are primarily trained to generalize to new data. A net with more neurons, and therefore more variational parameter, will in general fit better to the training set because it has more degrees of freedom to adjust. However, it does not necessary mean that the network will learn the correct features and will effectively generalize to new data. To the contrary, a large network easier adjusts its parameters to fit to unique features of the training data. In this way the network will exhibit great performance on the training set without learning the desired general features and it will not be applicable in practice. The tendency of a network to fit to the training set is generally known as overfitting.

There is no universal method for keeping the network from overfitting. However, one may still get good intuition on various approaches found to be practically effective in this regard. One way to deal with overfitting is to reserve some part of the training dataset for validating the performance throughout the training. This data, known as validation data, is separated from the training set and is not used at any stage of the training. The idea now is to keep the validation set away from the training process and only use it as an indicator to terminate the training before the overfitting takes place: The training is terminated when the accuracy evaluated on the training set still continues to improve but the performance evaluated on the validation

set saturates or even gets worse. This procedure is known as early stopping: We simply stop the training at the peak performance.

The overfitting can also be reduced by playing with the layout of the net. It is intuitive to expect the overall performance to improve by averaging outcomes of different learning experiments. To do the training multiple times by changing each time the layout or initial guess for variational parameters will most likely be effective, however a very costly procedure. One may obtain a similar effect at much smaller computational cost via employing a special layer called dropout [104]. At each minibatch training iteration, the dropout layer allows to propagate through only a random fraction of incoming neurons while nullifying the others. For validating and testing, all neurons are brought back. In this way the network becomes effectively very sparse and covers different layouts throughout the training. This is believed to have a positive effect on the training, in particular it helps against overfitting. A similar effect can be obtained by so-called pruning, where the weights that remain very close to zero are effectively removed [105, 106].

Another common approach to reduce the overfitting is based on the following observation: The networks with large weights amplify small changes in input and may drastically change the output. This means that any finest detail in the input may be used for differentiating between the data elements, which is precisely the sign of overfitting. Thus, one would want to prevent the weights of becoming very large. For doing so it is a common practice to add an extra term to the cost function penalizing the weights for being too large. There are two common ways for choosing penalizing terms, the so-called L_1- and L_2-Regularizations. Within these regularization schemes the cost function obtains the following extra terms

$$\Delta H_{L_1}(\mathbf{\Theta}) = \lambda_1 \sum_m |w_m|; \quad \Delta H_{L_2}(\mathbf{\Theta}) = \lambda_2 \sum_m |w_m|^2, \tag{7.9}$$

where w_m denotes all weights to which the regularization is applied, and $\lambda_{1/2}$ are hyperparameters needed to be well tuned by the user.

Let us also comment on another fundamental issue in deep learning, known as the vanishing-gradient problem. This problem is a direct consequence of the chain rule employed for calculating the weight and bias increments, cf. Sec. 7.2: Within the feed-forward net layout the increments are composed of sequentially multiplied derivatives and therefore the increments are generically very small for the layers close to the input. Similar

to the case with overfitting, the vanishing-gradient problem cannot be entirely resolved, however, by applying different techniques it can be made significantly weaker. In particular, to enhance the learning one may try to avoid the gradients to decrease, and therefore it is a common practice to give a preference to rectified linear activation with unit derivative, cf. $g_3(x)$ in Eq. 7.2, than to other activation functions. Another possibility is to purposely introduce new connections to the network that will connect layers that are far from each other, creating so-called residual neural networks [107]. One can also try to normalize the data. For example, the so-called batch normalization technique [108], where the outputs are normalized over the whole batch before feeding to the next layers, has been shown to be particularly helpful here.

To conclude, deep learning is a powerful machine learning tool specialized in finding highly nontrivial patterns in big datasets. In supervised learning one trains a net to fit to the known data, aiming the network to pick up only relevant features capable of generalizing to unseen data. The training is done via gradient-descent learning by minimizing the so-called cost function representing the deviation of the current network state from the desired performance. The networks are in general universal and can in principle reconstruct any continuous function. Their training, however, is a very nontrivial task because they tend to learn irrelevant features that do not generalize to new data. In this chapter we have also commented on two fundamental issues accompanied with the training procedure, the overfitting and vanishing-gradient problems. There is no universal solution to overcome these problems but with some understanding one can significantly reduce their effect on training by employing some of the techniques discussed above.

8 Convolutional Neural Networks

In Chapter 7 we have mainly looked at neural networks as being built entirely from the so-called fully-connected (dense) layers, where each node in a layer has a connection with each node in the next layer, see Fig. 7.1. One may then naively expect the feed-forward networks built entirely from dense layers to exhibit the most optimal performance because any type of a deep network can be represented as a large fully-connected net with some weights and biases restricted to be zero or else having some predefined relations between each other. However, this is not what we see in practice: In most cases the classification accuracy will not improve by simply adding more and more hidden layers to the networks, neither by taking more neurons per layer. By making a fully-connected network larger one drastically increases the number of variational parameters and the network becomes very hard to train and it tends to overfit very easily. Instead, the training can be boosted by changing the core layout of the network. Using a trial and error approach one then aims to design a neural network that will be acceptably efficient for a given task. A particular family of layouts, dubbed convolutional neural networks, has been found particularly effective in a wide range of recognition tasks. In Papers E-D we introduce a special type of convolutional nets, and use them in combination with concepts of 'learning by confusion', Sec. 8.2.1, and 'topological data augmentation', Sec. 8.2.2, to retrieve topological indices.

In 2012 Krizhesvky et al. [109] won the ImageNet challenge employing a convolutional network, dubbed the AlexNet. This revolutionized the field of machine learning. This work is considered to be one of the most influential papers of all times published in machine learning. The convolutional networks, first introduced by Fukushima [110] in 1980, facilitate the training by reducing the number of variational parameters per layer compared to fully-connected networks. By construction they inherit our basic understanding of pattern recognition via extraction of features. Most importantly the convolutional nets work well and considerably outperform the fully-connected networks in most recognition tasks. The base of any conventional convolutional network consists of several so-called convolution layers followed by so-called pooling layers. In fact, the AlexNet is constructed entirely from several convolution, pooling and dense layers. Below we briefly outline the operating principle of these layers.

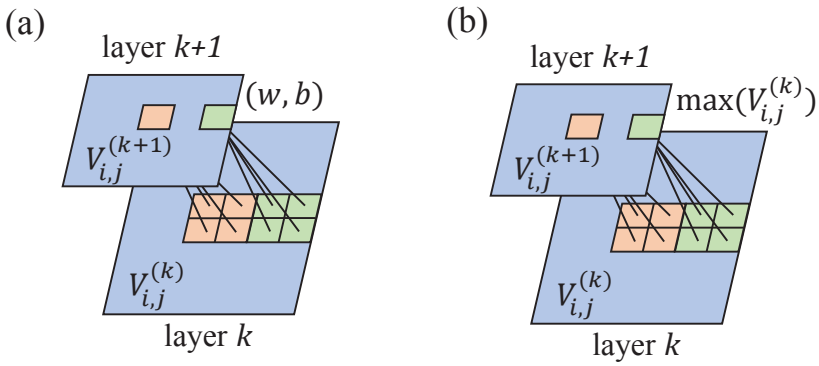

Figure 8.1: A schematic illustration of the operating principles of a) convolution and b) max-pooling layers.

8.1 Convolution and Pooling Layers

In Fig. 8.1a we give a simple example illustrating the operating principle of a convolution layer. Here the input nodes $V_{ij}^{(k)}$ and outputs $V_{ij}^{(k+1)}$ are related via

$$V_{ij}^{(k+1)} = g\left(\sum_{p=0}^{1}\sum_{q=0}^{1} w_{pq}^{(k)} V_{i+p,\,j+q}^{(k)} - b^{(k)}\right). \tag{8.1}$$

There are two essential details to highlight in Eq. (8.1). First of all, the input $V_{ij}^{(k)}$ and output $V_{ij}^{(k+1)}$ are given as two-dimensional arrays and only input nodes that are locally close to each other have connections to a common node in the output layer. In analogy to the terminology used in neuroscience, the area of input connecting to the same node is called the receptive field. Second, every output node is connected to the input using the same set of weights $w_{pq}^{(k)}$ and bias $b^{(k)}$. The set of such variational parameters builds up a so-called filter, also dubbed kernel. By convoluting the kernel with input nodes, Eq. (8.1), we extract a so-called feature and the entire output layer corresponding to one filter is then termed a feature map. Thus, one filter represents some local quantity of interest and the corresponding feature map points to where this quantity is located in the input layer. For example, in image recognition one may think of filters as of local visual objects like certain types of edges, lines, color gradients, etc. A two-dimensional input

layer to feature map transformation, Eq. (8.1), can be generalized by stacking multiple input arrays on top of each other, creating so-called input channels. Each channel is associated with an independent set of weights but with the same bias. The output layer can also be generalized to a stack of feature maps, each corresponding to a different filter. The size of receptive field and rate at which the receptive field sweeps through the input layer, the so-called stride, are hyperparameters to be specified by the user.

The convolution layers often work better in combination with so-called pooling layers. A pooling layer downsamples the feature maps by applying to them a local predefined function, typically average or max. Similar to the convolution layers, a pooling layer applies its operation in correspondence to some receptive field and stride specified by the user. In Fig. 8.1b we schematically illustrate a commonly used max-pooling layer. Explicitly, in this example it transforms the feature map $V_{ij}^{(k)}$ using the rule

$$V_{ij}^{(k+1)} = \max_{p,q\in\{0,1\}} V_{p+i,q+j}^{(k)}, \tag{8.2}$$

where the size of the receptive field is 2×2 and that of the stride is 1×1. Analogously to the convolution layers, the pooling layers can be generalized to take in a stack of two-dimensional data arrays. The predefined pooling function then operates separately on each channel, creating multiple two-dimensional output arrays. Note that pooling layers do not bring any variational parameters to the network ansatz. However, since the pooling function is piecewise continuous it can be straightforwardly incorporated into the training via forward-backward propagation. This layer is believed to be effective because it makes the convolution layers less sensitive to small changes in the input and therefore it helps to reduce the risk of overfitting.

To sum up, a convolution layer (extended by a pooling layer) produces a number of feature maps out of a stack of two-dimensional input data. These feature maps create a representation of input in terms of some local quantities and how they are positioned in the data. One can then use a sequence of convolution layers where an output of one layer is inputted into the next layer, in this way representing more and more abstract local features and their relations between each other, eventually reaching a highly nontrivial level. The sequence of convolution-pooling layers is conventionally extended by a few fully-connected layers just before the output layer. They allow the network to intertwine the learned local features corresponding to different spatial

parts of the two-dimensional input. These three ingredients of convolutional neural networks constitute the base of many widely applied neural networks.

8.2 Neural-network-based topological classification

In Papers D and E we introduce a novel neural-network-based protocol for finding topological indices without using any explicit supervision. Below we briefly summarize the main ideas and results of these two works.

8.2.1 Learning by confusion

Recall that the neural-network-based protocols commonly require training on some already labeled data before they become capable of recognizing any relevant features. The labeling is usually done manually and this seems to limit the applicability of these methods for tasks of learning properties unknown to the user. The 'learning by confusion' concept, recently introduced in Ref. [21], opens new horizons for the conventional supervised training by adapting it for revealing user-unknown information.

The core principle of 'learning by confusion' methodology bases on 1) creating presumably distinct sets of data and 2) training a neural network to distinguish the data by assigning 'dummy' labels to the datasets and then doing standard supervised training. In other words, we 'pretend' that the datasets are distinguishable and ask the network to find the corresponding distinguishing feature(s). Now, there are only two possibilities: 1. The considered datasets contain a distinguishing property. In this case the neural net will generically find it and the classification attempt will be successful. 2. The datasets are indistinguishable. The network will fail and the resulting test accuracy will always fluctuate around a value corresponding to random guessing. In the first case the procedure outputs a quantity that was not used anywhere for the training, i. e. the trained network has learned a new property generally unknown to the user. In the second case the procedure does not provide any useful quantity and one may then try out another 'guess' and repeat the training. Thus, the procedure is successful only if the network is successfully trained. Note that for retrieving the learned quantity from the network one would want to use an 'interpretable' layout of the network which assists in the interpretation of the learned feature. This, however, is in general not so easy to implement.

8.2.2 Topological data augmentation

In Papers D and E we adopt the 'learning by confusion' principle for finding topological indices. Our basic idea here is to use the defining property of topology, namely its invariance under continuous deformations, to generate extensive training data for a neural network. The method is inspired by data augmentation techniques widely used to boost datasets in the context of supervised machine learning. We dub this methodology 'topological data augmentation'. The datasets are explicitly produced by performing random continuous deformations on some selected seed states. By allowing only continuous deformations, the generated datasets are intrinsically composed of topologically equivalent data. Note that it is not necessary to know anything about the topological indices for continuously deforming the objects. Thus, the produced data may contain information on topological indices that was not put in by the user. To retrieve it, we suggest to use the 'learning by confusion' methodology: A neural network, trained to separate the datasets by assigning to them 'dummy' labels, picks up a quantity capable of differentiating between the topological equivalence classes. The learned quantity can then be numerically retrieved with the help of a special type of customized convolutional neural networks introduced in Papers D and E, cf. Sec 8.2.3.

A numerical representation of any geometric object is intrinsically discrete, creating a problem with the numerical definition of equivalence via continuous transformations. In Paper D we resolve this problem by explicitly developing two continuity conditions for the considered classes of systems, 2-band insulators in 1d and 2d. These two conditions embed the discrete objects into the formalism of continuous transformations. We should note here that the derived continuity conditions do not explicitly depend on the knowledge of the topological indices, but do require good understanding of the structure of the state space. Thus, one may argue that in a more complicated scenario these conditions might be very challenging to derive. In Paper E we take an attempt to generically solve this difficulty. We suggest to only look at the states which vary slowly with the momentum at all stages of the deformation procedure. In this way we automatically ensure the continuity of the transformations and the states. This is a significant improvement over the procedure described in Paper D which allows us to consider more complicated objects.

Importantly, in Paper E we extend the parent states by a number of trivial atomic bands before deforming them. This process removes the possibility for a network to learn irrelevant local topological indices that distinguish between trivial atomic-limit Hamiltonians. The deformation procedure is also significantly improved in Paper E compared to Paper D: The matrices corresponding to the neighboring momentum sites are now recursively deformed in groups, with the primary aim to preserve the discrete representation of continuity at each deformation step. This process is analogous to the bending of a rubber string where, due to its continuity, the points close to each other must move together.

8.2.3 Representing topological indices via customized convolutional nets

An important component in our protocol for finding topological indices is the layout of employed neural networks developed by us specifically for this purpose. Recall that the neural networks are used here as a tool for finding a distinguishable feature in the created sets of topologically equivalent data. In general, one can take any deep neural network and train it to differentiate between the datasets. However, this approach will have two big disadvantages: We will not be able to interpret the established feature and the trained network will generically not generalize to new data, for example, coming from other equivalence classes. Thus, the applicability of the methodology will be limited. To overcome this limitation we suggest to slightly change the strategy of the 'learning by confusion' principle: Instead of aiming to simply distinguish between the datasets at all costs, i.e. using any neural network suited to do the job, we suggest to do the training using a neural network capable of representing only a particular generic type of functions. Within this perspective, in a case of successful training the layout of the network allows us to interpret the quantity which has been identified. A failure to train the network to separate the datasets may now mean two things: 1) The chosen datasets were from the same equivalence class or 2) The feature distinguishing the datasets can not be represented by a neural network used for the training. Both failure scenarios require the procedure to be repeated. This is done either by repeating the data augmentation procedure or trying out a different interpretable network layout. We note here that we can identify the failure problem, deciding whether it is of type 1) or 2), by checking if

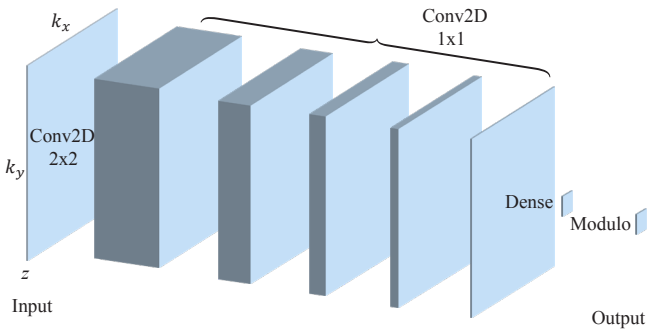

Figure 8.2: A 2d layout of a neural network introduced in Paper E. The input layer takes in data in format $k_x \times k_y \times z$, where k_x and k_y correspond to the momentum points in 2d, and z denotes all local degrees of freedom. The network consists of a Conv2D layer with receptive field of size 2×2 followed by several Conv2D layers with receptive field of size 1×1, a Dense layer, and a Modulo layer.

the datasets are distinguishable by a general type of network: If the training in this case has also failed then there is no feature distinguishing the datasets (type 1), otherwise, there is a feature but the interpretable layout could not capture it (type 2).

In Papers D and E we introduce two neural network layouts that can be used for finding topological indices via the interpretable 'learning by confusion' principle described above. In Paper D we describe a layout of convolutional neural networks that can be seen as a discrete representation of an integral over some curvature function: The network consists of several convolution layers, applying the same operation at each momentum site and then summed up to give a label. This layout was extended in Paper E to represent a more general type of functions. The idea was to replace the summation layer by a dense layer with one output node, linear activation function, zero bias, and weights restricted to take values between 0 and 1. Note that within this construction the network obtains a possibility to identify high-symmetry points, lines, and planes, and then perform summations over them. In Paper E we also added a new layer performing a predefined function on the output of a last dense layer. The predefined function is intended to mimic a modulo function over some integer value specified by the user. A

schematic illustration of the network layout employed in Paper E is depicted in Fig. 8.2. Thus, the introduced networks represent a limited but generic class of functions that most importantly represents a large class of topological indices.

9 Summary and Outlook

9.1 Floquet topological quantum matter

The topological classification of time-periodic band insulators is done on the same conceptual footing as for the time-independent ones. The idea is to define equivalence classes for the evolution operators and put the operators into the same class if and only if they can be continuously transformed into each other without closing the quasienergy band gap. The periodicity of the quasienergies then brings additional topological structure to the Floquet phases and formally results in two independent topological invariants counting the midgap boundary states in, respectively, conventional gap and in the so-called anomalous gap (at the edge of the Floquet zone). The latter topological index is non-trivial only in time-periodic systems and has no analogue in the time-independent formalism. Its appearance can also be explained by looking at two different mechanisms to obtain quasienergy band inversions, with the Floquet-conduction band or the Floquet-valence band on top of the other one. The band inversions usually signal the topological phase transitions and by this suggest that there indeed should exist two independent topological invariants associated with two different gaps, at 0 an $\hbar\Omega/2$.

In Papers A-C we focused on studying the topological robustness of the edge states in 1d Floquet topological band insulators. The emphasis was placed on the protection against time-periodic perturbations, allowed to be considered within the Floquet theory on fundamentally the same grounds as conventional time-independent perturbations. The robustness was established for both flavors of the Floquet symmetry-protected edge states, at quasienergies 0 and $\hbar\Omega/2$. Moreover, the protection of the edge states in conventional time-independent topological insulators against time-periodic perturbations was also discussed. Notably, it was found that the time-independent edge states show resilience to a large class of symmetry-breaking time-periodic perturbations, in addition to symmetry-preserving ones.

A step closer to an experimental test to observe the discussed topological robustness was taken in Paper B and Sec. 6.3. In Paper B we suggested that the fingerprints of the topological symmetry protection of the edge states can be seen in transport measurements. The idea is to bring the edge states close

to each other and by this open up a channel for transport in an otherwise insulating system. The exponential localization of the edge states then still protects them from symmetry-preserving perturbations, including time-periodic ones. All transport calculations were done by using the approach developed in Paper C, where the expressions for time-periodic currents and densities were derived within the Floquet-Sambe formalism. In Sec. 6.3 we discussed another probe of the Floquet symmetry protection: It is based on tracking the state time-evolution initialized at one edge of a perturbed finite-sized topological insulator. It was shown that the state exhibits an oscillatory behaviour between the edges if the chiral symmetry is maintained, but generically localizes at one of the edges if this symmetry is broken.

The robustness of the boundary states against symmetry-preserving time-periodic perturbations can be established also for Floquet topological insulators in higher dimensions and this provides a nice direction for further exploration. Probably the most accessible choice will be to study time-periodic perturbations of the edge states in 2D systems from the AII symmetry class. This is because the preservation of time-reversal symmetry can be analyzed in a similar fashion to the chiral symmetry. The particle-hole symmetry, however, is not expected to bring any interesting phenomena in this context because it is local in time, and therefore its maintenance is trivial: it has to be preserved at each instance of time. It would also be interesting to bring interactions into consideration. Interactions, if not too strong, are not expected to fundamentally invalidate the results presented in this thesis, however, details of their effect will be quite challenging to address.

9.2 Topological classification using neural networks

An artificial neuron is a basic computational unit performing a simple non-linear operation on the input. Each neuron has a set of its variational parameters, the so-called weights and bias. The neurons can then be combined into layers and the neuron layers constitute a neural network. The elementary operations of neurons are then transferred across the layers and the network can be seen as a high-dimensional variational ansatz. In supervised learning we are given a dataset of input-label pairs and aim to train the network to output a correct label for each input. The training is done via minimizing a cost function representing the deviation of the network's output from the

given labels. The minimization is commonly done using gradient decent learning. A neural network is a powerful construction allowing one to efficiently find patterns in data, and in Papers D-E we employed neural nets for finding topological indices.

Specifically, we developed a novel neural-network-based protocol based on topological data augmentation and the interpretable 'learning by confusion' principle to retrieve topological indices with minimal external supervision. The main idea lies in an observation that it is not necessary to know anything about the construction of topological indices in order to create datasets of topologically equivalent states; such data can be obtained by continuously deforming individual seed states. We then aim to retrieve a quantity distinguishing between the created datasets of topologically equivalent states. In order to do so we apply the 'learning by confusion' concept where the datasets are pretended to be separable, i.e. labeled using 'dummy' labels, and a neural network is trained to pick up the corresponding feature. The successful training of the network will point to distinguishability of the datasets, but in general we will not be able to retrieve the learned quantity from the network. In order to be able to retrieve the quantity we suggest to use a special interpretable network layout representing a particular class of functions covering a wide range of topological indices. With this in hand we can then interpret the quantity found by the neural net.

Our protocol is explicitly exemplified on several symmetry classes of systems in Papers D-E, where the corresponding topological indices were numerically identified. However, the considered examples of topological band insulators are admittedly very simple and carry only illustrative significance. It will be more interesting to apply our protocol to more complex systems, with highly nontrivial or even unknown analytic expressions for the topological indices. We have in mind multiple directions to be explored, including Floquet topological systems, non-Hermitian topological phases, 2d and 3d crystalline topological matter, adding spatial disorder and interactions into the picture.

A Wilson loop

The Berry phase can be generalized to degenerate cases using the notion of a Wilson loop [59]. Let $\mathscr{H}_n(\boldsymbol{r})$ be a degeneracy subspace spanned by orthonormal eigenstates $|n,a;\boldsymbol{r}\rangle$ associated with the energy level ε_n, where a indices the basis states within the subspace. The eigenstates $|n,a;\boldsymbol{r}\rangle$ are defined up to a unitary transformation $|n,a;\boldsymbol{r}\rangle' = \sum_b U^{(n)}_{a,b}(\boldsymbol{r})|n,b;\boldsymbol{r}\rangle$ for any unitary matrix $U^{(n)}(\boldsymbol{r})$. Matrices do not commute in general and therefore the transformation of degenerate states give us a non-Abelian generalization of an ordinary (a.k.a. 'Abelian') gauge transformation.

Let us take some specific set of eigenstates $|n,a;\boldsymbol{r}\rangle$ and assume every state to be single-valued. The non-Abelian Berry connection $\boldsymbol{A}^n(\boldsymbol{r})$ then takes the form of a matrix with elements

$$\boldsymbol{A}^n_{a,b}(\boldsymbol{r}) = i\langle n,a;\boldsymbol{r}|\nabla_{\boldsymbol{r}}|n,b;\boldsymbol{r}\rangle \tag{A.1}$$

It follows that the non-Abelian Berry connection $\boldsymbol{A}^n(\boldsymbol{r})$ transforms under the non-Abelian gauge transformation $|n,a;\boldsymbol{r}\rangle' = \sum_b U^{(n)}_{a,b}(\boldsymbol{r})|n,b;\boldsymbol{r}\rangle$ as

$$\boldsymbol{A}'^n(\boldsymbol{r}) = [U^{(n)}(\boldsymbol{r})]^{-1}\boldsymbol{A}^n(\boldsymbol{r})U^{(n)}(\boldsymbol{r}) + [U^{(n)}(\boldsymbol{r})]^{-1}\nabla_{\boldsymbol{r}}U^{(n)}(\boldsymbol{r}). \tag{A.2}$$

In analogy to the conventional Berry phase we intend to construct a gauge invariant quantity by integrating the non-Abelian Berry connection over a closed curve $p \subset \mathscr{P}$. The non-Abelian generalization of the Berry phase, denoted by the matrix $U^{(n)}_p$, is formally defined by the expression

$$U^{(n)}_p = \mathscr{P}\exp\left[-i\oint_p \boldsymbol{A}^n(\boldsymbol{r})\cdot d\boldsymbol{r}\right], \tag{A.3}$$

where $\mathscr{P}$ denotes a path ordering operator. It is needed here precisely because the matrices $\boldsymbol{A}^n(\boldsymbol{r})$ do not necessary commute for different values of $\boldsymbol{r}$.

We are interested in constructing a gauge invariant quantity and therefore it is essential to investigate how the non-Abelian Berry phase $U^{(n)}_p$ transforms under non-Abelian gauge transformations. This calculation is often skipped in the literature but we find it important and interesting to briefly outline it. To begin with, we discretize the closed curve p using infinitesimal ordered

intervals $(\boldsymbol{r}_i, \boldsymbol{r}_{i+1}) \subset p$. The non-Abelian Berry phase $U_p^{(n)}$, Eq. (A.3), then reduces to

$$U_p^{(n)} = \prod_i \exp(L(\boldsymbol{r}_i, \boldsymbol{r}_{i+1})); \quad [L(\boldsymbol{r}_i, \boldsymbol{r}_{i+1})]_{ab} = \langle n, a; \boldsymbol{r}_i | (|n, b; \boldsymbol{r}_{i+1}\rangle - |n, b; \boldsymbol{r}_i\rangle). \tag{A.4}$$

Since the basis states $|n, a; \boldsymbol{r}\rangle$ are assumed to be smooth and single-valued, i.e. they have well-defined derivatives, the matrices $L(\boldsymbol{r}_i, \boldsymbol{r}_{i+1})$ are infinitesimal as long as the corresponding intervals are infinitesimal. The exponents $\exp(L(\boldsymbol{r}_i, \boldsymbol{r}_i + 1))$ can therefore be expanded to first order in $L(\boldsymbol{r}_i, \boldsymbol{r}_{i+1})$ and all higher-order terms can be neglected. It follows that $U_p^{(n)} = \prod_i (I + L(\boldsymbol{r}_i, \boldsymbol{r}_{i+1}))$. Thus,

$$[U_p^{(n)}]_{a,b} = \langle n, a; \boldsymbol{r}_0 | \left(\prod_i \sum_c |n, c; \boldsymbol{r}_i\rangle\langle n, c; \boldsymbol{r}_i| \right) |n, b; \boldsymbol{r}_0\rangle, \tag{A.5}$$

where $\boldsymbol{r}_0$ denotes the starting point of the integration over p. In this form it is easy to see that the non-Abelian Berry phase $U_p^{(n)}$ is invariant under any gauge transformation $|n, a; \boldsymbol{r}_i\rangle' = \sum_b U_{a,b}^{(n)}(\boldsymbol{r}_i)|n, b; \boldsymbol{r}_i\rangle$ at any point $\boldsymbol{r}_i \neq \boldsymbol{r}_0$; this is simply due to the cyclic property of the trace. However, the transformation at $\boldsymbol{r}_0$ may in general have some nontrivial impact. The non-Abelian Berry phase $U^{(n)}(\boldsymbol{r})$ is therefore not gauge invariant and transforms under any gauge transformation as $U'^{(n)}(\boldsymbol{r}) = [U^{(n)}(\boldsymbol{r}_0)]^{-1} U_p^{(n)} U^{(n)}(\boldsymbol{r}_0)$. On the other hand, the quantity $W_p^{(n)} = \mathrm{Tr}\, U_p^{(n)}$, dubbed the *Wilson loop*, is gauge invariant and can be used to characterize continuous transformations of the degenerate bands. An element of the non-Abelian Berry curvature tensor $\Omega_{\mu,\nu}^n(\boldsymbol{r})$ is defined as the non-Abelian Berry phase acquired after completing an infinitesimal parallelogram in the (μ, ν)-plane of the manifold $\mathscr{P}$ and can be expressed as:

$$\Omega_{\mu,\nu}^n(\boldsymbol{r}) = \Big[\partial_\mu A_\nu^n(\boldsymbol{r}) - \partial_\nu A_\mu^n(\boldsymbol{r}) + [A_\nu(\boldsymbol{r}), A_\mu(\boldsymbol{r})] \Big]. \tag{A.6}$$

Note that in the non-Abelian case, the Berry phase cannot be simply expressed as the surface integral over the curvature $\Omega_{\mu,\nu}^n(\boldsymbol{r})$ due to the inapplicability of the Stoke's theorem.

Bibliography

[1] L. D. Landau. *On the Theory of Phase Transitions*. Zh. Eksp. Teor. Fiz. **7**: 19-32 (1937).

[2] L. D. Landau, E. M. Lifshitz. *Statistical Physics Part 1, 3rd Ed. (Course of Theoretical Physics Vol. 5)*, Pergamon Press, Oxford (1994).

[3] X.-G. Wen, *Zoo of quantum-topological phases of matter*, Rev. Mod. Phys. **89**, 41004 (2017).

[4] K. v. Klitzing, G. Dorda, M. Pepper, *New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance*, Phys. Rev. Lett. **45**, 494 (1980).

[5] D. C. Tsui, H. L. Stormer, A. C. Gossard, *Two-Dimensional Magnetotransport in the Extreme Quantum Limit*, Phys. Rev. Lett. **48**, 1559 (1982).

[6] D. J. Thouless, M. Kohmoto, M. P. Nightingale, M. den Nijs, *Quantized Hall Conductance in a Two-Dimensional Periodic Potential*, Phys. Rev. Lett. **49**, 405 (1982).

[7] R. B. Laughlin, *Quantized Hall conductivity in two dimensions*, Phys. Rev. B **23**, 5632 (1981).

[8] B. I. Halperin, *Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential*, Phys. Rev. B **25**, 2185 (1982).

[9] J. E. Avron, R. Seiler, B. Simon, *Homotopy and Quantization in Condensed Matter Physics*, Phys. Rev. Lett. **51**, 51 (1983).

[10] R. B. Laughlin, *Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations*, Phys. Rev. Lett. **51**, 1395 (1983).

[11] X.-G. Wen, *Symmetry protected topological phases in non-interacting fermion systems*, Phys. Rev. B **85**, 085103 (2012).

[12] A. Mesaros, Y. Ran, *A classification of symmetry enriched topological phases with exactly solvable models*, Phys. Rev. B **87**, 155115 (2013).

[13] A. Kitaev, *Periodic table for topological insulators and superconductors*, AIP Conf. Proc. 1134, 22, (2009).

[14] D. S. Freed, G.W. Moore, *Twisted equivariant matter*. Ann. Henri Poincaré **14**, 1927-2023 (2013).

[15] G. C. Thiang, *Topological phases: isomorphism, homotopy and K-theory*, Int. J. Geom. Methods Mod. Phys. **12**, 1550098 (2015).

[16] G. C. Thiang, *On the K-theoretic classification of topological phases of matter*, Ann. Henri Poincaré **17(4)**, 757-794 (2016).

[17] A. Kapustin, R. Thorngren, A. Turzillo, Z. Wang. *Fermionic symmetry protected topological phases and cobordisms*. J. High Energ. Phys. **2015**, 1-21 (2015).

[18] D. Freed, M. Hopkins, *Reflection positivity and invertible topological phases*, arXiv:1604.06527 (2016).

[19] J. Kaidi, J. Parra-Martinez, Y. Tachikawa, *Topological superconductors on superstring worldsheets*, arXiv:1911.11780 (2019).

[20] B. Mehlig, *Artificial Neural Networks*, arXiv:1901.05639 (2019).

[21] E. P. Van Nieuwenburg, Y.-H. Liu, S. D. Huber, *Learning phase transitions by confusion*, Nature Physics, **13**, 435 (2017).

[22] P. Zhang, H. Shen, H. Zhai, *Machine Learning Topological Invariants with Neural Networks*, Phys. Rev. Lett. **120**, 066401 (2018).

[23] N. Sun, J. Yi, P. Zhang, H. Shen, H. Zhai, *Deep learning topological invariants of band insulators*, Phys. Rev. B, **98**, 085402 (2018).

[24] D. Carvalho, N. A. García-Martínez, J. L. Lado, J. Fernández-Rossier, *Real-space mapping of topological invariants using artificial neural networks*, Phys. Rev. B **97**, 115453 (2018).

[25] M. D. Caio, M. Caccin, P. Baireuther, T. Hyart, M. Fruchart, *Machine learning assisted measurement of local topological invariants*, arXiv:1901.03346 (2019).

[26] T. Mano, T. Ohtsuki, *Application of Convolutional Neural Network to Quantum Percolation in Topological Insulators*, J. Phys. Soc. Jpn. **88**, 123704 (2019),

[27] B. Wu, K. Ding, C. T. Chan, Y. Chen, *Machine prediction of topological transitions in photonic crystals*, arXiv:1907.07996 (2019).

[28] J. F. Rodriguez-Nieva, M. S. Scheurer, *Identifying topological order through unsupervised machine learning*, Nature Physics **15**, 790 (2019).

[29] Y. Zhang, R. G. Melko, and E.-A. Kim, *Machine learning Z2 quantum spin liquids with quasiparticle statistics*, Phys. Rev. B **96**, 245119 (2017).

[30] Y. Zhang and E.-A. Kim, *Quantum loop topography for machine learning*, Phys. Rev. Lett. **118**, 216401 (2017).

[31] Y. Zhang, P. Ginsparg, E.-A. Kim, *Interpreting machine learning of topological quantum phase transitions*, Phys. Rev. Research **2**, 023283 (2020).

[32] M. S. Scheurer, R.-J. Slager, *Unsupervised machine learning and band topology*, Phys. Rev. Lett. **124**, 226401 (2020).

[33] Y. Long, J. Ren, H. Chen, *Unsupervised manifold clustering of topological phononics*, Phys. Rev. Lett. **124**, 185501 (2020).

[34] Y. Che, C. Gneiting, T. Liu, F. Nori, *Topological quantum phase transitions retrieved from manifold learning*, arXiv:2002.02363 (2020).

[35] Y. Ming, C. Lin, S.D. Bartlett, and W.-W. Zhang, *Quantum topology identification with deep neural networks and quantum walks*, Npj Comput. Mater. **5**, 88 (2019).

[36] C.-K. Chiu, J. C. Y. Teo, A. P. Schnyder, S. Ryu, *Classification of topological quantum matter with symmetries*, Rev. Mod. Phys. **88**, 035005 (2016).

[37] Z. Gong, Y. Ashida, K. Kawabata, K. Takasan, S. Higashikawa, M. Ueda, *Topological Phases of Non-Hermitian Systems*, Phys. Rev. X **8**, 031079 (2018).

[38] K. Kawabata, K. Shiozaki, M. Ueda, and M. Sato, *Symmetry and Topology in Non-Hermitian Physics*, Phys. Rev. X *9*, 041015 (2019).

[39] S. Pai, A. Prem, *Topological states on fractal lattices*, Phys. Rev. B *100*, 155135 (2019).

[40] Z. Yang, E. Lustig, Y. Lumer, M. Segev, *Photonic Floquet topological insulators in a fractal lattice*, Light Sci Appl *9*, 128 (2020).

[41] R. Verresen, R. Thorngren, N. G. Jones, F. Pollmann, *Gapless topological phases and symmetry-enriched quantum criticality*, Phys. Rev. Lett. *120*, 057001 (2018).

[42] J.-W. Rhim, J. H. Bardarson, R.-J. Slager, *Unified bulk-boundary correspondence for band insulators*, Phys. Rev. B **97**, 115143 (2018).

[43] T. Fukui, K. Shiozaki, T. Fujiwara, S. Fujimoto, *Bulk-Edge Correspondence for Chern Topological Phases: A Viewpoint from a Generalized Index Theorem*, J. Phys. Soc. Jpn. **81**, 114602 (2012).

[44] R. Mong, V. Shivamoggi, *Edge states and the bulk-boundary correspondence in Dirac Hamiltonians*, Phys. Rev. B **83**, 125109 (2011).

[45] T. Kitagawa, E. Berg, M. Rudner, E. Demler, *Topological characterization of periodically-driven quantum systems*, Phys. Rev. B **82**, 235114 (2010).

[46] N. H. Lindner, G. Refael, V. Galitski, *Floquet Topological Insulator in Semiconductor Quantum Wells*, Nature Physics **7**, 490–495 (2011).

[47] M. S. Rudner, N. H. Lindner, *Floquet topological insulators: from band structure engineering to novel non-equilibrium quantum phenomena*, arXiv:1909.02008 (2019).

[48] M. S. Rudner, N. H. Lindner, E. Berg, M. Levin, *Anomalous edge states and the bulk-edge correspondence for periodically-driven two dimensional systems*, Phys. Rev. X **3**, 031005 (2013).

[49] A. Bansil, H. Lin, T. Das, *Colloquium: Topological band theory*, Rev. Mod. Phys. **88**, 021004 (2016).

[50] J. K. Asbóth, L. Oroszlány, and A. Pályi, *A Short Course on Topological Insulators* (Springer-Verlag, Berlin Heidelberg, 2016).

[51] W. P. Su, J. R. Schrieffer, and A. J. Heeger, *Solitons in Polyacetylene*, Phys. Rev. Lett. **42**, 1698 (1979).

[52] H. Takayama, Y. R. Lin-Liu, and K. Maki, *Continuum model for solitons in polyacetylene*. Phys. Rev. B **21**, 2388 (1980).

[53] W. P. Su, J. R. Schrieffer, and A. J. Heeger, *Soliton excitations in polyacetylene*, Phys. Rev. B **22**, 2099 (1980).

[54] A. J. Heeger, S. Kivelson, J. R. Schrieffer, and W. P. Su, *Solitons in conducting polymers*, Rev. Mod. Phys. **60**, 781 (1988).

[55] M. V. Berry. *Quantal phase factors accompanying adiabatic changes*, Proc. Roy. Soc. London, (A392):45-57 (1984).

[56] I. Mondragon-Shem, T. L. Hughes, J. Song, E. Prodan, *Topological Criticality in the Chiral-Symmetric AIII Class at Strong Disorder*, Phys. Rev. Lett. **113**, 046802 (2014).

[57] M. Nakagawa, R.-J. Slager, S. Higashikawa, T. Oka, *Wannier representation of Floquet topological states*, Phys. Rev. B **101**, 075108 (2020).

[58] Y. Aharonov and D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory*, Phys. Rev. **115**, 485 (1959).

[59] T.D. Stanescu, *Introduction to Topological Quantum Matter Quantum Computation*. CRC Press, Taylor Francis Group (2017).

[60] M. Nahahara: *Geometry, Topology and Physics*, 2ed., IOP Publishing Ltd (2003).

[61] J. C. Budich, Björn Trauzette, *From the adiabatic theorem of quantum mechanics to topological states of matter*, physica status solidi (RRL) **7**, 109 (2013).

[62] O. Balabanov, *Time-periodic topological band insulators: Robustness of the edge states*, Licentiate Thesis, Department of Physics, University of Gothenburg (2019).

[63] R. S. K. Mong, J. H. Bardarson, J. E. Moore, *Quantum transport and two-parameter scaling at the surface of a weak topological insulator*, Phys. Rev. Lett. **108**, 076804 (2012).

[64] Z. Ringel, Y. E. Kraus, A. Stern, *The strong side of weak topological insulators*, Phys. Rev. B **86**, 045102 (2012).

[65] E. P. Wigner. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Friedrich Vieweg und Sohn, Braunschweig, Germany, pp. 251-254 (1931).

[66] Y. Ando, L. Fu, *Topological Crystalline Insulators and Topological Superconductors: From Concepts to Materials*, Annual Review of Condensed Matter Physics **6**, 361-381 (2015).

[67] B.C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction, Graduate Texts in Mathematics*, Springer **222** (2015).

[68] A. Altland and M. R. Zirnbauer, *Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures*, Phys. Rev. B **55**, 1142 (1997).

[69] E. Cartan, *Sur une classe remarquable d'espaces de Riemann, I*, Bulletin de la Société Mathématique de France, **54**, 214-216 (1926).

[70] E. Cartan, *Sur une classe remarquable d'espaces de Riemann, II*, Bulletin de la Société Mathématique de France, **55**, 114-134 (1927).

[71] C. L. Kane, E. J. Mele, *Z_2 Topological Order and the Quantum Spin Hall Effect*, Phys. Rev. Lett. **95**, 146802 (2005).

[72] G. Floquet, *Sur les équations différentielles linéaires à coefficients périodiques*, Annales de l'École Normale Supérieure **12**: 47–88 (1883).

[73] J. H. Shirley, *Solution of the Schrödinger Equation with a Hamiltonian Periodic in Time*, Phys. Rev. **138**, B979 (1965).

[74] H. Sambe, *Steady States and Quasienergies of a Quantum-Mechanical System in an Oscillating Field Phys*, Rev. A **7**, 2203 (1973).

[75] D. E. Liu, A. Levchenko, H. U. Baranger, *Floquet Majorana Fermions for Topological Qubits in Superconducting Devices and Cold-Atom Systems*, Phys. Rev. Lett. **111**, 047002 (2013).

[76] R. Roy, F. Harper, *Periodic Table for Floquet Topological Insulators*, Phys. Rev. B **96**, 155118 (2017).

[77] R. Roy, F. Harper, *Floquet topological phases with symmetry in all dimensions*, Phys. Rev. B **95**, 195128 (2017).

[78] M. Bukov, L. D'Alessio, A. Polkovnikov, *Universal High-Frequency Behavior of Periodically Driven Systems: from Dynamical Stabilization to Floquet Engineering*, Advances in Physics, **64-2**, 139-226 (2015).

[79] N. Goldman and J. Dalibard. *Periodically Driven Quantum Systems: Effective Hamiltonians and Engineered Gauge Fields*, Phys. Rev. X, **4**, 031027 (2014).

[80] T. Mori, *Floquet resonant states and validity of the Floquet-Magnus expansion in the periodically driven Friedrichs models*, Phys. Rev. A **91**, 020101 (2015).

[81] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Comm. Pure and Appl. Math., VII, 649–673 (1954).

[82] S. Blanes, F. Casas, J.A. Oteo, J. Ros, *The Magnus expansion and some of its applications*, Physics Reports **470**, 151-238 (2009).

[83] E. S. Mananga, T. Charpentier, *Introduction of the Floquet-Magnus expansion in solid-state nuclear magnetic resonance spectroscopy*, Chem. Phys. **135**, 044109 (2011).

[84] V. Dal Lago, M. Atala, and L. E. F. Foa Torres, *Floquet topological transitions in a driven one-dimensional topological insulator*, Phys. Rev. A **92**, 023624 (2015).

[85] L. Jiang, T. Kitagawa, J. Alicea, A. R. Akhmerov, D. Pekker, G. Refael, J. I. Cirac, E. Demler, M. D. Lukin, and P. Zoller, *Majorana Fermions in*

Equilibrium and in Driven Cold-Atom Quantum Wires, Phys. Rev. Lett. **106**, 220402 (2011).

[86] M. Thakurathi, A. A. Patel, D. Sen, and A. Dutta, *Floquet generation of Majorana end modes and topological invariants*, Phys. Rev. B **88**, 155133 (2013).

[87] J. K. Asbóth, B. Tarasinski, and P. Delplace, *Chiral symmetry and bulk-boundary correspondence in periodically driven onedimensional systems*, Phys. Rev. B **90**, 125143 (2014).

[88] F. Nathan, M. S. Rudner, *Topological singularities and the general classification of Floquet-Bloch systems*, New J. Phys. **17**, 125014 (2015).

[89] D. Carpentier, P. Delplace, M. Fruchart, and K. Gawedzki, *Topological Index for Periodically Driven Time-Reversal Invariant 2D Systems*, Phys. Rev. Lett. **114**, 106806 (2015).

[90] C. Jörg, F, Letscher, M, Fleischhauer, G. von Freymann, *Dynamic defects in photonic Floquet topological insulators*, New J. Phys. **19**, 083003 (2017).

[91] Z. Fedorova (Cherpakova), C. Jörg, , C. Dauer, F, Letscher, M, Fleischhauer, S. Eggert, S. Linden, G. von Freymann, *Limits of topological protection under local periodic driving*, Light Sci Appl **8**, 63 (2019).

[92] T. Ozawa, H. M. Price, A. Amo, N. Goldman, M. Hafezi, L. Lu, M. C. Rechtsman, D. Schuster, J. Simon, O. Zilberberg, and I. Carusotto, *Limits of topological protection under local periodic driving*, Rev. Mod. Phys. **91**, 015006, (2019).

[93] R. Drost, T. Ojanen, A. Harju, P. Liljeroth, *Topological states in engineered atomic lattices*, Nat. Phys. **13** 668 (2017).

[94] I. Belopolski, S-Y Xu, N. Koirala, C. Liu, G. Bian, V. N. Strocov, G. Chang, M. Neupane, N. Alidoust, D. Sanchez *et al.*, *A novel artificial condensed matter lattice and a new platform for one-dimensional topological phases*, Sci. Adv. **3** 1501692 (2017).

[95] M. Atala, M. Aidelsburger, J. T. Barreiro, D. Abanin, T. Kitagawa, E. Demler, I. Bloch, *Direct Measurement of the Zak phase in Topological Bloch Bands*, Nat. Phys. **9** 795 (2013).

[96] W. Zheng, H. Zhai, *Floquet Topological States in Shaking Optical Lattices*, Phys. Rev. A **89** 061603(R) (2014)

[97] M. Leder, C. Grossert, L. Sitta, M. Genske, A. Rosch, M. Weitz, *Real-space imaging of a topologically protected edge state with ultracold atoms in an amplitude-chirped optical lattice*, Nat. Commun. **7** 13112 (2016).

[98] S. Haykin, *Neural Networks: a comprehensive foundation*, New Jersey: Prentice Hall (1999).

[99] J. Duchi, E. Hazan, Y. Singer. *Adaptive subgradient methods for online learning and stochastic optimization*, J. Mach. Learn. Res. **12**, 2121-2159 (2011).

[100] T. Tieleman, G. Hinton, *Lecture 6.5 - RMSProp, COURSERA: Neural Networks for Machine Learnin*, Technical report (2012).

[101] M. D. Zeiler: *Adadelta: An adaptive learning rate method*, arXiv:1212.5701 (2012).

[102] H. B. McMahan. *Follow-the-regularized-leader and mirror descent: Equivalence theorems and L1 regularization*. In Proc. AISTATS (2011).

[103] D. P. Kingma, J. L. Ba, *Adam : A method for stochastic optimization*, arXiv:1412.6980v9 (2014).

[104] N. Srivastav, G. Hinton, A. Krishevsky, I. Sutskever, R. Salkhutdinov, *Dropout: A simple way to prevent neural networks from overfitting*. J. Mach. Learn. Res. **15(56)**, 1929-1958 (2014).

[105] S. J. Hanson, L. Y. Pratt, *Comparing biases for minimal network construction with backpropagation*, Advances in Neural Information Processing (1), Denver, pp. 177-185 (1989).

[106] A. Krogh, J. A. Hertz, *A simple weight decay can improve generalization*, Advances in Neural Information Processing (4), pp. 951-957 (1992).

[107] K. He, X. Zhang, S. Ren, J. Sun. *Deep residual learning for image recognition*, arxiv:1512.03385 (2015).

[108] S. Ioffe, C. Szegedy, *Batch normalization: Accelerating deep network training by reducing internal covariate shift*, arxiv:1502.03167 (2015).

[109] A. Krizhevsky, I. Sutskever, G. E. Hinton, *Imagenet classification with deep convolutional neural networks*, Advances in Neural Information Processing (25) (2012).

[110] K. Fukushima, *Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position*, Biol. Cybern. **36**, 193-202 (1980).