

Wechselwirkungen zwischen Kantenzuständen
auf dem
topologisch kristallinen Isolator
 $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$



Dissertation

zur Erlangung des naturwissenschaftlichen Doktorgrades
der Julius-Maximilians-Universität Würzburg

vorgelegt von
Johannes Jung
aus Speyer

Würzburg 2022



Eingereicht am
bei der Fakultät für Physik und Astronomie

1. Gutachter: Prof. Dr. Matthias Bode
 2. Gutachter: Prof. Dr. Hartmut Buhmann
 3. Gutachter:
- der Dissertation

Vorsitzende(r):

1. Prüfer: Prof. Dr. Matthias Bode
 2. Prüfer: Prof. Dr. Hartmut Buhmann
 3. Prüfer:
- im Promotionskolloquium

Tag des Promotionskolloquiums:.....
Doktorurkunde ausgehändigt am:.....

Zusammenfassung

Einerseits besteht die einfachste Möglichkeit zum Ladungs- und Informationstransport zwischen zwei Punkten in deren direkter Verbindung durch eindimensionale Kanäle. Andererseits besitzen topologische Materialien exotische und äußerst vorteilhafte Eigenschaften, weshalb es nahe liegt, dass schon bald neue Anwendungen aus ihnen realisiert werden. Wenn diese beiden Entwicklungen zusammenkommen, dann ist ein grundlegendes Verständnis von Quanteninterferenz oder Hybridisierungseffekten in eindimensionalen, topologischen Kanälen von fundamentaler Wichtigkeit. Deshalb werden in der vorliegenden Arbeit Wechselwirkungen von eindimensionalen, topologisch geschützten Kantenzuständen, die an ungeradzahligen Stufenkanten auf der (001)-Oberfläche von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ auftreten, untersucht.

Aufgrund der lateralen Lokalisierung auf wenige Nanometer um eine Stufenkante herum und der Notwendigkeit zwischen gerad- und ungeradzahligen Stufenkantenhöhen zu unterscheiden, bieten sich die Rastertunnelmikroskopie und -spektroskopie als Methoden an. Die neu entdeckten Kopplungs- bzw. Wechselwirkungseffekte zwischen benachbarten Kantenzuständen treten auf, sobald der Stufe zu Stufe Abstand einen kritischen Wert von $d_{\text{kri}} \approx 25 \text{ nm}$ unterschreitet. Dieses Kriterium kann durch verschiedene räumliche Anordnungen von Stufenkanten erfüllt werden. Infolgedessen werden sich kreuzende, parallel verlaufende und zusammenlaufende Stufenkanten genauer untersucht.

Bei letzteren verändert sich entlang der Struktur kontinuierlich der Abstand und damit die Kopplungsstärke zwischen den beiden Randkanälen. Infolgedessen wurden drei Koppelungsregime identifiziert. (I) Ausgehend von einer schwachen Wechselwirkung zeigt der für die Kantenzustände charakteristische Peak im Spektrum zunächst eine Verbreiterung und Verminderung der Intensität. (II) Mit weiter zunehmender Wechselwirkung beginnt sich der Zustand in zwei Peaks aufzuspalten, sodass ab $d_{\text{kri}} \approx 15 \text{ nm}$ an beiden Stufenkanten durchgehen eine Doppelpeak zu beobachten ist. Mit weiter abnehmendem Abstand erreicht die Aufspaltung Werte von einigen 10 meV , während sich die Intensität weiter reduziert. (III) Sobald zwei Stufenkanten weniger als etwa 5 nm voneinander getrennt sind, konvergieren aufgrund der schwindenden Intensität und des sinkenden energetischen Abstands der beiden Peaks zu den van Hove Singularitäten die Spektren an den Stufenkanten gegen das Spektrum über einer Terrasse.

Die Aufspaltung verläuft in den Bereichen I und II asymmetrisch, d. h. ein Peak verbleibt ungefähr bei der Ausgangsenergie, während der andere mit zunehmender Kopplung immer weiter weg schiebt. Bezüglich der Asymmetrie kann kein Unterschied festgestellt werden, ob die zusammenlaufenden Stufenkanten eine Insel oder Fehlstelleninsel bilden oder ob die Stufenkanten sogar gänzlich parallel verlaufen. Es zeigt sich keine Präferenz, ob zunächst der niederenergetische oder der hochenergetische Peak schiebt. Erst im Regime starker Kopplung (III) kann beobachtet werden, dass beide Peaks die Ausgangsenergie deutlich verlassen.

Im Gegensatz dazu kann bei sich kreuzenden Stufen ein erheblicher Einfluss der Geometrie, in Form des eingeschlossenen Winkels, auf das Spektrum beobachtet werden. Unabhängig vom Winkel existiert am Kreuzungspunkt selbst kein Kantenzustand mehr. Die Zustände an den vier Stufen beginnen, abhängig vom Winkel, etwa 10-15 nm vor dem Kreuzungspunkt abzuklingen. Überraschenderweise zeigt sich dabei, dass im Fall rechtwinkliger Stufen gar keine Aufspaltung zu beobachten ist, während bei allen anderen Winkeln ein Doppelpeak festgestellt werden kann. Diese Entdeckung deutet auf Orthogonalität bezüglich einer Quantenzahl bei den beteiligten Kantenzustände hin. Neben einer nur theoretisch vorhergesagten Spinpolarisation kann dieser Effekt auch von dem orbitalem Charakter der beteiligten Dirac-Kegel verursacht sein.

Da der topologische Schutz in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ durch Kristallsymmetrien garantiert ist, wird als letzter intrinsischer Effekt der Einfluss von eindimensionalen Defekten auf die Kantenzustände untersucht. Berücksichtigt werden dabei ein nicht näher klassifizierbarer, oberflächennaher Defekt und Schraubversetzungen. In beiden Fällen kann ebenfalls eine Aufspaltung des Kantenzustands in einen Doppelpeak gezeigt werden.

Im zweiten Teil dieser Arbeit werden die Grundlagen für eine Wiederverwendung von (Pb,Sn)Se-Oberflächen bei zukünftige Experimenten mit (magnetischen) Adatomen geschaffen. Durch Kombination von Inoenzerstäubung und Tempern wird dabei nicht nur eine gereinigte Oberfläche erzeugt, sondern es kann auch das Fermi-niveau gezielt erhöht oder gesenkt werden. Dieser Effekt beruht auf eine Modifikation der Sn-Konzentration und der von ihr kontrollierten Anzahl an Defektelektronen.

Als letztes sind erste Messungen an Cu- und Fe-dotierte Proben gezeigt. Durch die Adatome tritt eine n-Dotierung auf, welche den Dirac-Punkt des Systems in Richtung des Fermi-niveaus verschiebt. Sobald er dieses erreicht hat kommt es zu Wechselwirkungsphänomenen an freistehenden Stufenkanten. Dies führt zu einer Doppelpeakstruktur mit einer feinen Aufspaltung von wenigen meV. Das Phänomen ist auf ein schmales Energiefenster beschränkt, bei dem die Lage des Dirac-Punkts nur etwa 5 meV (in beide Richtungen) von der des Fermi-niveaus abweichen darf.

Abstract

First, the simplest possibility of transporting charges and information between two points is given by their direct connection due to one-dimensional channels. Second, topological materials have exotic and extremely advantageous properties, which make them suitable for further applications. If these two come together, then a basic understanding of quantum interference or hybridization effects in one-dimensional, topological channels is of fundamental importance. Therefore, in the present work, interactions of one-dimensional, topologically protected edge states, hosted at odd-numbered step edges on the (001) surface of (Pb,Sn)Se, are investigated.

Due to the lateral localization to a few nanometers around a step edge and the need to differentiate between even and odd-numbered step heights, scanning tunneling microscopy and spectroscopy are the tools of choice. The newly discovered coupling or interaction effects between neighboring edge states appear as soon as their distance decreases below a critical value of $d_{\text{cri}} \approx 25$ nm. This criterion can be met by various spatial arrangements of step edges. As a result, crossing, parallel and converging step edges are examined more closely.

With the latter, the distance and thus the coupling strength between the two edge channels changes continuously along the structure. As a result, three coupling regimes were identified. (I) Starting from a weak interaction, the peak in the spectrum that is characteristic of the edge states initially shows a broadening and reduction in intensity. (II) With increasing interaction, the state begins to split into two peaks, so that from $d_{\text{cri}} \approx 15$ nm a double peak can be observed at both step edges. As the distance continues to decrease, the splitting reaches values of a few 10 meV, while the intensity continues to drop. (III) As soon as two step edges are separated by less than about 5 nm, the spectra at the step edges converge against the spectrum over a terrace due to the decreasing intensity and the decreasing energetic distance of the two peaks to the van Hove singularities.

The split is asymmetrical in areas I and II, which means that one peak remains roughly at the original energy, while the other shifts further and further away with increasing coupling. With regard to the asymmetry, no difference can be determined whether the converging step edges form an island, a vacancy island or even run completely parallel. There is no preference as to whether the low energy or high energy peak shifts. Only in the regime of strong coupling (III) both peaks clearly leave the initial energy.

In contrast to this, a considerable influence of the geometry on the spectrum can be observed, with the included angle as parameter, for intersecting steps. Independent of the angle, there is no longer an edge state at the intersection itself. The states at the four edges start to decay, depending on the angle, about 10-15 nm before the point of intersection. Surprisingly, it turns out that in the case of right angled steps no splitting at all can be observed, while a double peak can be found for all other angles. This discovery indicates orthogonality with respect to a quantum number in the edge states involved. In addition to a theoretically predicted spin polarization, this effect can also be caused by the orbital character of the Dirac cones involved.

Since the topological protection in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ is guaranteed by crystal symmetries, the last intrinsic effect to be examined is the influence of one dimensional defects on the edge states. A near-surface defect, which cannot be classified in any more detail and a screw dislocation are taken into account. In both cases, a splitting of the edge state into a double peak can also be shown.

In the second part of this thesis the basis for reuse of surfaces in future experiments with (magnetic) adatoms is created. The combination of sputtering and annealing not only creates a cleaned surface, but in addition it tunes the Fermi level in a controllable way. This effect is based on a modification of the Sn concentration and the associated number of holes.

Finally, the first measurements on Cu and Fe-doped samples are shown. The adatoms cause n-doping, which shifts the Dirac point of the system in the direction of the Fermi level. As soon as he has achieved this, there is an interaction phenomenon at the freestanding step edges. This leads to a double peak structure with a fine split of a few meV. This phenomenon is limited to a narrow energy window in which the position of the Dirac point may only deviate by about 5 meV (in both directions) from that of the Fermi level.

Inhaltsverzeichnis

1	Einleitung	1
2	Topologie in der Festkörperphysik	3
2.1	Topologie	3
2.2	Das Su–Schrieffer–Heeger–Modell	5
2.3	Altland–Zirnbauer–Klassifizierung	9
2.4	Erweiterung der AZ–Klassifizierung	12
3	Rastertunnelmikroskopie	19
3.1	Quantenmechanischer Tunneleffekt	20
3.2	Bardeen–Modell	23
3.2.1	STM–Bilder im Tersoff–Hamann–Modell	25
3.2.2	STS–Spektren	26
4	Pb_{1–x}Sn_xSe	33
4.1	Kristall– & Bandstruktur	33
4.2	Topologische Phasenübergänge	37
4.3	Kantenzustände in Pb _{1–x} Sn _x Se	40
5	Kopplung von 1D Kantenzustände	53
5.1	Datenaufbereitung	53
5.2	Kantenzustände an geraden Stufen	55
5.3	Koppelungseffekte von Kantenzuständen an ungeraden Stufen	58
5.3.1	Zusammenlaufende Stufenkanten	58
5.3.2	Parallellaufende Stufenkanten	74
5.3.3	Kreuzende Stufenkanten	77
5.4	Kristalldefekte	82
6	Dotierung von Pb_{1–x}Sn_xSe mittels Adatome	91
6.1	<i>In situ</i> Reinigung von Pb _{1–x} Sn _x Se	92
6.2	3 <i>d</i> –Übergangsmetallatome auf Pb _{1–x} Sn _x Se	102

7 Zusammenfassung	105
Anhang	109
Abbildungsverzeichnis	110
Tabellenverzeichnis	110
Abkürzungsverzeichnis	111
Verzeichnis eigener Publikationen	112
Literatur	113

1 Einleitung

Symmetrien spielen eine entscheidende Rolle in der Physik: Unter anderem bei der Ableitung von Erhaltungsgrößen mittels des Noether-Theorems, der Entartung von Energieniveaus (z.B. bezüglich des Drehimpulses im Wasserstoffatom) oder bei quantenmechanischen Vielteilchenwellenfunktionen, wo die (Anti-) Symmetrie bezüglich der Vertauschung zweier Teilchen zu fundamental unterschiedlichen Besetzungsstatistiken führt. Bis zur Entdeckung des Quanten Hall Effekts (Quantum Hall Effect, QHE) 1980 konnten alle Phasen und Phasenübergänge durch Symmetrien bzw. Symmetriebrechungen beschrieben werden [1–3]. Retrospektiv war dies der Beginn eines Paradigmenwechsels, weil nun die Notwendigkeit bestand ein neues Konzept zur Klassifizierung von Phasen zu entwickeln, welches über eine bloße Betrachtung von Symmetrien hinausgeht.

Trotzdem dauerte es 25 Jahre bis das dem QHE zugrundeliegende Konzept der Topologie zur Vorhersage neuer Phasen führte. Der zweidimensionalen topologischen Isolator (Topological Insulator, TI)¹ wurde 2005 durch Kane und Mele [4] sowie durch Bernevig und Zhang [5] unabhängig voneinander vorhergesagt. Den Durchbruch in der modernen Festkörperphysik schafften TIs in den folgenden beiden Jahren, als zunächst von Bernevig, Hughes und Zhang [6] mit in CdTe eingebetteten HgTe–Quantentöpfen ein experimentell realisierbares Materialsysteme vorgeschlagen wurden (2006). An solchen Proben wurde dann im Rahmen von Transportmessungen an der Universität Würzburg dieser neue, exotische Zustand erstmals nachgewiesen [7] (2007). Seitdem werden TI intensiv erforscht und sind bis heute eines der aktivsten Forschungsfelder der modernen Festkörperphysik. Das zeigt sich u.a. an einer Vielzahl neu entdeckter Phasen, wie den 3D TIs [8–12], den topologischen Kondo-Isolatoren (Topological Kondo Insulator, TKI), den topologischen Dirac/Weyl–Halbmetallen (topological Dirac/Weyl Semimetals, WS) [13–17] oder den kristallinen topologischen Isolatoren (Topological Crystalline Insulator, TCI) [18–22]. Bedingt wird dieses immense Interesse durch zwei Tatsachen. Einerseits eröffnet die Anwendung topologischer Konzepte aus theoretischer Sicht völlig neue Herangehensweisen in der Festkörperphysik, weil Details über die Wechselwirkungen in einem Festkörper nicht bekannt sein müssen. Als Konsequenz dieser Konzepte ergibt sich, dass die elektronische Struktur an den Grenzflächen eines TIs nicht alleine von der exakten atomaren Anordnung abhängt, sondern z. T. durch Eigenschaften im Volumen des Materials bestimmt wird. Die Eigenschaften dieser Oberflächen-

¹2D TIs sind in der Literatur oft unter dem Name Quanten–Spin–Hall Zustand/Isolator anzutreffen.

zustände wiederum motivieren andererseits experimentelle Untersuchungen, weil sie von hoher Signifikanz für zukünftige Anwendung sind [7, 23, 24]. Hervorzuheben sind hierbei zum einen die Kopplung des Spins der Elektronen an deren Ausbreitungsrichtung (Wellenvektor) [23–27] und die Robustheit der Oberflächenzustände gegen Störungen verschiedenster Art.

Letztere Eigenschaft resultiert aus der Tatsache, dass die Oberflächenzustände eine Konsequenz aus einem räumlichen Phasenübergang (zwischen den elektronischen Strukturen im Volumen auf den beiden Seiten der Grenzfläche) sind. Sie können also nicht durch Störgrößen wie z.B. Unordnung entfernt werden, außer es kommt auf einer Seite zu einem Phasenübergang im Volumen selbst. Da die unterschiedlichen Phasen jedoch mit fundamentalen Symmetrien verknüpft sind, werden sie durch deren Anwesenheit geschützt. Dementsprechend sind auch die Oberflächen durch die selben Symmetrien geschützt und können nur durch eine Symmetriebrechung entfernt werden [23–25, 27, 28].

Bei TI kommt dieser Schutz durch die Zeit–Umkehr–Symmetrie (Time Reversal Symmetry, TRS) zustande. Zusammen mit der Ladungskonjugationssymmetrie konnten bereits 1997 zehn unterschiedliche Klassen aus Symmetriekombinationen identifiziert werden und für jede Klasse in Abhängigkeit der Dimension angegeben werden, ob topologisch nicht triviale Phasen vorliegen können oder nicht. Obwohl diese nach Altland und Zirnbauer benannte Klassifizierung in 50 % der Fälle nicht triviale Einträge liefert, wurde die Bedeutung topologischer Phasen lange Zeit nicht erkannt [23, 29, 30]. Allerdings beschränkten sich die Konzepte der Topologie nicht auf diese speziellen Symmetrien. So nehmen im Fall von TCIs diskrete Kristallsymmetrien die Rolle der TRS ein [18].

Da in dieser Arbeit ein TCI untersucht wurde, werden zum Ende des nächsten Kapitels die für $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ relevanten Kristallsymmetrien erstmals aufgegriffen. Zuerst werden in Kap. 2 jedoch wichtige Begriffe aus dem Gebiet der Topologie erläutert und einige Anwendungsmöglichkeiten in der Festkörperphysik aufgezeigt. Danach wird in Kap. 3 die Theorie der Messmethode eingeschoben, weil die Entdeckung der in dieser Arbeit untersuchten 1D Kantenzustände ebenfalls mittels eines Rastertunnelmikroskops (scanning tunneling microscope, STM) erfolgte. Neben der genauen Beschreibung der Kristallstruktur enthält Kap. 4 auch eine Zusammenfassung aller bisherigen Erkenntnisse zu der (Pb,Sn)(Se,Te)–Materialklasse. Kap. 4 fasst aber nicht nur den Forschungsstand zu Beginn meiner Promotion zusammen, sondern greift auch alle währenddessen entstanden Veröffentlichungen auf und bildet so den Übergang zu den Ergebnissen dieser Arbeit: Zunächst werden die intrinsischen Wechselwirkungen von 1D Kantenzuständen untereinander in Kap. 5 vorgestellt. In Kap. 6 werden dann die Grundlagen für extrinsische Störgrößen, in Form von Adatomen, gelegt und erste Zwischenergebnisse für Fe- und Cu-dotierte Proben vorgestellt. Zuletzt werden in Kap. 7 die wichtigsten Ergebnisse nochmals zusammengefasst.

2 Topologie in der Festkörperphysik

Die Einleitung wurde von dem Begriff des topologischen Isolators geprägt, ohne auf dessen Konzepte einzugehen. In diesem Kapitel sollen, ohne formelle Definitionen, einige grundlegende Begriffe (wie den der topologischen Invariante) und Ideen (wie die der stetigen Transformation) aus der Topologie kurz eingeführt werden. Daraufhin werden dann verschiedene Anknüpfungspunkte zwischen diesem abstraktem mathematischem Konzept und der Festkörperphysik aufgezeigt und anhand eines konkreten Beispiels, des Su–Schrieffer–Heeger–Modells (Su–Schrieffer–Heeger, SSH), vorgestellt. Der Schwerpunkt bildet die daran anschließende Klassifizierung topologischer Phasen gemäß fundamentaler Symmetrien. Diese sogenannten Altland–Zirnbauer–Klassen werden dann noch um bestimmte räumliche Symmetrien erweitert, welche für den in dieser Arbeit untersuchten topologisch kristallinen Isolator $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ besonders relevant sind.

2.1 Topologie

Die Topologie ist ein Teilgebiet der Mathematik, dass sich mit Mengen (von Mengen) und deren stetigen Transformationen in andere Mengen (von Mengen) befasst. Dabei liegt der Schwerpunkt auf solchen Eigenschaften, die durch die Transformation unverändert bleiben und als topologische Invarianten bezeichnet werden. Die Bedingung der Stetigkeit einer Transformation ist dabei an ein verallgemeinertes Konzept der Nähe geknüpft. Erfüllen zwei Elemente (Mengen) der Ausgangsmenge die Bedingung der Nähe, so müssen sie dies auch in der Zielmenge tun.

Ein Standardbeispiel für Topologie stellen die Oberflächen dreidimensionaler Objekte (als Elemente von Mengen) und deren Deformation ineinander (als Transformationen) dar. Anschaulich gesprochen können stetige Transformationen dann benachbarte Punkte einer Oberfläche nicht voneinander trennen, wie es z.B der Fall wäre, wenn ein Loch in die Fläche gemacht oder ein Ring aufgeschnitten wird. Beispiele für stetige Transformationen sind hingegen alle Arten von Verformungen wie Streckung oder Stauchung. Aus topologischer Sicht ist nun interessant, durch welche unveränderliche Eigenschaften die Oberflächen beschrieben werden können, welche Werte diese Invarianten annehmen und wie die Zuordnung bzw. Verteilung zwischen diesen Werten und den Oberflächen verschiedener Objekte aussieht. Durch den Satz von Gauß-Bonnet existiert ein Zusammenhang zwischen der Krümmung K und dem Geschlecht g einer Fläche, also der Anzahl

ihrer Löcher:

$$\int K dA = 2 \cdot \pi(2 - 2 \cdot g) \tag{2.1}$$

Die Integration über die lokale Eigenschaft der Oberflächenkrümmung ist also quantisiert in ganzen Zahlen und mit der globalen Eigenschaft des Geschlechts verknüpft. Die Anzahl der Löcher ist dabei die gesuchte Invariante, da unsere Bedingung für stetige Transformationen das Kreieren von Löchern implizit verbietet. Mit dieser Invariante können also Oberflächen anhand ihrer Anzahl an Löchern in verschiedene Mengen bzw. Klassen eingeteilt werden. So ist also eine Kugel zu einem Würfel topologisch äquivalent, nicht jedoch zu einem Torus.

Wo sind die Anknüpfungspunkte dieses Konzepts an die Festkörperphysik? Ein möglicher Zugang ist die Suche nach einem Analogon zu Gl. 2.1, bei dem ebenfalls die Geometrie in Form einer lokalen Eigenschaft mit der Topologie (in Form einer globale Eigenschaft) verknüpft wird. So lässt sich aus einer Gaußkrümmung im k -Raum die sogenannte Berry-Phase berechnen, deren Fluss wiederum tatsächlich mit einer topologischen Invariante, der Chernzahl verknüpft ist. Anstatt diese Zusammenhänge nun ausführlich mathematisch herzuleiten, soll hier die Frage betrachtet werden, wie diese zusätzliche Information aus der Topologie in die Wellenfunktion eingeht.

In der Quantenmechanik ist bekannt, dass nur das Betragsquadrat einer Wellenfunktion eine messbare Observable ist. Dementsprechend besitzt jede Wellenfunktion noch einen Phasenfaktor, der in der Regel schlicht vernachlässigt wird, da dessen Betragsquadrat gleich Eins ist. In diesem Faktor stecken jedoch sehr wohl physikalische Informationen, wie z.B. bei der Interferenz des Aharonov-Bohm-Effekts deutlich wird. Tatsächlich lässt sich eine enge Analogie zwischen der Berry-Phase und dem Aharonov-Bohm-Effekt finden, die verdeutlicht, dass die topologische Information der Berry-Phase über den Phasenfaktor in der Wellenfunktion berücksichtigt wird. Einen weiteren Zugang bietet die Frage, ob und wenn ja wo ein direkter Zusammenhang zwischen Topologie und der Wellenfunktionen selbst besteht. Allgemein sind Funktionen durch eine Abbildungsvorschrift zwischen zwei Mengen (Werte- und Definitionsmenge) definiert. Daher ist es naheliegend, dass die Konzepte der Topologie auch auf diese Mengen angewandt werden können. So ist z.B für einen 2D-Kristall der k -Raum periodisch, sodass dieser Wertebereich durch die Oberfläche eines 3D-Torus beschrieben wird. Es handelt sich also geometrisch um das selbe Objekt wie zuvor bei dem Donut, nun jedoch im Impuls- statt im Ortsraum. Diese Unterscheidung ist jedoch für die Konzepte der Topologie irrelevant. Interessanterweise ist auch die exakte Funktionsvorschrift unerheblich, was letztlich bedeutet, dass die ihr zugrundeliegenden Details der Wechselwirkungen des konkreten Systems nicht bekannt sein müssen. Hierin liegt die Mächtigkeit des Ansatzes, Topologie und Festkörperphysik miteinander zu kombinieren.

Wenn es also nicht nötig ist jede Wechselwirkung im Detail zu kennen, dann sollte es ebenfalls nicht nötig sein für jedes wechselwirkende System einen neuen Hamiltonoperator \mathcal{H} aufstellen und lösen zu müssen. Stattdessen sollten die Konzepte von stetigen Transformationen und topologischen Invarianzen auch auf Hamiltonoperatoren anwendbar sein. Dies soll im Folgenden kurz anhand des SSH-Modells gezeigt werden, wobei das Augenmerk zum Schluss auf die Bandstruktur respektive Dispersionsrelation gelegt wird.

2.2 Das Su–Schrieffer–Heeger–Modell

In diesem Kapitel werden die zuvor vorgestellten Konzepte auf ein einfaches Modell, das Su–Schrieffer–Heeger–Modell übertragen. Ziel ist es die wesentlichen Charakteristika eines topologischen Zustands herauszuarbeiten. Diese Eigenschaften sind allgemeingültig und beschreiben daher auch die Oberflächenzustände von (Pb,Sn)Se.

Das SSH-Modell besteht aus einer eindimensionalen Kette von Gitterplätzen, die alternierend zu zwei verschiedenen Untergittern A und B gehören². Während in A alle Plätze vollständig besetzt sind, sind sie in B vollständig unbesetzt. Jeweils zwei benachbarte Plätze bilden eine Zelle, sodass jede Zelle aus einem A - und einem B -Gitterplatz besteht. Da es gleich viele A - und B -Plätze gibt erfolgt die Paarbildung so, dass keine ungepaarten Plätze übrig bleiben [siehe Abb. 2.1(a)]. Der Hüpfparameter innerhalb einer Zelle (v) und der zwischen zwei Zellen (w) sind die Parameter dieses Modells. Die einzelnen Gitterplätze werden eindeutig durch einen Zustand $|m, i\rangle$ beschrieben, wobei die erste Quantenzahl dem Zellindex m und die zweite dem Untergitterindex $i \in \{A, B\}$ entspricht. Der Einteilchen-Hamiltonoperator dieses Systems kann mit Hilfe des äußeren Produkts dieser Zustände folgendermaßen beschrieben werden [31]:

$$\mathcal{H} = v \sum_{m=1}^N (|m, B\rangle \langle m, A| + h.c.) + w \sum_{m=1}^{N-1} (|m+1, A\rangle \langle m, B| + h.c.). \quad (2.2)$$

Der ausgeschriebene Teil des ersten Terms beschreibt den Übergang vom Untergitter B in A innerhalb der m -ten Zelle, während h.c. für den adjungierten Ausdruck steht. Für $|m, B\rangle \langle m, A|$ lautet dieser $|m, A\rangle \langle m, B|$ und beschreibt somit den umgekehrten Übergang in der selben Zelle. Der zweite Term beschreibt den Übergang aus der nächsten Zelle ($m+1$) in die vorherige, sowohl erneut für den Fall, dass dieser von A nach B oder umgekehrt stattfindet. In ausgeschriebener Form führt Gl. 2.2 zu einer $2N \times 2N$ -Matrix mit Nullen auf der Hauptdiagonalen und alternierenden v und w auf den beiden angrenzenden Nebendiagonalen [31].

²Die Ableitung basiert auf Ref. [31] und ist zusätzlich auch als Open Access auf arXiv.org unter dem Index 1509.02295 zu finden.

Zur Lösung des dazugehörigen Eigenwertproblems bieten sich periodische Randbedingungen an, um die Eigenschaften des Volumens zu untersuchen. Anschaulich entspricht das dem Schließen der Kette zu einem Ring, sodass zusätzlich noch ein Übergang zwischen dem A -Gitter der $m = 1$ und dem B -Gitter der Zelle $m = N$ erlaubt wird. Diese zusätzlichen w tauchen in der Matrix ganz unten links und ganz oben rechts auf. Die Lösung des Eigenwertproblems führt dann zu einer Dispersionsrelation³ der Form [31]

$$E(k) = |v + e^{-ik} \cdot w| = \pm \sqrt{v^2 + w^2 + 2 \cdot v \cdot w \cdot \cos(k)}. \quad (2.3)$$

Diese Dispersion entspricht in diesem Modell der Bandstruktur unseres Festkörpers und ist in Abb. 2.1(b) für fünf verschiedene Parameterkonstellationen gezeigt. Im Ersten Fall ist kein Übergang zwischen Zellen möglich ($w = 0$), sodass der \cos -Term aus Gl. 2.3 verschwindet und die Zustände nicht mehr dispergieren. Dasselbe Bild ergibt sich auch, wenn der Austausch nur zwischen den Zellen möglich ist ($v = 0$). Bei allen endlichen Wertepaaren (v, w) für die $v \neq w$ gilt, wird eine endliche Dispersion und überall eine Bandlücke, vor allem auch am Rand der Brillouin-Zone (Brillouin Zone, BZ), beobachtet. Lediglich für den Spezialfall $v = w$ ist die Bandlücke bei $k = \pm\pi$ geschlossen. Es handelt sich nun also nicht länger um einen Isolator.

Das gleiche Ergebnis ergibt sich durch eine topologische Betrachtung, aber diese liefert, wie nun gezeigt werden soll, sogar noch Erkenntnisse darüber hinaus. Während die exakte Rechnung der Bandstruktur keinen signifikanten Unterschied zwischen den Fällen $v = 0$ und $w = 0$ ergibt, so wird nun gezeigt, dass die beiden zu verschiedenen topologischen Phasen gehören. Als Konsequenz daraus folgt, dass keine stetige Transformation diese beiden Phasen ineinander überführen kann ohne währenddessen einen Phasenübergang zu durchschreiten. In diesem Beispiel passiert dies offensichtlich an dem Punkt, an dem aus einem Isolator ein Leiter wird.

Für die Analyse der Topologie wird als erstes ausgenutzt, dass jeder Hamiltonoperator als Produkt aus Pauli-Spin-Matrizen $\vec{\sigma}$ und einem impulsabhängigen Vektor $d(\vec{k})$ dargestellt werden kann [31]:

$$\mathcal{H} = d(\vec{k}) \cdot \vec{\sigma} + d_0(\vec{k}) \cdot \mathbb{1}. \quad (2.4)$$

Durch Einsetzen der expliziten Form der Pauli-Matrizen und mit $d_0(\vec{k}) = 0$ ($\mathbb{1}$ ist die Einheitsmatrix) ergibt sich dann für die Komponenten des $d(\vec{k})$ -Vektors im SSH-Modell [31]:

$$d_x = v + w \cdot \sin(k), \quad d_y = w \cdot \sin(k) \quad \text{und} \quad d_z = 0. \quad (2.5)$$

³Da im SSH-Modell der Gitterabstand $a = 1$ gesetzt wird, ist k dimensionslos und die Grenzen der Brillouin-Zone liegen bei $\pm \frac{\pi}{a} = \pm\pi$.

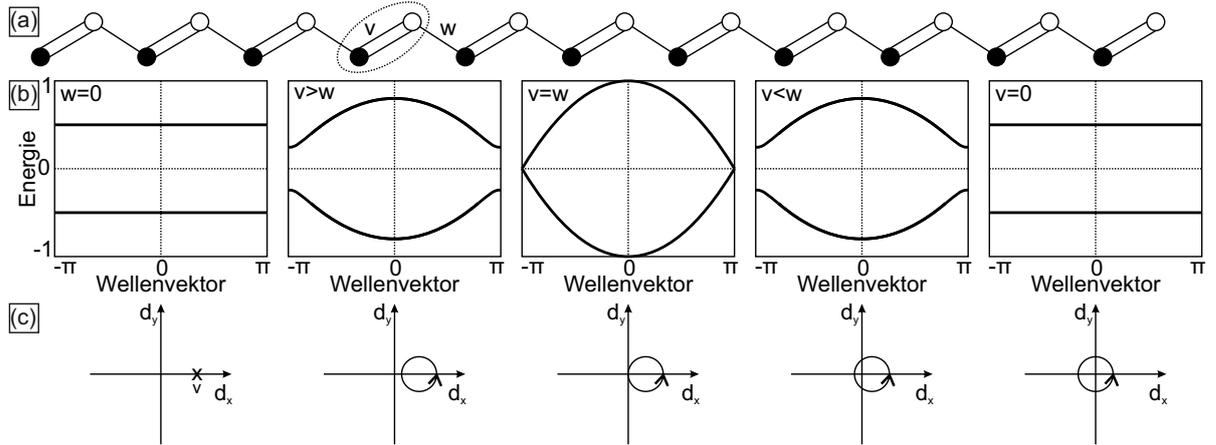


Abbildung 2.1 Su–Schrieffer–Heeger Modell: (a) Skizze des SSH–Modells. Die 1D–Kette besteht aus zwei gleichgroßen Untergittern A und B , von denen eines vollständig be- und das andere unbesetzt ist. Je ein Untergitterplatz zusammen bilden eine Zelle, wie durch die gestrichelte Linie angedeutet. Die Hüpfparameter innerhalb (v) und zwischen den Zellen (w) sind die Modellparameter. (b) Dispersion für verschiedene Verhältnisse der Parameter gemäß Gl. 2.3. (c) Trajektorien des Endpunkts vom $d(\vec{k})$ –Vektor bei Variation von $k = 0 \rightarrow 2 \cdot \pi$ gemäß Gl. 2.5, unter Verwendung der gleichen Parameter wie in (b). Die Anzahl der Windungen um den Ursprung ist eine topologische Invariante ν . Im mittleren Fall ist ν nicht definiert, da ein Phasenübergang zwischen Isolator und Leiter vorliegt. Links der Mitte gilt $\nu = 0$ und rechts $\nu = 1$, sodass die identischen Dispersionen in (b) trotzdem zu topologisch unterschiedlichen Klassen gehören. Adaptiert nach [31].

Wie jeder Vektor wird auch der d –Vektor durch zwei Eigenschaften beschrieben: Länge und Richtung. Die Länge gibt die Energie der dazugehörigen Eigenzustände an und enthält daher keine zusätzlichen Informationen. Diese steckt in der Richtung des d –Vektors, bzw. genauer gesagt in der Trajektorie, welche der Endpunkt als Funktion von $k = 0 \rightarrow 2 \cdot \pi$ durchläuft⁴. Fasst man k als Winkel auf, so parametrisiert Gl. 2.5 einen Kreis mit Radius w und Mittelpunkt $(v, 0)$ in Polarkoordinaten. Im Fall des SSH–Modells schreitet der d –Vektor also eine Kreisbahn in der $d_x d_y$ –Ebene ab, was in in Abb. 2.1(c) für die gleichen Parameterkombinationen (v, w) wie in (b) dargestellt ist. So fällt für den ersten Fall ($w = 0$) die Abhängigkeit von $\sin(k)$ weg, da der Radius des Kreises verschwindet. Der d –Vektor zeigt somit für alle k –Werte auf den Punkt $(v, 0)$. Im zweiten Fall ($v > w$) ergibt sich ein Kreis mit endlichem Radius. Allerdings liegen alle Werte rechts der Null, weil $\sin(k) \geq -1$ ist und $d_{x, \min} = v - w$ gilt. Für $v > w$ ist somit der Wert von $d_{x, \min}$ positiv.

Die Lage des Kreises relativ zum Ursprung ist entscheidend, denn analog zu dem Geschlecht einer Fläche kann für geschlossene Pfade eine topologische Invariante gefunden und eine Klassifizierung vorgenommen werden. In diesem Fall handelt es sich dabei um die Windungszahl ν , welche angibt wie oft der Pfad den Ursprung umschließt. Während in den ersten beiden Fällen in (c) der Ursprung gar nicht umschlossen ist, wird er es in den letzten beiden Fällen jeweils

⁴Die Brillouin–Zone läuft zwar von $-\pi$ nach $+\pi$, aber da es sich bei der Trajektorie um einen geschlossenen Pfad und bei $E(k)$ um eine periodische Funktion handelt, sind Anfangs- und Endwert irrelevant solange eine komplette Periode durchlaufen wird.

einmal. Im Fall $v = w$ schneidet der Kreis den Ursprung, sodass ν nicht definiert ist. Aus der topologischen Betrachtung geht also hervor, dass die beiden Grenzfälle $v = 0$ und $w = 0$ zu unterschiedlichen topologischen Klassen gehören und deren Hamiltonoperatoren deshalb nicht durch eine stetige Transformation ineinander übergeführt werden können, obwohl beide die identische Bandstrukturen generieren.

Aus Gl. 2.5 geht hervor, dass der d -Vektor auf die $d_x d_y$ -Ebene beschränkt ist. Diese Randbedingung wird durch $d_z = 0$ beschrieben und ergibt sich als direkte Konsequenz aus der chiralen⁵ Symmetrie des Hamiltonoperators. Da der d -Vektor in der $d_x d_y$ -Ebene liegen muss, kann ν in diesem Modell nur die Werte Null und Eins annehmen. Um die Windungszahl zu ändern gibt es nur zwei Möglichkeiten, zum einen den Ring von der $d_x d_y$ -Ebene anzuheben und ihn neben dem Ursprung wieder abzusetzen, oder ihn über den Ursprung zu ziehen [31]. Die erste Option bricht die gerade eingeführte, fundamentale Symmetrie des Hamiltonoperators, sodass eine solche Transformation nicht mehr als stetig bezeichnet werden kann. Die zweite Option tut dies nicht, jedoch ist sie gleichbedeutend mit der Tatsache, dass ein Punkt in der Brillouin-Zone existiert, wo der d -Vektor und somit auch der Hamiltonoperator verschwinden. Die physikalische Konsequenz daraus ist, dass sich an diesem Punkt die Bandlücke schließt und damit ein Phasenübergang vom Isolator zum Leiter eintritt. Folglich ist an eine stetige Transformation ebenfalls die Bedingung zu stellen, dass die Bandlücke niemals geschlossen wird. Zur Vollständigkeit sei an dieser Stelle angemerkt, dass als dritte Bedingung für stetige Transformationen eine kontinuierliche Variation der Parameter des Hamiltonoperators erfolgen muss [31].

Allgemein muss wegen der periodischen Randbedingungen der Pfad nach einem Durchlauf des Wellenvektors von $k = 0 \rightarrow 2 \cdot \pi$ lediglich geschlossen, nicht jedoch kreisförmig, sein. Insbesondere für höherdimensionale Problemstellungen schreitet der d -Vektor dann z.B auch Flächen ab. Auf jeden Fall nimmt dabei der Ursprung die gleiche Sonderrolle ein, denn falls dieser Teil von $d(\vec{k})$ ist, dann existiert ein Punkt in der Brillouin-Zone wo die Bandlücke verschwindet.

Bisher wurde nur die elektronische Struktur im Volumen der SSH-Kette betrachtet. Anstatt nun die Kette gedanklich wieder zu trennen und die Zustände an den Enden explizit zu berechnen, soll hier direkt der Fall einer Grenzfläche zwischen zwei Ketten mit unterschiedlicher topologischer Invariante diskutiert werden. Da die Zustände an der Grenzfläche nicht einfach verschwinden können, muss also analog zur Diskussion von Abb. 2.1(c) ein Weg gefunden werden, wie der Ring den Ursprung passieren kann [31].

Nicht triviale, also von Null verschiedene topologische Invarianten, sind nur für bestimmte Kombinationen von Symmetrien und Dimensionalität eines physikalischen Problems möglich. Zu der daraus resultierenden Altland-Zirnbauer-Klassifizierung (Altland Zirnbauer, AZ) folgt später mehr. Mit Ausnahme einer Klasse ist in jedem Fall die topologische Invariante mit mindestens

⁵Die chirale Symmetrie wird formal zu Beginn des Unterkapitels 2.3 definiert.

einer Symmetrie verknüpft. In all diesen Fällen existiert neben dem Schließen der Bandlücke auch das Brechen einer Symmetrie als Möglichkeit, um die elektronischen Zustände an der Grenzfläche ineinander überzuführen. Bei kristallinen Symmetrien, die Schwerpunkt dieser Arbeit sind, kann dies je nach Oberflächenorientierung des Kristalls der Fall sein oder auch nicht.

Im SSH-Modell bleibt die zugrundeliegende chirale Symmetrie jedoch erhalten. Somit ist die einzige Option ein Schließen der Bandlücke an der Grenzfläche, was bedeutet dass das Valenzband (Leitungsband) auf der topologisch trivialen Seite in das Leitungsband (Valenzband) auf der nicht trivialen Seite übergehen muss. Durch diese Zwangsbedingung muss es an der Grenzfläche also aufgrund der Topologie im Volumen zur Ausbildung von Oberflächenzuständen in der Volumenbandlücke kommen, die auch noch so lange robust sind, wie die zugrundeliegende Symmetrie erhalten bleibt.

Dies ist die Quintessenz eines topologisch geschützten Zustands: (i) Er sitzt an einer Grenzfläche und hat damit mindestens eine um Eins niedrigere Dimension als die ihn schützende Bandstruktur, (ii) seine Existenz leitet sich alleine aus der elektronischen Struktur im Volumen bzw. in einer höheren Dimension ab und (iii) er wird durch eine Symmetrie geschützt.

2.3 Altland–Zirnbauer–Klassifizierung

Nachdem im vorherigen Abschnitt die fundamentale Bedeutung von Symmetrien für die topologische Klassifizierung herausgearbeitet wurde, soll in diesem Kapitel ein grundlegendes Klassifizierungsschema aufgezeigt werden. Dieses beschränkt sich zunächst auf wenige Symmetrien und wird dann in Kap. 2.4 erweitert, sodass auch die für (Pb,Sn)Se relevanten Kristallsymmetrie berücksichtigt werden.

Im vorherigen Kapitel wurde angedeutet, dass es Ausnahmen gibt, bei denen gar keine fundamentale Symmetrie gegeben ist. Eine solche Ausnahme stellt der Quanten-Hall-Zustand dar. In diesem wird durch das externe Magnetfeld die Zeitumkehrsymmetrie gebrochen. Charakteristisch ist dann der Ladungstransport entlang der Kanten, der selbst bei einem hohen Maß an Unordnung im Festkörper erhalten bleibt. Diese Robustheit gegen Unordnung ist eine Schlüsseleigenschaft, die sich letztlich aus dem topologischen Schutz an der Grenzfläche ergibt.

Durch Unordnung wird dem Hamiltonoperator ein Term hinzugefügt, der die Translationssymmetrie bricht. Da auch im Fall von Unordnung der topologische Schutz besteht, muss folglich der topologische Effekt auf den nicht translationsinvarianten Hamiltonoperator zurückgeführt werden. Da jeder unitäre Operator jedoch mit dem Hamiltonoperator kommutiert, kann dieser in Blockform gebracht und die Wirkung der dazugehörigen Symmetrie in jedem Block separat betrachtet werden. In diesem Sinne handelt es sich dann um eine lokale Symmetrie, welche nicht zur Beschreibung der globalen Topologie geeignet ist. Durch die Beschränkung auf nicht unitäre

Operatoren bleiben genau zwei Symmetrien und deren Kombination übrig: TRS \mathcal{T} , Teilchen–Loch–Symmetrie \mathcal{C} (Particle Hole Symmetry, PHS) und als Kombination deren Produkt, die chirale Symmetrie⁶ $\mathcal{S} = \mathcal{T} \cdot \mathcal{C}$.

Chiralität bedeutet, dass eine Invarianz unter einer Paritätstransformation vorliegt oder nicht⁷. Deshalb kann \mathcal{S} nur zwei Werte annehmen: Invarianz ja (entspricht dem Wert 1 in Tab. 2.1) oder nein (Wert 0). Die Paritätstransformation kann unabhängig davon durchgeführt werden, ob andere Symmetrien existieren oder nicht, aber solange TRS, PHS oder beide vorliegen, bestimmt das Produkt $\mathcal{T} \cdot \mathcal{C}$ eindeutig das Ergebnis: (i) Sobald entweder die TRS oder die PHS gebrochen ist, kann das System nicht chiral sein. (ii) Sobald weder die TRS noch die PHS gebrochen ist, muss das System chiral sein. Für den Fall, dass weder PHS noch TRS vorliegen, kann trotzdem eine Paritätstransformation durchgeführt werden, allerdings lässt sich dann das Ergebnis nicht mehr mit Hilfe dieser beiden Symmetrien vorhersagen. Daher kann es in diesem Fall sowohl chirale als auch nicht chirale Hamiltonoperatoren geben.

Für TRS und PHS existiert hingegen nicht nur die Unterscheidung zwischen gegeben oder abwesend. Falls eine dieser Symmetrien vorliegt, dann kann der dazugehörige Operator entweder bei zweifacher Anwendung zur Identität werden oder das Vorzeichen ändern, d.h. $\mathcal{T}^2 = \pm 1 = \mathcal{C}^2$. Es gibt also für TRS und PHS jeweils drei Fälle (-1, 0, +1), wobei für den Fall $\mathcal{C} = \mathcal{T} = 0$ nicht eindeutig definiert ist, ob das System chiral ist oder nicht. Folglich führen diese Symmetrien also zu $(3 \cdot 3 - 1) + 2 = 10$ verschiedenen Kombinationen, welche mit ihren dazugehörigen Werten für \mathcal{T} , \mathcal{C} und \mathcal{S} in Tab. 2.1 gezeigt sind.

Die Tabelle ist als Altland–Zirnbauer–Klassifizierung oder auch als Periodensystem topologischer Invarianten bekannt [23, 29]. Besagte Autoren entdeckten, dass es für jeden Einteilchen–Hamiltonoperator eine 1:1–Zuordnung zu einem sogenannten symmetrischen Raum⁸ gibt. Bereits 1926 zeigt der Mathematiker Élie Cartan, dass es genau zehn solcher Räume gibt, sodass die AZ–Klassifizierung vollständig bezüglich dieser drei Symmetrien ist [30].

Die Unterteilung in komplexe und reelle Klassen kommt daher, dass in Abwesenheit von TRS und PHS keine Bedingungen an \mathcal{H} gestellt wird und damit keine mathematische Beziehung zu dessen hermitesch transponierte \mathcal{H}^* angegeben werden kann [32]. Die Anordnung der reellen Klassen in dieser Reihenfolge ergibt sich nicht nur aus dem sich wiederholenden Muster innerhalb

⁶Die PHS ist oft unter dem Namen Ladungskonjugationssymmetrie oder kurz \mathcal{C} –Symmetrie zu finden. Die chirale Symmetrie wird gelegentlich als Untergittersymmetrie (englisch sublattice symmetry) bezeichnet. Im Fall des SSH–Modells liegen zwar tatsächlich Untergitter vor, jedoch existiert diese Symmetrie auch in Systemen ohne solche. Daher wird hier durchgehend die Definition ($\mathcal{S} = \mathcal{T} \cdot \mathcal{C}$) und Bezeichnung chiral verwendet.

⁷Eine Paritätstransformation beschreibt die Transformation zweier gleichartiger Teilchen mit unterschiedlicher Chiralität ineinander. Für masselose Teilchen ist die Chiralität mit der Helizität identisch und gibt an, ob das Teilchen rechts- oder linkshändig ist. Definiert wird diese Eigenschaft durch das Vorzeichen der Projektion des Spins- auf den Impulsvektors.

⁸Ein symmetrischer Raum ist eine spezielle Riemannsche Mannigfaltigkeit, bei der an jedem Element eine Spiegelung durchgeführt werden kann, d.h. jeder Punkt ist ein Inversionszentrum.

Tabelle 2.1 Periodensystem topologischer Invarianten: Klassifizierung von topologischen Invarianten gemäß verschiedenen Symmetrieklassen und der Dimension d . Für jede Klasse ist der Cartan-Name und die dazugehörige Kombination aus Zeit-Umkehr- (\mathcal{T}), Teilchen-Loch- (\mathcal{C}) und chiraler Symmetrie (\mathcal{S}) angegeben. Hamiltonoperatoren mit ausschließlich trivialer Bandstruktur sind mit einer 0 markiert; nicht triviale können durch eine ganze Zahl \mathbb{Z} oder einer Restklasse Modulo 2 \mathbb{Z}_2 beschrieben werden. Reelle Klassen zeichnen sich durch mindestens eine Symmetrie (\mathcal{T} oder \mathcal{C}) und eine achtfache Periode bezüglich d aus, komplexe Klassen durch das Fehlen einer mathematischen Verknüpfung von \mathcal{H} mit \mathcal{H}^* und eine zweifache Periode in d .

Cartan-Name	Symmetrie			$d = \dots$							
	\mathcal{T}	\mathcal{C}	\mathcal{S}	0	1	2	3	4	5	6	7
Komplexe Klassen:											
A	0	0	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}	0
AIII	0	0	1	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}
Reelle Klassen:											
AI	+1	0	0	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2
BDI	+1	+1	1	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2
D	0	+1	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0
DIII	-1	+1	1	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$
AII	-1	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0
CII	-1	-1	1	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0
C	0	-1	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0
CI	+1	-1	1	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}

der Zeilen (welches in den Spalten durch geschickte Anordnung wiederholt werden kann), sondern es folgt aus einem Reduktionsverfahren. Dieses stellt einen Zusammenhang zwischen einer topologischen Eigenschaft in der Klasse X mit Dimension D und einer anderen Invarianten in der Dimension $D - 1$ und einer neuen Klasse Y her. Als Konsequenz daraus kommt es zu einer diagonalen Anordnung, bei der die topologische Invariante um eine Dimension steigt, wenn von einer Symmetriekombination bzw. Zeile zur nächsten übergegangen wird.

Eine solche Periodizität kann jedoch nur erreicht werden, weil in jede Dimension exakt fünf Invarianten auftreten (inklusive der komplexen Klassen). Aber nicht nur ihre Gesamtzahl stimmt überein, sondern auch die Verteilung auf verschiedene Arten von Zahlenräumen. Je zweimal wird eine Invariante durch eine Restwertklasse Modulo 2 (\mathbb{Z}_2) beschrieben, während in den anderen drei Fällen eine bzw. zwei ganze Zahlen \mathbb{Z} die Invariante charakterisieren. Auffallend ist dabei, dass die beiden Restwertklassen stets zusammen mit einer \mathbb{Z} -Invarianten als Triplet auf treten und zwar nicht nur innerhalb jeder Dimension, sondern auch innerhalb jeder Klasse. Dieser Umstand legt den Verdacht nahe, dass ein enger Zusammenhang zwischen diesen Invarianten besteht. Tatsächlich lassen sich die Restwertklassen durch ein Dimensionsreduktionsverfahren aus der \mathbb{Z} -Invarianten direkt ableiten.

Zum Abschluss der AZ-Klassifizierung ist noch anzumerken, warum acht Dimensionen berücksichtigt werden. Dies liegt daran, dass die zugrundeliegenden Symmetriegruppen dem Bott-

Periodizität–Theorem genügen. Aus diesem leitet sich dann nicht nur das zuvor festgestellte Muster ab, sondern auch eine achtfache Periodizität bezüglich der Dimension. Dementsprechend wiederholen sich die Invarianten IV innerhalb einer Klasse für höhere Dimensionen gemäß $d_n \rightarrow d_{n+8}$, z.B. also $IV(d = 0) \rightarrow IV(d = 8)$.

2.4 Erweiterung der AZ–Klassifizierung

Im vorherigen Abschnitt wurden nur wenige Symmetrien mit globaler Wirkung auf den Einteilchen–Hamiltonoperator berücksichtigt. Nun sollen zwei Erweiterungen vorgestellt werden, bei denen zusätzliche Symmetrien berücksichtigt werden. Als nächstes werden räumliche Symmetrien zweiter Ordnung, wie z.B Spiegelungen oder zweifache Rotationen, betrachtet. Im darauffolgendem Schritt werden dann sogenannte nicht symorphe Raumgruppen zweiter Ordnung einbezogen. Die dazugehörigen Symmetrien sind die Partner von den räumlichen Symmetrien zweiter Ordnung, da sie sich lediglich durch eine zusätzliche Translation um einen Bruchteil eines primitiven Gittervektors von diesen unterscheiden. Ein Beispiel für eine nicht symorphe Symmetrie (NonSymmorphic Symmetry, NSS) ist die Spiegelung an einer Ebene und eine anschließende Verschiebung z. B. senkrecht zum dazugehörigen Normalenvektor. Diese Kombination wird als Glide–Symmetrie bezeichnet und am Ende dieses Abschnitts unter Berücksichtigung der erweiterten Klassifizierung nochmals diskutiert.

Symorphe Symmetrien zweiter Ordnung

Die hier gezeigten Ergebnisse leiten sich aus einem Spezialfall der K–Theorie, der sogenannten twisted equivariant K theory, ab. Allgemein befasst sich die K–Theorie mit Vektorbündeln und topologischen Räumen. Um eine weitere Symmetrie in die Klassifizierung aufzunehmen, muss zunächst spezifiziert werden, welchen Eigenschaften sie bzw. ihr Operator besitzt. Diese können entweder durch eine unitär Matrix U oder eine antiunitäre Matrix A repräsentiert werden. Außerdem kann die K–Theorie auch angewandt werden, wenn die Antisymmetrien \bar{U} oder \bar{A} betrachtet werden, wobei sich diese durch ein zusätzliches, negatives Vorzeichen vor dem Hamiltonoperator auszeichnen:

$$\begin{aligned}
 U\mathcal{H}(\vec{r}, \vec{k})U^{-1} &= \mathcal{H}(-\vec{r}_{||}, \vec{r}_{\perp}, -\vec{k}_{||}, \vec{k}_{\perp}) & \bar{U}\mathcal{H}(\vec{r}, \vec{k})\bar{U}^{-1} &= -\mathcal{H}(-\vec{r}_{||}, \vec{r}_{\perp}, -\vec{k}_{||}, \vec{k}_{\perp}) \\
 A\mathcal{H}(\vec{r}, \vec{k})A^{-1} &= \mathcal{H}(-\vec{r}_{||}, \vec{r}_{\perp}, -\vec{k}_{||}, \vec{k}_{\perp}) & \bar{A}\mathcal{H}(\vec{r}, \vec{k})\bar{A}^{-1} &= -\mathcal{H}(-\vec{r}_{||}, \vec{r}_{\perp}, -\vec{k}_{||}, \vec{k}_{\perp}).
 \end{aligned} \tag{2.6}$$

Da ggf. nur ein Teil der Koordinaten durch die Symmetrie zweiter Ordnung geändert wird, ist dieser Teil des Ortsvektors \vec{r} durch den Index $||$ gekennzeichnet. Dementsprechend werden

auch die selben Komponenten des Impulsvektors \vec{k} durch die Symmetrie beeinflusst. Als Konsequenz ergibt sich, dass nicht mehr zwingend die Dimension D des Gesamtsystems alleine entscheidend ist, sondern die Anzahl der invertierten Koordinaten ebenfalls zur Bildung einer effektiven Dimension berücksichtigt werden müssen. Während hier Ort und Impuls klassisch behandelt wurden, d.h. unter der Annahme dass ihre Operatoren miteinander kommutieren, so muss für jede Extrasymmetrie deren Vertauschungsrelationen mit \mathcal{T} und \mathcal{C} berücksichtigt werden. Ferner muss das Ergebnis einer zweifachen Hintereinanderausführung betrachtet werden, weil $U^2 = A^2 = \bar{U}^2 = \bar{A}^2 = \pm 1$ entweder zur Identität führt oder ein Phasenfaktor auftreten kann.

Insgesamt gibt es also ohne Vertauschungsrelationen bereits acht Fälle zu berücksichtigen, was zu 16 bzw. 32 Kombinationen führt, falls die Kommutatoren von einer oder beider Symmetrien (\mathcal{T} und \mathcal{C}) betrachtet werden müssen (je nach Klasse gibt es z.T. nur eine von beiden). In jeder Symmetrieklasse lassen sich diese Kombinationen gleichmäßig auf 4 Unterklassen aufteilen, sodass nur jeweils 4 bzw. 8 Operatoren pro Unterklasse möglich sind. Im Folgenden wird sich auf Operatoren der Form U und die reellen Klassen beschränkt, nicht jedoch auf eine spezielle Unterklasse. Dabei wird als Notation $U_{b,c}^a$ mit $a, b, c \in \{+, -\}$ verwendet, wobei der obere Index das Vorzeichen von U^2 angibt und die unteren beiden den Kommutator (+) bzw. Antikommutator (-) für \mathcal{T} (b) und \mathcal{C} (c) angeben. Sofern eine Symmetrieklasse nur entweder \mathcal{T} oder \mathcal{C} besitzt, wird der andere Index weggelassen⁹.

Als konkrete Beispiele sollen hier die Spiegelsymmetrie (Tab. 2.2) und ihr nicht symorpher Partner, die Gleitsymmetrie (Tab. 2.3), betrachtet werden. Bei beiden kommt es zur Invertierung von einer Raumkoordinate, was in Ref. [33] und [34] dem Fall $\delta_{||} = 1$ entspricht. Analoge Tabellen finden sich dort auch für $\delta \in \{0, 2, 3\}$ und für die hier nicht gezeigten komplexen Klassen und antiunitären Operatoren.

Auch unter Anwesenheit einer Spiegelsymmetrie zeigt sich in Tab. 2.2 das bekannte Muster aus dem $\mathbb{Z}-\mathbb{Z}_2-\mathbb{Z}_2$ -Tripletts gefolgt von einer Null und einer $2\mathbb{Z}$ -Invarianten innerhalb jeder Klasse, sofern alle Kommutatoren das gleiche oder das entgegengesetzte (Vor-) Zeichen wie U^2 besitzen (1. und 3. Block). Allerdings ist der Anfang des Triplets verschoben. Im ersten (dritten) Block werden u. a. Symmetrien der Form U_+^+ (U_-^+) behandelt und in jeder Klasse sind die Invarianten um eine Dimension gestiegen (gesunken), sodass alle Einträge um eins nach rechts (links) gerutscht sind. So beginnt das Tripletts in der Klasse AI nun nicht mehr bei $d = 0$, sondern die \mathbb{Z} -Invariante sitzt nun bei $d = 1$ ($d = 7$). Außerdem bleibt die Bott-Periodizität erhalten.

Neu sind hingegen nicht triviale Phasen die durch zwei Invarianten beschrieben werden und in der Form $\mathbb{Z} \oplus \mathbb{Z}$ bzw. $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ in Tab. 2.2 auftauchen. Diese sind nur in solchen Klassen zu

⁹ Beispiel: Die Symmetrieklassen C und D besitzt gemäß Tab. 2.1 nur die PHS, sodass der untere Index von U in diesem Fall den (Anti-) Kommutator mit \mathcal{C} angibt.

Tabelle 2.2 Symorphe Erweiterung der AZ-Klassifizierung: Klassifizierung von topologischen Invarianten gemäß verschiedenen Symmetrieklassen in Anwesenheit einer unitären, räumlichen Symmetrie zweiter Ordnung mit einer invertierten Koordinate, z.B einer Spiegelsymmetrie. Der obere Index von U gibt das Vorzeichen von U^2 an, die unteren ob U mit \mathcal{T} und/oder \mathcal{C} kommutiert (+) oder antikommutiert (-). Die Null symbolisiert triviale Phasen, während Invarianten bei \mathbb{Z} durch ganze Zahlen und bei \mathbb{Z}_2 durch die Restwertklasse Modulo 2 beschrieben werden. Das \oplus -Zeichen deutet an, dass zwei entsprechende Invarianten die Phase charakterisieren. Analoge Tabellen für mehr oder weniger invertierte Koordinaten, anunitäre Operatoren und die komplexen Klassen sind in Ref. [33].

Klasse	Symmetrie	$d = \dots$							
		0	1	2	3	4	5	6	7
AI		\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2
BDI		\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0
D		0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$
DIII	$U_{+,+}^+, U_{-,-}^-$	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0
AII	$U_{-,-}^-, U_{+,+}^+$	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0
CII		0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0
C		0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}
CI		\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2
BDI	$U_{+,-}^+, U_{-,+}^-$	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}$	0	0	0	$2\mathbb{Z} \oplus 2\mathbb{Z}$	0	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$
DIII	$U_{-,+}^+, U_{+,-}^-$	0	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}$	0	0	0	$2\mathbb{Z} \oplus 2\mathbb{Z}$
CII	$U_{+,-}^+, U_{-,+}^-$	0	$2\mathbb{Z} \oplus 2\mathbb{Z}$	0	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}$	0	0
CI	$U_{-,+}^+, U_{+,-}^-$	0	0	0	$2\mathbb{Z} \oplus 2\mathbb{Z}$	0	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}$
AI		0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}
BDI		\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2
D		\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2
DIII	$U_{-,-}^-, U_{+,+}^+$	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$	0
AII	$U_{+,+}^+, U_{-,-}^-$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0	$2\mathbb{Z}$
CII		$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0	0
C		0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0	0
CI		0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}	0
BDI, CII	$U_{-,+}^+, U_{+,-}^-$	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}
DIII, CI	$U_{+,+}^+, U_{-,-}^-$	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}	0	\mathbb{Z}

beobachten wo sowohl \mathcal{T} - als auch \mathcal{C} -Symmetrie vorliegt. Für die Klassen BDI und CII sind solche Phasen nur beobachtbar, wenn U^2 und \mathcal{T} das selbe Vorzeichen haben, während in den Klassen DIII und CI die selbe Bedingung für U^2 und \mathcal{C} gilt. Auch in diesem zweiten Block gibt es die gleiche Abfolge von Invarianten (\mathbb{Z} - \mathbb{Z}_2 - \mathbb{Z}_2 -0- $2\mathbb{Z}$ -0-0-0) innerhalb einer Klasse, welche nun jedoch in gedoppelter Form vorliegen. Außerdem ist die Diagonalform, bei der eine Invariante um eine Dimension steigt wenn in die nächste Zeile gegangen wird, ebenfalls gegeben. Es ist dabei aber zu berücksichtigen, dass zwischen BDI und DIII noch die Klasse D liegt, welche jedoch nur PHS hat und deshalb im zweiten Block fehlt. Dementsprechend muss nicht nur eine, sondern jeweils zwei Dimensionen höher gegangen werden, wenn in diesem Block die nächste Zeile betrachtet wird.

Eine weitere Neuheit ist aus Tab. 2.2 alleine nicht ersichtlich, weil hier nur für eine invertierte Koordinate die Invarianten angegeben wurden¹⁰. Wenn auch die weiteren Tabellen betrachtet werden, so ist auch eine Periodizität in der Anzahl der invertierten Koordinaten feststellbar. So gibt es in der Klasse AI für $d = 0$ und eine Symmetrie der Form U_+^+ folgende Reihenfolge von Invarianten ($d = 3 \rightarrow 0$): $0-\mathbb{Z}-\mathbb{Z}_2-\mathbb{Z} \oplus \mathbb{Z}$. Die selbe Abfolge findet sich auch bei einer Symmetrie der Form U_-^+ (ebenfalls $d = 0$), allerdings um zwei Dimensionen verschoben ($d = 3 \rightarrow 0$): $\mathbb{Z}_2-\mathbb{Z} \oplus \mathbb{Z}-0-\mathbb{Z}$. Dieser Zusammenhang gilt auch in anderen Klassen, wobei dort die Abfolge der Invarianten sich ändert. So ist z.B. in der BDI-Klasse die Null und \mathbb{Z} vertauscht (gleiches U_b^a) oder es tauchen gänzlich andere Reihen auf wie für U_+^+ in CII ($d = 3 \rightarrow 0$): $2\mathbb{Z}-0-0-0$. Abschließend sei noch erwähnt, dass es nur vier invertierte Koordinaten gibt, weil sich die Tabellen danach, analog zur Bott-Periodizität, wiederholen. Diese vierfache Periode folgt aus der Tatsache, dass es vier Unterklassen mit Kombinationen aus Vertauschungsrelationen und Operatoren gibt.

Nicht symorphe Symmetrien zweiter Ordnung

Als zweite Erweiterung sollen nun die nicht symorphen Partner zu den räumlichen Symmetrien zweiter Ordnung vorgestellt werden, zu denen u.a. die Gleit- und eine zweifache Schraubsymmetrie gehören. Analog zu dem vorherigen Beispiel können diese durch unitäre oder antiunitäre Operatoren (und entsprechende Antisymmetrien) repräsentiert werden. Allerdings erübrigt sich die Unterscheidung nach dem Vorzeichen von U^2 , weil durch die zusätzliche Verschiebung das Gitter nicht wieder auf sich selbst sondern nur noch auf eine translationsinvariante Kopie abgebildet werden kann. Als eine weitere Konsequenz ergibt sich, dass die AZ-Gruppen nur in zwei statt vier Unterklassen aufgeteilt werden (zumindest in den reellen Klassen). Daraus folgt dann auch eine Periode von zwei bezüglich der invertierten Koordinaten statt von vier.

Die Ergebnisse für die reellen Klassen mit einer invertierten Koordinate sind in Tab. 2.3 gezeigt. Die niedrigste Dimension $d = 1$ ist höher als in Tab. 2.2, weil in Anwesenheit einer NSS zweiter Ordnung die K-Theorie ansonsten zu einer inkonsistenten Beschreibung führt. Die NSS verursacht eine \vec{k} -abhängige algebraische Beziehung zwischen den Operatoren verschiedener Raumgruppen (räumlicher Symmetrien). Diese Beziehungen weisen dann, in Analogie zum

¹⁰ Streng genommen sind die Dimension eines Problems und die Anzahl der invertierten Koordinaten alleine nicht ausreichend zur Klassifizierung, da die Gruppen in der K-Theorie nur von Differenzen aus Dimensionen abhängen. Die Subtrahenden beschreiben eigentlich die Dimension D^* eines topologischen Defekts und die Anzahl seiner invertierten Koordinaten. Einfachheitshalber werden in dieser Einführung jedoch keine Defekte behandelt und $D^* = 0$ gesetzt.

Tabelle 2.3 Nicht symorphe Erweiterung der AZ-Klassifizierung: Klassifizierung von topologischen Invarianten gemäß verschiedenen Symmetrieklassen in Anwesenheit einer nicht symorphen, räumlichen Symmetrie zweiter Ordnung mit einer invertierten Koordinate, z.B einer Gleitspiegelung. Die Symmetrien werden durch unitäre (U) oder antiunitäre (A) Operatoren und deren Antisymmetrien (\bar{U} , \bar{A}) gemäß Gl. 2.6 repräsentiert. Die Null symbolisiert triviale Phasen, während Invarianten mit \mathbb{Z} durch ganze Zahlen und \mathbb{Z}_x durch die Restwertklasse Modulo x beschrieben werden. Das \oplus -Zeichen deutet an, dass zwei entsprechende Invarianten die Phase charakterisieren. Blaue Einträge sind zusätzliche Invarianten, welche durch die NSS bedingt sind. Analoge Tabellen für mehr oder weniger invertierte Koordinaten und die komplexen Klassen sind in Ref. [34].

Klasse	Symmetrie	$d = \dots$							
		1	2	3	4	5	6	7	8
AI	U, A für (AI, AII)	\mathbb{Z}_2	0	0	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}_4	\mathbb{Z}_2
BDI		\mathbb{Z}_2	\mathbb{Z}_2	0	0	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}_4
D	U, \bar{A} für (D, C)	\mathbb{Z}_4	\mathbb{Z}_2	\mathbb{Z}_2	0	0	0	\mathbb{Z}_2	\mathbb{Z}_2
DIII		\mathbb{Z}_2	\mathbb{Z}_4	\mathbb{Z}_2	\mathbb{Z}_2	0	0	0	\mathbb{Z}_2
AII	$[U, \mathcal{TC}] = 0 = [\bar{U}, \mathcal{TC}]$,	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}_4	\mathbb{Z}_2	\mathbb{Z}_2	0	0	0
CII	$[A, \mathcal{TC}] = 0 = [\bar{A}, \mathcal{TC}]$	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}_4	\mathbb{Z}_2	\mathbb{Z}_2	0	0
C	für BDI, DII, CII, CI	0	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}_4	\mathbb{Z}_2	\mathbb{Z}_2	0
CI		0	0	0	\mathbb{Z}_2	\mathbb{Z}_2	\mathbb{Z}_4	\mathbb{Z}_2	\mathbb{Z}_2
AI	\bar{U}, \bar{A} für (AI, AII)	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$
BDI		$\mathbb{Z} \oplus \mathbb{Z}_2$	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$
D	\bar{U}, A für (D, C)	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$	0	0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2
DIII		\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$	0	0	0	$2\mathbb{Z}$	0
AII	$\{U, \mathcal{TC}\} = 0 = \{\bar{U}, \mathcal{TC}\}$,	0	\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$	0	0	0	$2\mathbb{Z}$
CII	$\{A, \mathcal{TC}\} = 0 = \{\bar{A}, \mathcal{TC}\}$	$2\mathbb{Z}$	0	\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$	0	0	0
C	für BDI, DII, CII, CI	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$	0	0
CI		0	0	$2\mathbb{Z}$	0	\mathbb{Z}_2	$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	$\mathbb{Z} \oplus \mathbb{Z}_2$	0

Möbiusband¹¹, einen nicht triviale Twist¹² auf, welcher eben nur ab $d = 1$ definiert ist.

Als erstes soll zunächst der untere Block betrachtet werden, welcher mit Ausnahme der Klassen C und D durch Antisymmetrien und Antikommutatoren beschrieben wird. Hier kann erneut das Triplet, beginnend bei $d = 8$ in der AI Klasse, beobachtet werden. Allerdings reicht eine \mathbb{Z} -Invariante alleine zur Charakterisierung nicht mehr aus, sondern es tritt zusätzlich noch eine \mathbb{Z}_2 -Größe auf. Diese Extrainvariante tritt auch in der zweiten, nicht aber der niedrigsten Dimension des Triplets auf. Alle zusätzlichen, durch die NSS bedingten Invarianten sind blau hervorgehoben.

Noch größer ist der Einfluss der NSS, wenn der erste Block betrachtet wird. Hier sind pro Klasse und Dimension fünf Invarianten mit der Abfolge \mathbb{Z}_2 - \mathbb{Z}_2 - \mathbb{Z}_4 - \mathbb{Z}_2 - \mathbb{Z}_2 zu sehen, von denen vier neu sind. Auch die Restwertklasse Modulo 4 stellt eine völlig neue Art von Invariante dar, welche für Systeme mit TRS und Gleitsymmetrie in drei Dimensionen auftritt (Klasse AII). Auch wenn

¹¹Ein Möbiusband entsteht, wenn bei einem Band eine Seite um 180° gedreht wird, bevor es zu einem Ring zusammengefügt wird. Seine Fläche ist dann nicht orientierbar, d.h. kann weder als oben oder unten noch als innen oder außen charakterisiert werden. Darin unterscheidet es sich von einem Band, dass auf gleiche Weise aber ohne Drehung zusammengeführt wird. Diese beiden Bänder lassen sich nicht durch stetige Transformationen ineinander überführen, da dafür das Band geöffnet werden muss.

¹²Die Verdrehung eines Möbiusbands wird als Twist bezeichnet.

Tab. 2.3 und Tab. 2.2 nicht so leicht zu vergleichen sind, so kann aufgrund der \mathbb{Z}_4 -Charakteristik sofort festgestellt werden, dass es für eine symorphe Symmetrie (Spiegelung) keinen Fall mit einer gleichwertigen Invarianten gibt. Topologische Phasen, die durch verschiedenartige Invarianten beschrieben werden, sind aber offenkundig durch einen topologischen Phasenübergang getrennt. In Kap. 4 wird zu Beginn die Kristallstruktur von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ gezeigt. Diese weist die Besonderheit auf, dass sie im Falle einer geraden Gesamtzahl eine Gleit- und im Falle einer ungeraden Anzahl von Atomlagen eine Spiegelsymmetrie besitzt. Die sich hieraus ergebende Frage, ob in dieser Materialklasse ein atomlagenabhängiger Phasenübergang beobachtbar ist, kann im Grenzfall einzelner Lagen bejaht werden. Bei SnTe liegt die Ursache hierfür in der Tatsache, dass nicht symorphe Symmetrien im Vergleich zu ihren Partnersymmetrien zusätzliche Entartungen in der Bandstruktur garantieren [35]. Diese Entartung verhindert für Filme mit gerade Anzahl an Atomlagen, dass sich die Bandlücke schließen kann, während dieser Schutz bei der Spiegelsymmetrie bzw. einer ungeraden Anzahl von Lagen fehlt. Ohne diesen Schutz kann es dann durch (Größen-) Quantisierungseffekt oder Hybridisierung von Ober- und Unterseite zu einer erneuten Invertierung der invertierten Bandstruktur kommen, was zu einer trivialen Phase führt. Im Grenzfall unendliche dicker Filme muss die Bandstrukturen einer symorphen Symmetrie konvergieren gegen die Bandstruktur von ihrem nicht symorphen Partner. Dann spielt es keine Rolle mehr, dass NSS im Vergleich zu ihren Partnern zusätzliche Entartungen garantieren, weil diese einfach immer auftreten.

Die in dieser Arbeit vermessenen Proben besitzen jedoch alle makroskopische Ausmaße, sodass diese Effekte aufgrund der Filmdicke keine Rolle spielen sollten. Die Kristalle können in sehr guter Näherung als unendlich bzw. deren Oberflächen als halbumendliche Gitter betrachtet werden. Die sich daraus ergebenden periodischen Randbedingungen lassen sich nur mit der Einheitszelle modellieren. Diese weist im Fall von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ die Spiegelsymmetrie auf, während die Gleit Spiegelung eine extra Atomlage benötigt. Aus der Gleitsymmetrie-Einheitszelle lässt sich durch Translation also kein periodisches Gitter aufbauen. Dementsprechend sollten alle Eigenschaften durch die Symmetrie der Einheitszelle beschrieben werden. Andererseits zeigen Kristalle an ihren Oberflächen eine Vielzahl von Stufenkanten, insbesondere einatomare. Unter der Annahme, dass die Kristallsymmetrie nur über einen endlichen Einflussbereich auf die elektronische Struktur verfügt (z.B. wegen der endlichen Kohärenz von Wellenfunktionen), stellt sich die Frage, ob an der Oberfläche nicht doch ein gerade-ungerade-Effekt auftreten kann. Außerdem fällt bei genauer Betrachtung auf, dass ungerade Stufen selbst Gleitebenen des Kristalls sind. An diesen wird das Gitter gespiegelt und um einen halben Einheitsvektor senkrecht zur Oberfläche verschoben. Unter diesen Aspekten ist es schon fast nicht mehr überraschend, dass $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ an ungeraden Stufenkanten topologisch geschützte Kantenzustände besitzt, um deren Wechselwirkungen es in dieser Arbeit geht.

3 Rastertunnelmikroskopie

In diesem Kapitel soll die grundlegende Theorie der verwendeten Messmethode erläutert werden. Dabei werden zwei Ziele verfolgt, einerseits eine leicht zugängliche Interpretation der topographischen und spektroskopischen Daten zu ermöglichen, zugleich aber andererseits auch die Grenzen dieser vereinfachten Sichtweise aufzuzeigen. Diese kommen daher, dass die exakte Modellierung der zentrale Messgröße, des Tunnelstroms I_T , aufgrund ihres quantenmechanischen Ursprungs hoher Voraussetzungen bedarf: Die Kenntnis über die Wellenfunktionen von Probe und Rastersonde (Spitze). Durch die inhärente Verknüpfung von Wellenfunktion und Zuständen in der Quantenmechanik kann trotzdem durch verschiedene Annahmen und Näherungen ein direkter Zusammenhang zwischen I_T und der elektronischen Struktur der Probe hergestellt werden. Der Aufbau dieses Kapitels ist in dem Flussdiagramm in Abb. 3.1 dargestellt. Beginnend mit einem 1D-Rechteckpotential wird der quantenmechanische Tunneleffekt eingeführt. Anschließend wird im Rahmen des Bardeen-Modells ein Ausdruck I_B für den Tunnelstrom vorgestellt, allerdings für planare Tunnelstrukturen, da diese Theorie älter als das STM ist. Dabei taucht ein Matrixübergangselement M_{if} auf, das auf den Wellenfunktionen beider Elektroden basiert. Durch vernachlässigen der expliziten Energieabhängigkeit dieses Matrixelements und unter Zu-

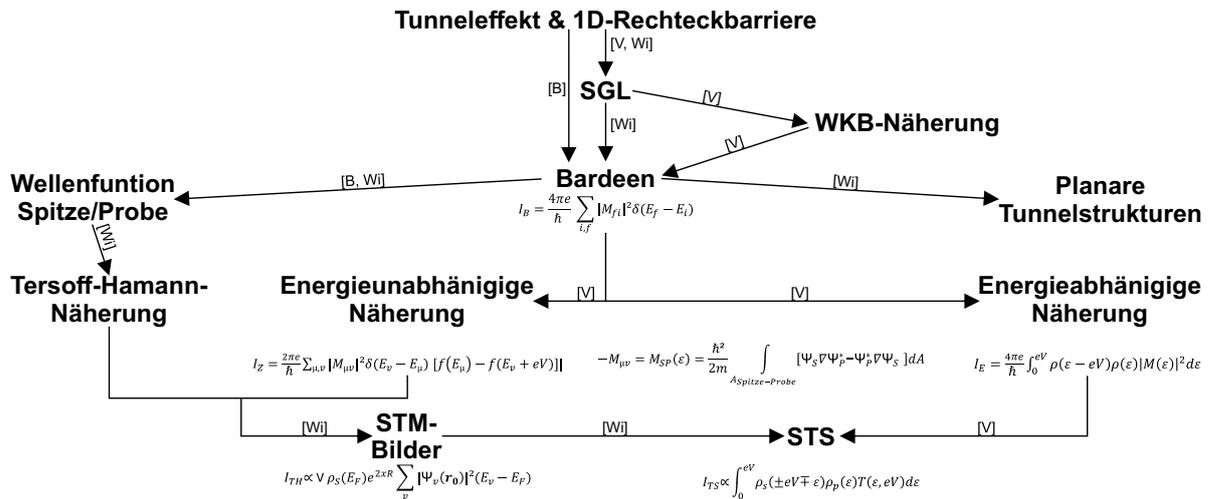


Abbildung 3.1 Schematischer Aufbau dieses Kapitels: Nach der Einführung des quantenmechanischen Tunneleffekts wird der Tunnelstrom im Bardeen-Modell vorgestellt und in verschiedenen Näherungen behandelt um STM-Bilder und Spektren interpretierbar zu machen. Wichtige Ausdrücke für den Tunnelstrom sind angegeben. Die Abkürzung SGL steht für die Schrödingergleichung. In Klammern wird, stark vereinfacht, dargestellt wie dieses Thema in der Fachliteratur behandelt wird: [B][36], [V][37] & [Wi][38].

hilfenahme des Tersoff–Hamann–Modells wird dann ein Ausdruck I_{TH} für die Interpretation von STM-Bildern abgeleitet. Im nächsten Schritt wird die Energieabhängigkeit berücksichtigt, um einen äquivalenten Ausdruck I_{TS} für die Rastertunnelspektroskopie (scanning tunneling spectroscopy, STS) zu erhalten.

3.1 Quantenmechanischer Tunneleffekt

Zu Beginn wird ein 1D–Rechteckpotential der Höhe U_0 und ein Teilchen mit Energie E betrachtet. Klassisch gesehen kann dieses die Energiebarriere nur dann überwinden, wenn $E > U_0$ gilt. In der Quantenmechanik existiert ein Welle–Teilchen–Dualismus, wodurch ein Teilchen Welleneigenschaften besitzen und durch eine Wellenfunktion beschrieben werden kann.

In Abb. 3.2 (a) ist für ein Teilchen mit $E < U_0$ als einfachster Ansatz eine nach rechts propagierende Welle der Form $\Psi(\vec{r}) = \Psi(z) = e^{-ikz}$ gezeigt. Im Bereich I gilt $U_0 = 0$ und Real- bzw. Imaginärteil der Welle werden durch Sinus und Kosinus beschrieben. Im Bereich II gilt $U_0 \neq 0$, sodass das Teilchen klassisch betrachtet hier nicht anzutreffen ist. Während die wesentlichen Schritte zur Lösung der Schrödingergleichung später in diesem Kapitel nachvollzogen werden, wird das wichtige Ergebnis für den Bereich II hier vorgezogen: Die Wellenfunktion fällt im Bereich II monoton ab, aber verschwindet nicht instantan. In Bereich III gilt wieder $U_0 = 0$, sodass dann wieder ein oszillatorisches Verhalten mit reduzierter Amplitude zu erwarten ist.

Da die Aufenthaltswahrscheinlichkeit das Betragsquadrat der Wellenfunktion ist, führt dieses Verhalten zu zwei klassisch unerwarteten Phänomenen. Erstens gibt es eine gewisse Wahrscheinlichkeit das Teilchen im klassisch verbotenen Bereich zu messen. Zweitens besteht die Möglichkeit das Teilchen hinter der Barriere in Bereich III zu detektieren, falls die Amplitude in Bereich II nicht vollständig abgeklungen ist.

Die Konsequenzen dieses Tunnelprozesses sind für das STM in Abb. 3.2 (b) & (c) veranschaulicht, wo Spitze und Probe zunächst als jeweils unabhängiges System mit eigenen Wellenfunktionen betrachtet werden. Solange ihr Abstand zueinander groß ist, werden Spitze und Probe jeweils durch ihre eigene Austrittsarbeit Φ und Fermienergie E_{F} charakterisiert (b). Die Fermienergie entspricht dabei der Energie bei der die Besetzungswahrscheinlichkeit eines elektronischen Zustands genau 50 % beträgt. Da Spitze und Probe in der Regel aus unterschiedlichen Materialien bestehen, liegt diese Grenze auch bei unterschiedlichen Energien. Ohne Beschränkung der Allgemeinheit wurde $\Phi_{\text{S}} > \Phi_{\text{P}}$ in der Abbildung angenommen. Weil Φ die Differenz von E_{F} zur Vakuumenergie beschreibt, gilt somit auch $E_{\text{F,S}} < E_{\text{F,P}}$. Im Energiebereich zwischen beiden Fermienergien existieren daher in der Probe überwiegend besetzte und in der Spitze unbesetzte Zustände.

Sobald Spitze und Probe soweit angenähert sind, dass ihre Wellenfunktionen überlappen, kön-

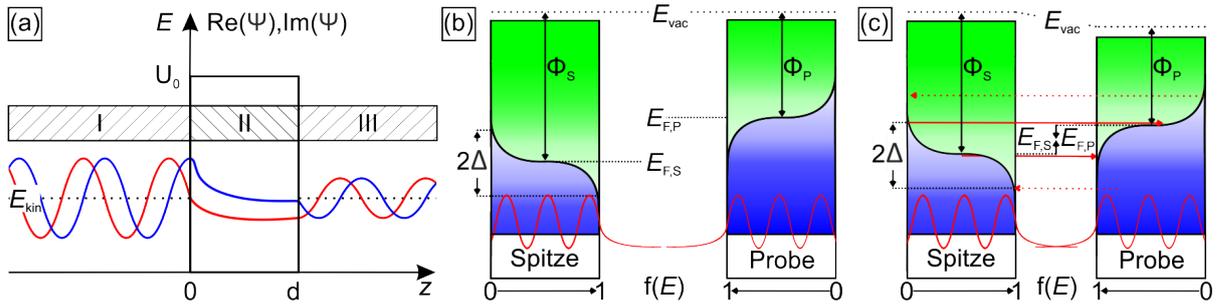


Abbildung 3.2 Tunnelprozess: (a) 1D-Rechteckpotential der Höhe U_0 und Real- und Imaginärteil einer nach rechts propagierenden ebenen Welle mit Energie $E < U_0$. Vor der Barriere oszillieren $Re(\Psi)$ und $Im(\Psi)$, im klassisch verbotenen Bereich II tritt eine starke Dämpfung auf und hinter der Barriere sind erneut Oszillationen, jedoch mit reduzierter Amplitude, zu beobachten. (b,c) Schematische Darstellung von Spitze und Probe als unabhängige Systeme mit jeweils eigener Wellenfunktion und Fermi-Dirac-Verteilung. Für große Abstände überlappen die Wellenfunktionen nicht (b), sodass Spitze und Probe jeweils ihre eigene Fermienergie haben. Sobald eine Überlagerung besteht (c), können Elektronen von besetzten in unbesetzte Zustände in beide Richtungen tunneln. Diese Bedingung ist für die Spitze (Probe) nur zwischen den durchgezogenen (gestrichelten) Pfeilen erfüllt. Durch die unterschiedliche Breite kommt es zu einem Nettoladungstransfer und die Fermienergien gleichen sich an. Erst wenn diese Angleichung erfolgt ist, liegt ein thermodynamisches Gleichgewicht vor. In (c) ist daher die Situation vor dem Erreichen des Gleichgewichts zu sehen.

nen sich Elektronen von der Spitze und Probe am selben Ort aufhalten. Da Elektronen quantenmechanisch ununterscheidbar sind und keiner klassischen Trajektorie folgen, gibt es keinen Automatismus zur Sortierung der Elektronen in dem gemeinsamen Aufenthaltsgebiet. Dementsprechend ist es möglich, dass ein Elektron, bevor es in diesen Bereich kommt, einen Zustand der Probe und danach einen der Spitze besetzt. Da der umgekehrte Weg der selben Argumentation folgt, kommt es zu einem Teilchenaustausch in beide Richtungen. Entscheidend für diesen Tunnelprozess ist, dass zu Beginn auf der einen Seite ein besetzter und auf der anderen Seite ein unbesetzter Zustand existiert. Aus dieser Einschränkung ergibt sich, dass aus der Spitze in Abb. 3.2 (c) nur solche Elektronen tunneln können, deren Energie zwischen den beiden durchgezogenen Pfeilen liegt. Oberhalb gibt es keine besetzten Zustände in der Spitze während bei Energien unterhalb freie Zustände in der Probe fehlen. Analog kann nur zwischen den beiden gestrichelten Pfeilen von der Probe in die Spitze getunnelt werden.

Die Breite dieses Energiebereichs hängt maßgeblich von dem Abstand der beiden Fermienergien zueinander und der Verbreiterung Δ der Fermi-Dirac-Verteilung ab. Wenn der energetische Abstand größer als die Summe beider Verbreiterungen ist, liegen die obersten besetzten Zustände der Spitze energetisch unterhalb der niedrigsten unbesetzten der Probe, sodass der Tunnelprozess vollständig unterdrückt wird. Gleichzeitig können aber weiterhin alle Zustände mit $E_{F,S} - \Delta < E < E_{F,P} + \Delta$ zum Tunnelprozess von der Probe in die Spitze beitragen. Durch den damit einhergehenden Ladungstransfer steigt das Fermienergielevel der Spitze während das der Probe fällt. Sobald sich die beiden angeglichen haben ist ein thermodynamisches Gleichgewicht entstanden, bei dem statistisch gleich viele Tunnelprozesse in beide Richtungen stattfinden.

Zum Abschluss dieser Einführung wird im Folgenden nun die Lösung der Schrödingergleichung nachvollzogen. Da das Potenzial zeitlich konstant ist, kann die Zeitabhängigkeit der Wellenfunktion mittels eines Separationsansatzes eliminiert werden, sodass die Schrödingergleichung die folgende Form annimmt:

$$\frac{\hbar^2}{2m} \nabla^2 \Psi(\vec{r}) = (U(\vec{r}) - E) \Psi(\vec{r}). \quad (3.1)$$

Während m die Teilchenmasse und \hbar das Wirkungsquantum beschreiben, gilt für das 1D-Rechteckpotential $\vec{r} = z$ und

$$U = \begin{cases} 0, & z < 0 & \text{(I)} \\ U_0, & 0 \leq z \leq d & \text{(II)} \\ 0, & z > d & \text{(III)}. \end{cases} \quad (3.2)$$

Für die allgemeine Lösung ist zu beachten, dass die Bereiche I und III äquivalent sind. Deswegen stellt auch eine nach links laufende Welle eine Lösung dar. Außerdem wirkt keiner der Operatoren aus Gl. 3.1 auf komplexe Zahlen, sodass jede Welle mit einer beliebigen komplexen Zahl als Amplitude ebenfalls eine Lösung ist. Darum ergibt sich als allgemeiner Ansatz:

$$\Psi(z) = \begin{cases} A \cdot e^{ikz} + B \cdot e^{-ikz} \\ C \cdot e^{i\gamma z} + D \cdot e^{-i\gamma z} \\ F \cdot e^{ikz} + G \cdot e^{-ikz}. \end{cases} \quad (3.3)$$

Durch einsetzen, differenzieren und Koeffizientenvergleich der linear unabhängigen Lösungen fallen die komplexen Amplituden heraus, sodass Auflösen nach k bzw. γ zu folgenden Ausdrücken führt:

$$k = \frac{\sqrt{2mE}}{\hbar}, \quad \gamma = \frac{\sqrt{2m(E - U_0)}}{\hbar}. \quad (3.4)$$

Wegen $U_0 > E$ wird γ im Bereich II rein imaginär, sodass das Argument der Exponentialfunktion aufgrund von $i^2 = -1$ reell wird. Die beiden Exponentialfaktoren beschreiben daher kein oszillatorisches Verhalten mehr, sondern ab $z = 0$ eine nach rechts- und ab $z = d$ eine nach links exponentiell abklingende Funktion.

Weil es im Bereich III keine Möglichkeit zur Reflexion gibt, kann durch die Annahme, dass kein Teilchen von links auf die Barriere trifft, $G = 0$ gesetzt werden. Zudem kann einer der Koeffizienten gleich eins gesetzt werden, was zu einer Reskalierung der anderen führt. Deshalb verbleiben vier unabhängige Koeffizienten zu bestimmen. Diese ergeben sich aus Stetigkeitsbedingungen

der Wellenfunktion und ihrer Ableitung an den beiden Rändern der Potenzialbarriere.

Durch $|\Phi_{\text{III}}|^2$ wird dann die Aufenthaltswahrscheinlichkeit hinter der Barriere und somit auch die Transmission T bestimmt:

$$T = |\Phi_{\text{III}}|^2 = |F|^2 = \frac{4k^2 \text{Im}^2(\gamma)}{(k^2 + \text{Im}^2(\gamma))^2 \sinh^2(\text{Im}(\gamma)d) + 4k^2 \text{Im}^2(\gamma)}. \quad (3.5)$$

Wegen $\gamma = \sqrt{2m\hbar^{-2}(E - U_0)} = i\sqrt{2m\hbar^{-2}(U_0 - E)}$, gilt für $\text{Im}(\gamma) = \sqrt{2m\hbar^{-2}(U_0 - E)}$. Wenn der zweite Summand des Nenners vernachlässigt und der Grenzfall $\text{Im}(\gamma)d \gg 1$ betrachtet wird, folgt mit der Näherung $\sinh^2(x) \approx \frac{1}{4}e^{2x}$

$$T \approx \frac{16k^2 \text{Im}^2(\gamma)}{[k^2 + \text{Im}^2(\gamma)]^2} e^{-2\text{Im}(\gamma)d}. \quad (3.6)$$

Aus Gl. 3.6 ergeben sich zwei wichtige Abhängigkeiten für die Transmission und daher auch für den Tunnelstrom. Erstens hängt dieser exponentiell von der Dicke der Barriere und somit dem Probe–Spitze–Abstand ab. Außerdem besteht eine Abhängigkeit von der Energie des Teilchens, weil über $\text{Im}(\gamma)$ auch die Wurzel der Energiedifferenz $U_0 - E$ in T eingeht.

Obwohl diese zwei Ergebnisse bereits sehr wertvolle Einblicke in den Tunnelprozess liefern, bleiben einige wichtige Aspekte des STM in diesem einfachen Modell jedoch unzureichend berücksichtigt. Zum einen wird weder eine Spannung zwischen Probe und Spitze noch die Anzahl der elektronischen Zustände, die Zustandsdichte (Density Of States, DOS), berücksichtigt. In das Modell geht nur eine monochromatische ebene Welle ein, die nicht normierbar ist und daher zur Interpretation als Wahrscheinlichkeit streng genommen nicht zulässig ist.

3.2 Bardeen–Modell

Nachdem zum Schluss des vorherigen Kapitels die Unzulänglichkeiten des Rechteckpotentials als Modell aufgezählt wurden, soll in diesem Kapitel nun ein differenzierteres Modell, basierend auf Bardeens Arbeit, vorgestellt werden. Grundlage ist der Tunnelprozess zwischen zwei Elektroden durch eine Oxidschicht, da das STM erst später erfunden wurde.

Ähnlich wie in Abb. 3.2 (b) werden die beiden Elektroden, bzw. Spitze und Probe, als unabhängige Systeme betrachtet. Deren Zustände $\Psi_{S,P}$ können durch Lösen der stationären Schrödingergleichung erhalten werden. Dabei gilt für das jeweilige Potential $U_{S,P}$ aus Gl. 3.1, dass es außerhalb der betrachteten Elektrode verschwindet. Zur Bestimmung des gekoppelten Systems wird dann $U_{\text{ges}} = U_S + U_P$ in die Schrödingergleichung (Schrödinger equation, SGL) eingesetzt und der Tunnelprozess mittels zeitabhängiger Störungstheorie erster Ordnung behandelt. Dieses Verfahren liefert eine Übergangsrate von einem Anfangszustand $|i\rangle$ in einen Endzustand $|f\rangle$ und

ist als Fermis Goldene Regel bekannt.

Die Wellenfunktion des Endzustands ergibt sich dann aus der Superposition vom Anfangszustand der Spitze und allen Zuständen der Probe [37]¹³:

$$\Psi_f(\vec{r}, t) = \Psi_{S,i}(\vec{r}, t) + \sum_k a_k(t) \Psi_{P,k}(\vec{r}, t). \quad (3.7)$$

Die Zeitabhängigkeit in Ψ ist durch einen Exponentialfaktor der Form $e^{i\frac{E}{\hbar}t}$ gegeben. Die Entwicklungskoeffizienten $a_k(t)$ verschwinden alle zum Zeitpunkt $t = 0$ und $|a_k(t)|^2$ gibt wie üblich die Wahrscheinlichkeit an, ein Elektron in dem dazugehörigen Zustand der Probe anzutreffen. Setzt man U_{ges} und Gl. 3.7 in die SGL 3.1 ein, so folgt für die Übergangsrate ω_{if} im Grenzwert

$\lim_{t \rightarrow \infty}$:

$$\omega_{if} = \frac{|a_f^\infty|^2}{t} = \frac{2\pi}{\hbar} |M_{if}|^2 \delta(E_f - E_i), \quad (3.8)$$

mit

$$M_{if} = \frac{-\hbar^2}{2m} \int_{A_{\text{Spitze-Probe}}} \Psi_f \nabla \Psi_i^* - \Psi_i^* \nabla \Psi_f d\vec{S}. \quad (3.9)$$

Das Integral wird über eine beliebige Fläche ausgeführt, welche Probe und Spitze separiert. Das Matrixübergangselement hängt somit nur von den Werten der Wellenfunktionen auf dieser Fläche ab. Da die Wellenfunktionen aus der SGL kommen und in diese das Potenzial U eingeht, besteht eine implizite Abhängigkeit von U [37].¹⁴

Durch Summieren über alle Kombinationen aus Anfangs- und Endzuständen ergibt sich dann die Übergangsrate von Spitze zu Probe. Diese muss mit $2e$ multipliziert werden um den Tunnelstrom I_B im Bardeen-Modell zu erhalten, weil jeder Zustand mit zwei Elektronenspins besetzt werden kann.

$$I_B = \frac{4e\pi}{\hbar} \sum_{i,f} |M_{if}|^2 \delta(E_f - E_i). \quad (3.10)$$

In den folgenden Unterkapiteln wird nun dieser Ausdruck genauer analysiert. Die Quintessenz dabei ist die Berechnung des Matrixelements und der dafür getroffenen Annahmen für die Wellenfunktionen. Da diese aus der SGL folgen, sind die Wellenfunktionen nicht unabhängig vom Potenzial $U(z)$. Deswegen können entweder direkt bestimmte Wellenfunktionen oder aber ein spezifischer Potenzialverlauf $U(z)$ angenommen werden. Im ersten Fall wird, wie im Tersoff-Hamann-Modell, die Energieabhängigkeit des Matrixelements vernachlässigt und im letzteren

¹³ Ohne Beschränkung der Allgemeinheit wird ein Tunnelprozess von der Spitze in die Probe betrachtet.

¹⁴ Für weitere Details zur Herleitung siehe Appendix B von Ref. [37]

explizit berücksichtigt.

3.2.1 STM-Bilder im Tersoff-Hamann-Modell

Zunächst sollen für die Wellenfunktionen von Spitze und Probe direkt explizite Formen angenommen werden. Dieser Ansatz ist jedoch auf kleine Tunnelspannungen beschränkt, da er die Energieabhängigkeit des Matrixelements vernachlässigt. Für die Probe setzten Tersoff und Hamann eine Fourierentwicklung einer ebenen Welle und für die Spitze ein s-förmiges Orbital mit Radius R um \vec{r}_0 an. Aus dem minimalen Abstand d_{\min} zwischen Probe, Spitze und R ergibt sich dann der Abstand zwischen Probe und dem Punkt r_0 zu $r_0 = d_{\min} + R$. Für tiefe Temperaturen lässt sich dann folgende Proportionalität für den Tunnelstrom im Tersoff-Hamann-Modell finden [38]:

$$I_{\text{TH}} \propto V \rho_s(E_{\text{F}}) e^{2xR} \sum_{\nu} |\Psi_{\nu}(\vec{r}_0)|^2 \delta(E_{\nu} - E_{\text{F}}). \quad (3.11)$$

Der Strom ist also proportional zur angelegten Spannung und zur Zustandsdichte der Spitze am Fermineau. Die Exponentialfunktion e^{2xR} wird durch $|\Psi_{\nu}(\vec{r}_0)|^2 \propto e^{-2x(d_{\min}+R)}$ aufgehoben, sodass die zu erwartende exponentielle Dämpfung mit d_{\min} folgt ($I_{\text{TH}} \propto e^{-2xd_{\min}}$). In die Abklingkonstante $x = \sqrt{2m\Phi}$ geht die Austrittsarbeit Φ ein. Der Faktor $\Psi_{\nu}(\vec{r}_0)$ beschreibt die Wahrscheinlichkeit, ein Teilchen mit Energie E_{ν} am Ort \vec{r}_0 , dem Ursprung des s-Orbitals der Spitze, zu messen.

Für ein System mit N diskreten Energieniveaus kann die Anzahl der Zustände in einem Energieintervall $[E_1, E_2]$ mit Hilfe der δ -Distribution ermittelt werden. Da der Ausdruck $\delta(E - E_n) = 1$ für $E = E_n$ liefert, kann N ermittelt werden indem einerseits das Energieintervall durchgefahen (Integration) und andererseits über alle Zustände summiert wird.

$$\int_{E_1}^{E_2} \sum_n \delta(E - E_n) dE = \sum_n \int_{E_1}^{E_2} \delta(E - E_n) dE = \sum_n 1 = N. \quad (3.12)$$

Die Zustandsdichte, definiert als Anzahl der Zustände pro infinitesimalem Energieintervall, $\rho = \frac{dn}{dE}$ kann mit Gl. 3.12 daher so dargestellt werden:

$$\rho(E) = \sum_n \delta(E - E_n). \quad (3.13)$$

Setzt man $E = E_{\text{F}}$, so entspricht dieser Term, bis auf einen Gewichtungsfaktor $|\Psi_{\nu}|^2$, der Summe in Gl. 3.11. Dementsprechend lässt sich die gesamte Summe als Zustandsdichte der Probe bei der Fermienergie interpretieren, jedoch am Ort der Spitze. Verallgemeinert laute die Formulierung

für diese sogenannte lokale Zustandsdichte (Local Density Of States, LDOS):

$$\rho_{\text{local}}(E, \vec{r}) = \sum_n |\Psi_n(\vec{r})|^2 \delta(E - E_n). \quad (3.14)$$

Wird also ein STM-Bild mit konstantem Tunnelstrom aufgenommen, so zeichnet die Spitze eine Kontur der LDOS am Fermi-niveau nach. Äquivalent kann dies auch als der Teil der Ladungsdichte interpretiert werden, zu dem nur Zustände am Fermi-niveau beitragen (generell tragen zur Ladungsdichte Zustände aller Energien bei).

3.2.2 STS-Spektren

In diesem Abschnitt soll die Energieabhängigkeit des Matrixelements direkt betrachtet werden. Da zu jeder Wellenfunktion in Gl. 3.9 eine wohldefinierte Energie gehört, kann die Abhängigkeit von den Wellenfunktionen durch eine Abhängigkeit von den Energien E_i & E_f ersetzt werden. Da in Gl. 3.10 die δ -Distribution steht, gibt es nur für $E_i = E_f$ von Null verschiedene Beiträge. Durch die δ -Distribution wird die Energieerhaltung sichergestellt, d.h. es werden nur elastische Tunnelprozesse betrachtet. Im Ergebnis führt dies dazu, dass das Matrixelement nur von der Energie des Endzustands abhängt: $M_{fi} = M_{fi}(E_f)$. Somit kann der gesamte Ausdruck in der Summe von Gl. 3.10 als Ergebnis der Integration über zwei δ -Distributionen bezüglich der Energie identifiziert werden:

$$I = \frac{4e\pi}{\hbar} \sum_{i,f} \int_{-\infty}^{\infty} |M_{fi}(E)|^2 \delta(E - E_i) \delta(E - E_f) dE \quad (3.15)$$

$$= \frac{4e\pi}{\hbar} \int_{-\infty}^{\infty} |M_{fi}(E)|^2 \sum_i \delta(E - E_i) \sum_f \delta(E - E_f) dE \quad (3.16)$$

$$= \frac{4e\pi}{\hbar} \int_{-\infty}^{\infty} |M_{fi}(E)|^2 \rho_i(E) \rho_f(E) [f(E - eV) - f(E)] dE \quad (3.17)$$

$$= \frac{4e\pi}{\hbar} \int_0^{eV} |M_{fi}(E)|^2 \rho_{\text{Spitze}}(E - eV) \rho_{\text{Probe}}(E) dE \quad (3.18)$$

Durch Vertauschen von Summe und Integration sowie durch Aufspaltung in zwei Summen wird ersichtlich, dass in Gl. 3.16 zweimal die Definition der Zustandsdichten (Gl. 3.12) auftritt. Einmal bezogen auf $|i\rangle$ und einmal auf $|f\rangle$. Da die beiden Ausdrücke symmetrisch im Integrand stehen sind die Zustandsdichten von Probe von Spitze gleichermaßen für den Tunnelstrom relevant.

In der Praxis ist jedoch die DOS der Spitze oft unbekannt und wird einfachheitshalber als konstant angenommen. In Gl. 3.17 wurde die Fermi-Dirac-Funktion $f(E)$ eingeführt, weil für endliche Temperaturen die alleinige Angabe einer Zustandsdichte physikalisch nicht sinnvoll ist.

Wegen des Pauli-Prinzips muss für Fermionen stets noch die Besetzungswahrscheinlichkeit berücksichtigt werden, da ein bereits besetzter Zustand nicht zur Verfügung steht. Andererseits ist auch die alleinige Angabe der Fermi-Dirac-Funktion unzureichend, da diese kontinuierlich bezüglich E ist, Zustände aber diskret sind. Sie ist folglich auch bei Energien definiert bei denen gar keine Zustände existieren. In Gl. 3.18 wurde die Besetzungswahrscheinlichkeit in den Integrationsgrenzen berücksichtigt, wobei $T = 0$ K angenommen wurde. Die obere Grenze entspricht dabei der von Außen angelegten Spannung zwischen Probe und Spitze, was in erster Näherung zu einer relativen Verschiebung der beiden DOS zueinander um $|eV|$ führt. Dementsprechend existiert kein gemeinsames Ferminiveau mehr, das als Referenz für Probe und Spitze dienen kann. Daher wurde im letzten Gleichheitszeichen dieser Versatz berücksichtigt, weil ansonsten die DOS der Spitze auf das Ferminiveau der Probe Bezug nimmt.

Um das Matrixelement zu berechnen muss nun als nächstes ein spezifischer Potenzialverlauf angenommen werden. Solange die Wellenlänge des Teilchens klein gegenüber der räumlichen Änderung des Potenzials ist, kann im Rahmen der Wentzel-Kramers-Brillouin-Näherung (Wentzel Kramers Brillouin, WKB) der Verlauf durch infinitesimale Rechtecke genähert werden. Dadurch wird die Abklingkonstante $\gamma(z)$ des 1D-Rechteckpotenzials z -abhängig und im Exponent der e -Funktion muss über sie integriert werden. Der Transmissionsfaktor $T(d, E)$ ersetzt dann das Matrixelement in Gl. 3.18.

Eine Abweichung von einem einfachen Rechteck entsteht z.B. durch unterschiedliche Austrittsarbeiten von Probe und Spitze oder durch eine angelegte Spannung. Beides führt jeweils zu einem trapezförmigen Potenzialverlauf, der in erster Näherung durch ein Rechteck mit Höhe des Mittelwerts beschrieben werden kann. Die effektive Barriere setzt sich dann aus der mittleren Austrittsarbeit und der mittleren Spannung zusammen, sodass $U_0 - E$ in $\text{Im}(\gamma)$ durch die effektive Barriere $\Phi_{\text{eff}} = \frac{\Phi_S + \Phi_P + eV}{2} - E$ ersetzt wird. Wegen der geringen Dicke der Tunnelbarriere im STM ist diese Näherung akzeptabel obwohl T exponentiell und nicht linear von der Barrierrhöhe abhängt.

Als wichtiges Ergebnis ist festzuhalten, dass das Matrixelement durch einen Transmissionsfaktor $T(d, E, \Phi_{\text{eff}})$ ersetzt wird. Infolgedessen haben Elektronen eine umso niedrigere Transmissionswahrscheinlichkeit, je kleiner ihre Energie ist, weil sie eine höhere effektive Barriere sehen. Dieser Effekt ist in Abb. 3.2 (c) und im Folgenden durch die Pfeillänge visualisiert.

Da bei tiefen Temperaturen nur Elektronen bis zur Energie eV tunneln können, sind für den Fall kleiner Spannungen $eV \ll \Phi_{S,P}$ die letzten beiden Terme von Φ_{eff} vernachlässigbar und T hängt nur noch von dem Abstand zwischen Spitze und Probe ab. Da die Austrittsarbeit von Festkörpern allgemein in der Größenordnung einiger Elektronenvolt liegt und die Tunnelspannung der in dieser Arbeit aufgenommene Daten bis zu wenigen 100 mV beträgt, ist davon auszugehen, dass der Einfluss der Spannung auf T nicht vollständig vernachlässigbar ist.

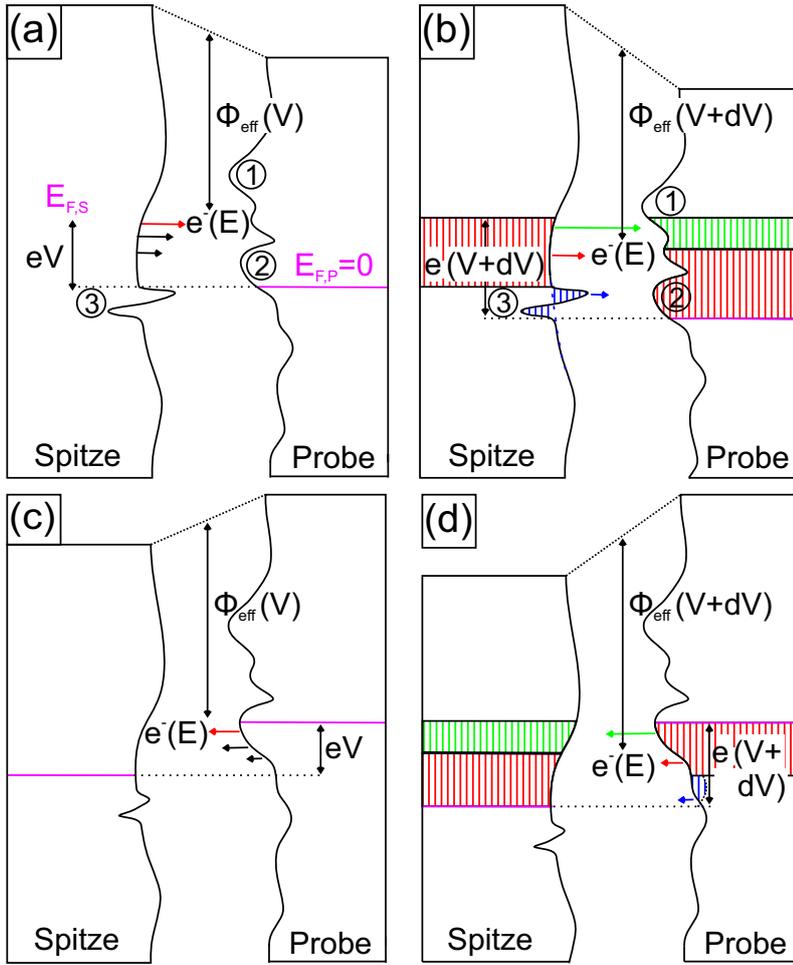


Abbildung 3.3 dI/dV -Beiträge: (a) Zustandsdichte von Probe und Spitze für $V > 0$. Wegen der Näherung $T = 0$ K findet der Tunnelprozess nur bei Energien zwischen beiden Fermi-niveaus statt. (b) Nach dem Anlegen einer zusätzlichen Spannung dV sind die beiden DOS relativ zueinander verschoben, sodass auf Seiten der Probe bei hohen (grün) und auf Seite der Spitze bei niedrigeren Energien neue Zustände (blau) zum Tunneln beitragen. Rote Bereiche tragen in (a) & (b) zum Tunnelprozess bei, besitzen in (b) jedoch einen kleineren Transmissionsfaktor. Die Farben repräsentieren die verschiedenen Beiträge zum dI/dV -Signal (vgl. Gl. 3.19): Grün markiert den ersten, rot den zweiten und blau den letzten Term. (c,d) Analoge Darstellung für $V < 0$. Adaptiert nach [37].

Die gewünschte Information über die Zustandsdichte der Probe ist in Gl. 3.18 im Integrand zu finden. Direktes differenzieren des Tunnelstroms nach der Spannung [differenzielle Leitfähigkeit, differential conductivity, (dI/dV)] ist jedoch auch nach der Substitution $dE = e dV$ nicht möglich, weil die obere Integralgrenze ebenfalls von V abhängt. Durch Anwenden der Leibnizregel für Parameterintegrale ergibt sich dann:

$$\begin{aligned}
 \frac{dI}{dV} = & \frac{4e\pi}{\hbar} [e\rho_S(0)\rho_P(eV)T(eV, V) \\
 & + \int_0^{eV} \rho_S(E - eV)\rho_P(E) \frac{\partial T(E, V)}{\partial V} dV \\
 & + \int_0^{eV} \frac{\partial \rho_S(E - eV)}{\partial V} \rho_P(E) T(E, V) dV].
 \end{aligned} \tag{3.19}$$

Die drei Beiträge sind in Abb. 3.3 durch verschiedenfarbige Pfeile markiert. In (a) ist eine Spannung $V > 0$ angelegt, sodass das Fermi-niveau der Spitze oberhalb des der Probe liegt. Der Referenzpunkt für die Energie ist das Fermi-niveau der Probe $E = 0 = E_{F,P}$. Wird nun eine zusätzliche Spannung dV angelegt (b), so kommt es zu einer Verschiebung der gesamten DOS relativ zueinander. Die Zustände der Spitze werden nach oben bzw. die der Probe nach unten

geschoben. Der Peak 1 liegt dabei zunächst oberhalb von $E_{F,S}$. Durch die zusätzliche Spannung können nun Elektronen mit $E_{F,S} + eV < E < E_{F,S} + e(V + dV)$ in die Zustände des Peaks 1 tunneln. Dieser zusätzliche Beitrag entspricht in (b) dem grün schraffierten Bereich und in Gl. 3.19 dem ersten Term ($V \gg dV \approx 0$). Er wird bestimmt durch die DOS der Spitze am Fermi-niveau der Spitze, der DOS der Probe bei eV oberhalb des Fermi-niveaus der Probe und gewichtet mit einem Transmissionsfaktor, dessen effektive Barriere bei eV bestimmt wird.

Der Transmissionsfaktor erfährt durch dV ebenfalls eine Änderung, weil Φ_{eff} um $\frac{dV}{2}$ zunimmt. Gleichzeitig gilt jedoch für den grünen Bereich, dass die Zustände der Spitze um edV nach oben verschoben sind. Da E vollständig in Φ_{eff} eingeht, führt dies dazu, dass für die obersten Zustände die effektive Barriere verringert ist. Daher tragen die Elektronen bei $E_{F,S}$ in (b) stärker zum Tunnelstrom bei als je in (a).

Der Transmissionsfaktor, genau genommen seine Änderung, ist für den zweiten Beitrag entscheidend. Der rote Bereich hebt denjenigen Anteil der DOS hervor, der bereits vor dem Anlegen von dV zum Tunnelstrom beigetragen hat. In dem Energiebereich wo sich die rot schraffierten Flächen überlappen, zeichnen sich Elektronen durch eine konstante Energie E aus, welche in (a) & (b) gleich ist. Veranschaulicht wird dies durch einen roten Pfeil, der in beiden Teilbildern den gleichen Abstand zu $E_{F,P}$ hat und somit ein Elektron mit der selben Energie zeigt. Für dieses ändert sich die effektive Barriere ausschließlich um $\frac{dV}{2}$, da E konstant bleibt. Daher nimmt die Transmission für diese Zustände mit zunehmender Spannung ab. Der Effekt dieser Änderung von T auf das dI/dV -Signal wird durch den zweiten Term beschrieben und kann prinzipielle sogar zu einer negativen differentiellen Leitfähigkeit führen, wenn die Abnahme durch den zweiten Term größer als die Zunahme durch den ersten Term ist.

Der dritte Term wird nur durch die Änderung der DOS der Spitze bestimmt. In diesem Energiefenster haben die vorhandenen Zustände der Probe (Peak 2) bereits vor dem Anlegen von dV zum Tunnelprozess beigetragen. Wenn nach der durch dV verursachten Verschiebung der Spitzen-DOS [solide, schwarze Linie in (b)] in diesem Energiebereich nun mehr Zustände sind als ursprünglich (gestrichelte, blaue Linie), dann nimmt der Tunnelstrom zu. Dies ist bei dem blau schraffierten Bereich von Peak 3 der Fall. Wenn die DOS durch die Verschiebung abnimmt (z.B durch die Vertiefung unterhalb von Peak 3), so wird dieser Beitrag sogar negativ, was im Extremfall insgesamt zu einer negativen differentiellen Leitfähigkeit führen kann. Da diese Zustände jedoch die größte effektive Barriere besitzen, ist ihr Beitrag zum dI/dV -Signal stark durch T gedämpft. Im Fall einer konstanten DOS würde eine Verschiebung um edV auf der Seite der Spitze zu keiner Änderung führen und der dritte Term in Gl. 3.19 verschwindet.

In den Teilbildern (c) & (d) ist das Vorzeichen der Tunnelspannung umgekehrt, sodass nun Elektronen im Energiefenster zwischen $E_{F,S}$ und $E_{F,P}$ aus der Probe in die Spitze tunneln. Der erste Term aus Gl. 3.19 beschreibt nun einen Tunnelprozess mit hohem Transmissionsfaktor in

die Zustände der Spitze. Der zweite Term die Modifikation durch die Änderung von T und der letzte Term die Anpassung aufgrund der Zustände, die auf Probenseite von unten in das Energiefenster geschoben wurden. Dadurch wird in diesem Fall das dI/dU -Signal stärker von der DOS der Spitze geprägt. Dieser Umstand ist aus zwei Gründen problematisch, weil zum einen der Experimentator primär an der DOS der Probe interessiert ist. Zum anderen ist die DOS der Spitze nur bedingt experimentell kontrollierbar und deren Struktur daher auch nur näherungsweise bekannt. Veranschaulicht wird diese Problematik in Abb. 3.3 (d) dadurch, dass der blaue Beitrag größer als in (b) ist. Dies liegt an der Struktur der angenommenen DOS, welche für die Spitze eher glatt ist und für die Probe mehrere ausgeprägte Maxima und Minima besitzt. Somit ergibt sich auch eine unterschiedlich starke Änderung $\frac{\partial \rho}{\partial V}$.

Zusammengefasst ergibt sich, dass für positive Probenspannungen ($V > 0$) die DOS der Probe und für negative Probenspannungen ($V < 0$) die DOS der Spitze einen dominierenden Beitrag zum dI/dV -Signal liefern. In der Folge kommt es oft zu Schwierigkeiten Tunnelspektren im Bereich negativer Spannungen zu reproduzieren, da im Experiment Wissen und Kontrolle über die DOS der Spitze begrenzt sind. Ursache hierfür ist die Spannungsabhängigkeit von T . Zwar tragen bei einer Spannungsänderung dV sowohl in der Spitze wie auch in der Probe neue Zustände zum Tunnelprozess bei, aber für positive (negative) Probenspannungen liegen die Zustände in der Probe energetisch oberhalb (unterhalb) von denen der Spitze. Da mit abnehmender Energie T kleiner wird, wird das Spektrum daher stärker von der DOS der Probe (Spitze) geprägt.

Zum Abschluss dieses Kapitels sollen noch auf einige Aspekte mit Bezug zu den durchgeführten Experimenten eingegangen werden. Prinzipiell sind zwei Wege möglich, um an das dI/dV -Signal zu gelangen. Es kann die I - V -Kurve gemessen und anschließend numerisch differenziert werden oder es wird, während die Gleichspannung durchgeföhren wird, zusätzlich eine kleine Wechselspannung angelegt und das Messsignal mit einem Lock-In-Verstärker gefiltert. In dieser Arbeit wurde nur letztere Methode verwendet, weil der Lock-In-Verstärker ein drastisch verbessertes Signal zu Rausch Verhältnis liefert. Dies wird erreicht, indem alle Signalanteile die nicht zur Modulationsspannung V_{Mod} in Phase sind oder deren Frequenz nicht mit der der Modulation übereinstimmt, herausgefiltert werden. Zusammen mit der Temperatur T bestimmt die Modulation auch die experimentelle Energieauflösung:

$$\Delta E = \sqrt{(2eV_{\text{Mod}}^{\text{RMS}})^2 + (0.28 \text{ meV/K} \cdot T)^2}. \quad (3.20)$$

Hieraus ergibt sich unter den üblichen experimentellen Bedingungen ($T = 5 \text{ K}$, $V_{\text{Mod}}^{\text{RMS}} \approx 1 \text{ mV}$) eine maximale theoretische Energieauflösung von 1.7 meV. Die tatsächliche Auflösung ist jedoch schlechter, weil durch die Messelektronik zusätzliches Rauschen entsteht.

Um den Einfluss von T zu minimieren kann die normierte differentielle Leitfähigkeit $\frac{dI/dV}{I/V}$

betrachtet werden. Als grober Richtwert ergibt sich, wenn man plausible Werte für die Parameter in aus Gl. 3.6 einsetzt, dass sich der Transmissionsfaktor nach etwa 300 mV halbiert hat. Dieser Wert ist größer als der typische Messbereich in dieser Arbeit, was zeigt, dass der Einfluss von T noch nicht dominant ist. Daher wird in der vorliegenden Arbeit in der Regel auf eine Normierung verzichtet.

4 Der topologische kristalline Isolator $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$

Bevor im folgenden Kapitel die neu gewonnenen Ergebnisse diskutiert werden, soll zunächst noch der aktuelle Stand der Forschung bezüglich des Materialsystems $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ zusammengefasst werden. Da sich die vorliegende Arbeit mit topologisch kristallinen Isolatoren und deren elektronischer Struktur befasst, wird als erstes die Kristall- und die Bandstruktur vorgestellt. Besonderes Augenmerk wird dabei auf die (001)–Oberfläche gelegt, da die Ergebnisse aus Kap. 5 auf dieser gemessen wurden. Anschließend werden die topologischen Aspekte in der Bandstruktur und in verschiedenen Legierungen der (Sn,Pb)(Se,Te)–Materialklasse diskutiert. Zum Abschluss werden die charakteristischen Punkte der Bandstruktur mit Features in den Tunnelspektren assoziiert, um dann die exotischen, eindimensionalen Kantenzustände dieser Arbeit einzuführen. Das Kapitel endet mit der Diskussion der gegenwärtigen Modelle zur topologischen Beschreibung der Kantenzustände

4.1 Kristall- & Bandstruktur

Die Mischkristalle SnTe, PbSe und PbTe kristallisieren alle in einem kubisch flächenzentrierten Gitter mit zweiatomiger Basis [siehe Abb. 4.1 (a)]¹⁵. Diese Struktur entspricht der des Kochsalzes NaCl und analog dazu umfasst die Basis je ein Anion [Se^{2-} oder Te^{2-} , blau in Abb. 4.1 (a)] und ein Kation (Sn^{2+} oder Pb^{2+} , rot). Aufgrund der identischen Kristallstruktur sind auch Legierungen möglich, die aus drei oder allen vier verschiedenen Ionen bestehen. Hierbei kommt es zu einer zufälligen Verteilung der Sn- und Pb–Ionen im Kationen-, bzw. von Se- und Te–Ionen im Anionen–Untergitter. Durch die Stöchiometrie der Legierung lassen sich physikalische Eigenschaften, wie beispielsweise Gitterabstand, Leitfähigkeit oder Bandlücke, kontinuierlich variieren. Die Grenzen sind dabei durch die Eigenschaften der reinen Mischkristalle gegeben. $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ ist ein solches Beispiel, bei dem sich insbesondere die Bandlücke an einem Hochsymmetriepunkt der Volumen–Brillouin–Zone als Funktion der Sn–Konzentration x ändert. Dieser Umstand wurde für den experimentellen Nachweis topologisch kristalliner Isolatoren genutzt [20–22].

Um diesen topologischen Phasenübergang zwischen PbSe und PbSe zu verstehen, muss zunächst noch die Bandstruktur genauer betrachtet werden. Dabei liegt der Fokus auf der (001)–

¹⁵ Der vierte binäre Mischkristall SnSe kristallisiert in einem orthorhombischen Gitter (drei rechte Winkel, aber auch drei verschiedene Gitterkonstanten). Für die später betrachteten Legierungen ist die Sn–Konzentration jedoch hinreichend klein, sodass diese durch ein fcc–Gitter beschrieben werden.

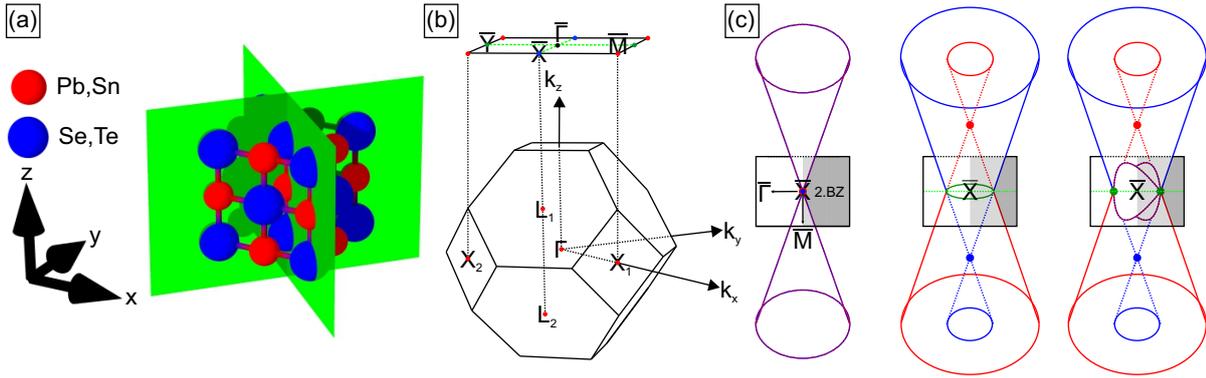


Abbildung 4.1 Gitter, Brillouin-Zone und Dirac-Kegel von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$: (a) Einheitszelle des fcc-Gitters mit zweiatomiger Basis der (Sn,Pb)(Se,Te) Materialklasse. In grünen sind die (110)- und (1 $\bar{1}$ 0)-Spiegelebenen zu sehen, welche senkrecht zur (001)-Oberfläche stehen. (b) Volumen-Brillouin-Zone und ausgewählte Hochsymmetriepunkte, sowie deren Projektion auf die (001)-Oberfläche. Die Punkte \bar{Y} und \bar{X} sind für das unverzerrte Gitter äquivalent und liegen auf den Projektionen der in (a) grün gezeigten Ebenen. (c) Entwicklung der Oberflächenzustände, ausgehend von zwei aufeinander projizierten Dirac-Kegeln am \bar{X} -Punkt (links). Durch Streuprozesse zwischen L_1 und L_2 wird ein Kegel energetisch nach oben (rot) und der andere nach unten (blau) verschoben. Die gestrichelte Linie verläuft entlang von $\bar{\Gamma}\bar{X}$, die dunkel grüne Linie markiert die Schnittpunkte der zwei Dirac-Kegel (mitte). Durch Hybridisierung werden alle Schnittpunkte außer den beiden auf der $\bar{\Gamma}\bar{X}$ -Spiegelinie eliminiert. Die Spiegelsymmetrie bezüglich $\bar{\Gamma}\bar{M}$ beschützt die Kreuzungspunkte (grüne Punkte). Diese Punkte zeigen eine lineare Dispersion, liegen jedoch nicht mehr am \bar{X} -Punkt. Sie sind äquidistant entlang von $\bar{\Gamma}\bar{X}$ verschoben (links). Graue Bereiche markieren die zweite Brillouin-Zone.

Oberfläche, da der überwiegend ionische Bindungscharakter diese unpolaren Oberflächen bei Spalten von Einkristallen energetisch bevorzugt [39].

In Abb. 4.1 (b) ist die Volumen-Brillouin-Zone sowie die Projektion auf die (001)-Oberfläche zu sehen. Während die Brillouin-Zone der (001)-Oberfläche, genauso wie die Einheitszelle des Realraums, quadratisch ist, ergibt sich für die Brillouin-Zone des Volumens durch die Konstruktion der Wigner-Seitz-Zelle ein Gebilde aus mehreren Quadraten und Sechsecken. Das Zentrum der Quadrate wird durch den Schnittpunkt der Brillouin-Zone mit den Impulsachsen definiert, sodass es insgesamt sechs Quadrate gibt ($k_{\pm x}, k_{\pm y}, k_{\pm z}$). Diese Schnittpunkte sind wegen der 90°-Rotationssymmetrie des fcc-Gitters äquivalent zueinander und stellen gleichzeitig den Hochsymmetriepunkt X der Volumen-Brillouin-Zone dar. Die Quadrate sind durch Sechsecke miteinander verbunden, wobei nur jede zweite Kante eines Sechsecks an ein jeweils anderes Quadrat grenzt. Die anderen drei Kanten grenzen jeweils an ein benachbartes Sechseck. Hieraus ergibt sich, dass die Volumen-Brillouin-Zone durch

$$\frac{\text{Anzahl Quadrate} \times \text{Anzahl der Quadratkanten}}{\text{Anzahl Sechseckkanten} - \text{Anzahl Sechseck-Sechseck-Kanten}} = \frac{6 \times 4}{6 - 3} = 8 \quad (4.1)$$

Sechsecke vervollständigt wird. Die Mittelpunkte dieser Sechsecke sind ebenfalls Hochsymmetriepunkte und werden konventioneller Weise mit L bezeichnet. Sie markieren die Schnittpunkte der Raumdiagonalen der kubischen Einheitszelle mit dem Rand der 1. Brillouin-Zone. Bis auf

das Zentrum der Brillouin-Zone Γ sind aus Gründen der Übersichtlichkeit keine weiteren Hochsymmetriepunkte in Abb. 4.1 (b) eingezeichnet.

Die Oberflächen-Brillouin-Zone der (001)-Oberfläche wird in den Ecken durch die Projektion der X -Punkte begrenzt und mit \bar{M} bezeichnet. Die Richtung $\bar{\Gamma M}$ entspricht somit der einer Impulsachse im reziproken Raum und einer $\langle 100 \rangle$ -Richtung im Realraum. Der allgemeinen Konvention entsprechend werden alle Hochsymmetriepunkte der Oberflächen-Brillouin-Zone mit einem Querstrich versehen. Diese Konvention ist besonders wichtig, da bei der (001)-Oberfläche der zweite Hochsymmetriepunkt üblicherweise mit \bar{X} bezeichnet wird und weder mit dem X -Punkt der Volumen-Brillouin-Zone verwechselt werden sollte, noch direkt mit diesem zusammenhängt. Tatsächlich zeichnen sich die \bar{X} -Punkte dadurch aus, dass sie die Projektion von zwei L -Punkten der Volumen-Brillouin-Zone sind. Sie liegen zwischen zwei \bar{M} -Punkten, also auf dem Rand der Brillouin-Zone.

Die Richtung $\bar{\Gamma X}$ ist gegenüber $\bar{\Gamma M}$ um 45° gedreht und entspricht daher der $\langle 110 \rangle$ -Richtung im Ortsraum. Somit markieren die gestrichelten grünen Linien aus Abb. 4.1 (b) die Projektionen der (110)- und ($1\bar{1}0$)-Spiegelebenen aus Abb. 4.1 (a). In $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ wurde experimentell eine Gitterverzerrung festgestellt, bei der die beiden Untergitter entlang der [110]-Richtung zueinander hin verschoben sind. Dadurch bleibt die (110)-Spiegelsymmetrie erhalten, während sie bezüglich der ($1\bar{1}0$)-Ebene gebrochen wird [40, 41]. Daher werden gelegentlich zwei der vier \bar{X} -Punkte entlang einer Richtung auch mit \bar{Y} gekennzeichnet.

Für topologisch nicht triviale Vertreter der (Sn,Pb)(Se,Te)-Materialklasse kommt es zu einer Inversion der Volumenbandstruktur am L -Punkt [18, 19, 21]. Da es acht äquivalente L -Punkte in der Brillouin-Zone gibt, kommt es also insgesamt zu einer geraden Anzahl an Bandinversionen. Für einen konventionellen topologischen Isolator folgt hieraus eine triviale Klassifizierung, was die Notwendigkeit einer eigenständigen topologischen Phase für solche Materialien unterstreicht. Die zwei Punkte L_1 und L_2 in Abb. 4.1 (b) unterscheiden sich lediglich in ihrer Impulskomponente senkrecht zur Oberfläche, werden jedoch auf den selben \bar{X} -Punkt projiziert. Der einfachste Ansatz für die Oberflächenzustände ist daher eine Überlagerung der beiden Dirac-Kegel, wie in Abb. 4.1 (c) gezeigt. Im Rahmen der $k \cdot p$ Theorie kann für diese Situation folgender Hamiltonoperator angegeben werden [42]:

$$H_{\bar{X}}(\mathbf{k}) = (v_x \cdot k_x \cdot s_y - v_y \cdot k_x \cdot s_x) \otimes I. \quad (4.2)$$

Der Index \bar{X} verdeutlicht, dass der Hamiltonoperator in der Umgebung des \bar{X} -Punkts entwickelt wurde, wobei I die Einheitsmatrix, s_i Paulimatrizen und \vec{v} die Fermigeschwindigkeit sind. Für die Komponenten des Wellenvektors \vec{k} gilt, dass k_x parallel zu $\bar{\Gamma X}$ und k_y parallel zu $\bar{X M}$ orientiert sind. Das Koordinatensystem ist also relativ zu dem aus Abb. 4.1 (b) um 45° gedreht,

sodass die Achsen mit den Projektionen der (110)- und (1 $\bar{1}$ 0)-Spiegelebenen zusammenfallen. Der entartete Zustand mit zwei überlagerten Dirac-Kegeln ist jedoch energetisch nicht stabil, da sich bei Streuprozessen zwischen den Punkten L_1 und L_2 lediglich der Impuls senkrecht zur Oberfläche ändert¹⁶. Solche Prozesse können in Gl. 4.2 durch einen Term der Form $m \cdot \tau_x$ berücksichtigt werden, wobei $\tau_x = \pm 1$ gilt. Infolge dieser Wechselwirkung wird ein Dirac-Kegel energetisch um m nach oben und der andere um m nach unten verschoben [siehe Abb. 4.1 (c)]. Während die Dirac-Kegel weiterhin im reziproken Raum bei \bar{X} lokalisiert sind, so ergibt sich für deren Schnittmenge nun eine Ellipse. Diese wird durch die Gleichung $m^2 = v_x^2 \cdot k_x^2 + v_y^2 \cdot k_y^2$ parametrisiert und umschließt \bar{X} in ihrem Mittelpunkt. Der \bar{X} -Punkt weist neben der TRS noch drei weitere Symmetrien auf: Reflexion bezüglich x , y und zweizählige Rotationssymmetrie. Mit letzteren beiden geht ein Austausch zwischen L_1 und L_2 einher, sodass sie in nullter Ordnung durch den τ_x -Term beschrieben werden. Für die Spiegelung bezüglich x bleiben die L -Punkte jedoch unverändert, sodass noch ein weiterer Term Gl. 4.2 hinzugefügt werden kann, der allen Symmetrien genügt [42]:

$$H_{\bar{X}}(\mathbf{k}) = (v_x \cdot k_x \cdot s_y - v_y \cdot k_y \cdot s_x) + m \cdot \tau_x + \delta \cdot s_x \cdot \tau_y. \quad (4.3)$$

Dieser δ -Term beschreibt die Hybridisierung zwischen den beiden Dirac-Kegeln, welche zu einer \vec{k} -abhängigen Bandlücke führt. Dass eine \vec{k} -Abhängigkeit vorliegt, lässt sich mittels der Eigenwerte der Spiegelsymmetrien M_x und M_y erklären. Entlang von $\overline{X\Gamma}$ haben die beteiligten Zustände die Eigenwerte i und $-i$ bezüglich M_x , während sie entlang von $\overline{X\bar{M}}$ identische Eigenwerte bezüglich M_y haben. Da die Hybridisierung zwischen Zuständen unterschiedlicher Eigenwerte verboten ist, sind die zwei Schnittpunkte der Spiegellinie mit der Ellipse topologisch geschützt, während sich an allen anderen Stellen des reziproken Raums eine Hybridisierungslücke öffnet. Wie in Abb. 4.1 (c) zu sehen ist, zeichnen sich die beiden Schnittpunkte (grüne Punkte) durch eine (näherungsweise) lineare, Dirac-artige Dispersion aus. Einer dieser zwei neuen Dirac-Punkte liegt jedoch nicht mehr in der ersten Brillouin-Zone. Ferner sind durch die Hybridisierung vier Sattelpunkte, je zwei energetisch ober und unterhalb der neuen Dirac-Punkte in der Bandstruktur entstanden [42, 43]. Diese Punkte werden später in Kap. 4.3 noch detailliert diskutiert. So sind z.B die beiden Sattelpunkte des Valenzbands in der 3D-Darstellung der Bandstruktur in Abb. 4.4(b) markiert.

¹⁶Eine genaue Betrachtung bezieht nicht nur die L -Punkte sondern auch die sie umgebenden Elektronen- bzw. Lochtaschen mit ein.

4.2 Topologische Phasenübergänge

Bereits kurz nach der Vorhersage von TCIs [18], wurden die Vertreter der (Sn,Pb)(Se,Te)-Materialklasse als mögliche Realisierung identifiziert und genauer untersucht [19–22]. Obwohl die Inversion der Bandstruktur am L -Punkt bereits um 1965 herum intensiv untersucht wurde [44–46], war eine der ersten Fragestellungen, welche binären Mischkristalle topologisch nicht trivial sind und wo der Phasenübergang hin zu trivialer Bandstruktur liegt. Für die Telluride wurde als erstes vorhergesagt, dass PbTe trivial und SnTe nicht trivial ist [19].

Verglichen mit konventionellen TIs ist diese Tatsache kontraintuitiv, da bei TIs oft die Spin-Bahn-Kopplung eine entscheidende Rolle für die Bandinversion einnimmt und diese für Pb ($Z = 82$) größer als für Sn ($Z = 50$) ist. Einerseits sind aber noch weitere Korrekturen zu berücksichtigen, wie der Darwin-Term und die relativistische Massenkorrektur. Andererseits weisen alle drei Effekte eine starke Abhängigkeit vom Bahndrehimpuls und damit auch vom Wellenvektor auf. Somit ergibt sich, dass im Gegensatz zu TIs, deren Bandinversion oft um den Γ -Punkt herum stattfindet, das leichtere Element eine invertierte Bandordnung am Rand der Brillouin-Zone begünstigt [45].

In Abb. 4.2 sind die experimentellen Ergebnisse zu den beiden Legierungen $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ und $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ zusammengefasst. Analysiert wurden Proben unterschiedlicher Sn-Konzentration x um so die Bandlücke kontinuierlich zu variieren. Ausgehend von z.B. PbSe nimmt die Bandlücke mit zunehmendem x ab und beginnt ab $x \approx 0.25$ sich wieder zu öffnen [21]. Der genaue Wert für x hängt wegen der thermischen Ausdehnung des Gitters auch von der Temperatur ab. Daher kann, wie durch die rote Linie angedeutet, auch ein temperaturinduzierter Phasenübergang auftreten. Sobald sich die Bandlücke wieder öffnet ist das topologisch nicht triviale Regime erreicht und die charakteristischen Dirac-Kegel bilden sich aus [siehe Abb. 4.2 (a)].

In winkelaufgelösten Photoelektronenspektroskopie-Messungen (ARPES) lässt sich um den Phasenübergang herum nicht nur das plötzliche Auftreten von Oberflächenzuständen beobachten, sondern auch die typische lineare Dispersion der Dirac-Kegel. Dies zeigt der Vergleich von zwei Schnitten entlang der $\overline{\Gamma X}$ Richtung aus Abb. 4.2 (b) & (c). Für eine triviale Sn-Konzentration sind nur die Beiträge des Volumenmaterials zu sehen, wobei das Valenzband bei einer Bindungsenergie von etwa 50 meV beginnt. Für ein nicht triviales x wird bei allen Photonenergien signifikant mehr Intensität im Bereich der Bandlücke gemessen [Abb. 4.2 (c)], wobei eine w-förmige Dispersion auffällt. Noch markanter als die lineare Dispersion der zwei Dirac-Punkte ist jedoch die Spinstruktur. Der Spin ist an den Impuls gekoppelt und senkrecht dazu orientiert. Durch gleichzeitiges Messen des Spins der Photoelektronen kann dieses Charakteristikum der Dirac-Kegel nachgewiesen werden. In Abb. 4.2 (d) wurde entlang der beiden gestrichelten Linien die Spinpolarisation überprüft und eingezeichnet. Während für die Oberflächenzustände ein

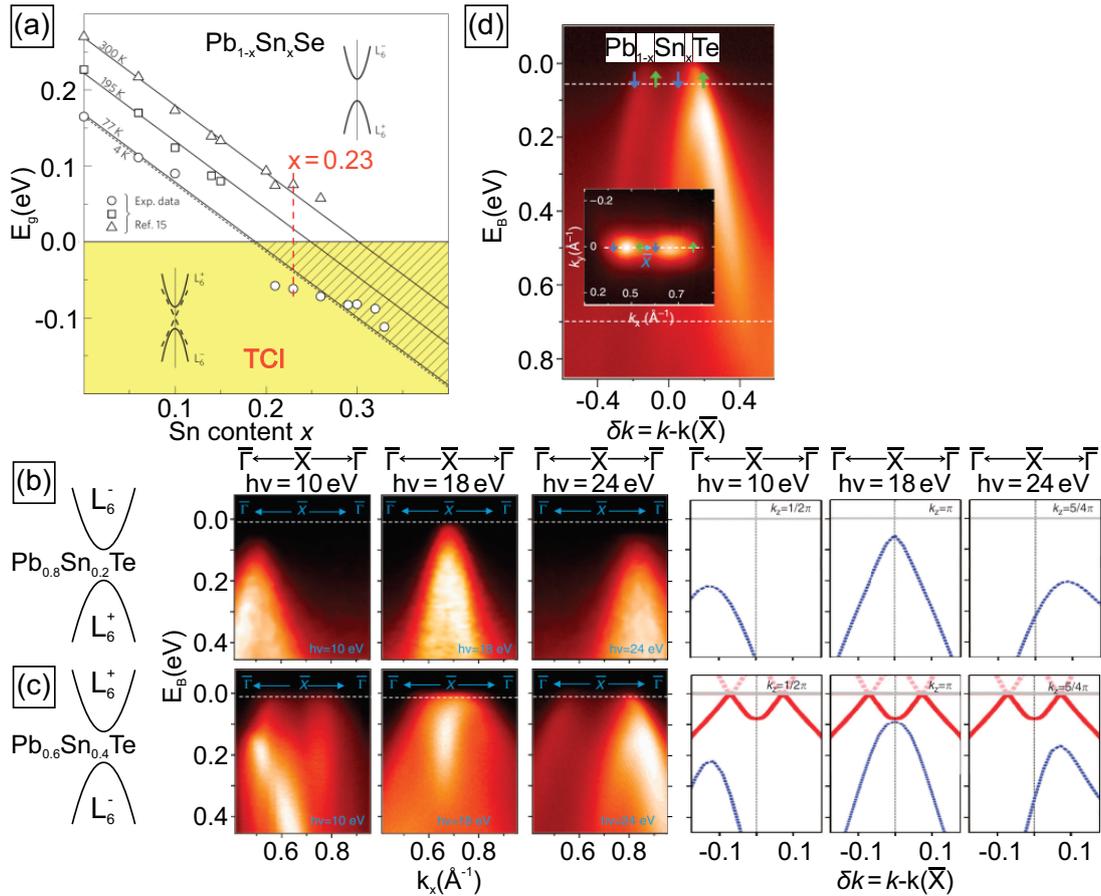


Abbildung 4.2 Topologischer Phasenübergang von $\text{Pb}_{1-x}\text{Sn}_x(\text{Se,Te})$: (a) Experimentell bestimmte Bandlücke von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ in Abhängigkeit von x und der Temperatur. Ab einer kritischen Konzentration von $x \approx 0.25$ öffnet sich die Bandlücke mit invertierter Ordnung. (b) ARPES-Schnitt von trivialem $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ bei verschiedenen Energien und berechnete Bandstruktur (blaue Linien). (c) Analoges Datensatz zu (b) für eine nicht triviale Probe, wobei die auftretenden zusätzlichen Oberflächenzustände rot in der Berechnung dargestellt sind. (d) Pfeile markieren zusätzlich die gemessene Spinpolarisation entlang der gestrichelten Linien. Im Volumen ist keine Polarisation messbar. Adaptiert von [21] (a) & [20] (b-d).

alternierendes Spinsignal beobachtet wurde, konnte tief im Valenzband keine effektive Spinpolarisation gemessen werden [20].

Ein weiterer topologischer Phasenübergang kann auftreten, wenn die Schichtdicke des Materials soweit reduziert wird, dass die Oberflächenzustände von Ober- und Unterseite miteinander wechselwirken. Durch die Kristallstruktur ergibt sich jedoch für die $(\text{Sn,Pb})(\text{Se,Te})$ -Materialklasse eine Besonderheit, da bei einer endlichen Anzahl N von Atomlagen deren Parität die zugrundeliegende Symmetrie beeinflusst. Gezeigt ist das in Abb. 4.3 (a), wo exemplarisch einmal für $N = 4$ und $N = 3$ die Spiegelebenen eingezeichnet sind. Da für ungerade N die Spiegelebene mit einer Atomlage zusammenfällt, wird jedes Untergitter in sich selbst gespiegelt. Bei geradem N liegt die Ebene zwischen zwei Atomlagen, sodass die beiden Untergitter ineinander gespiegelt werden. Erst durch eine zusätzliche Verschiebung bildet eine solche Gleitsymmetrie jedes

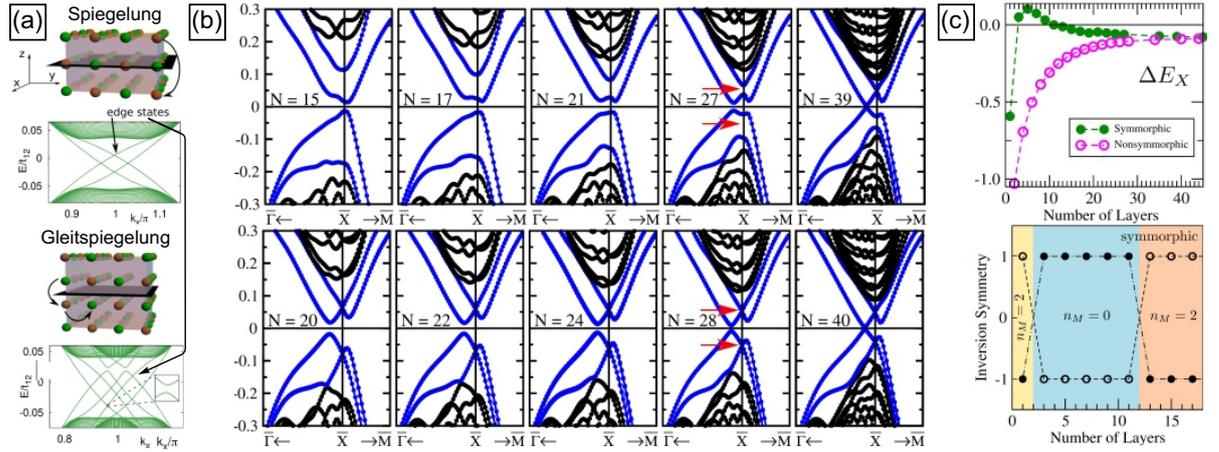


Abbildung 4.3 Topologischer Phasenübergang durch Gleitspiegelsymmetrie: (a) Strukturmodell und Spektrum der Kanten von SnTe für einen Film bestehend aus drei und vier Atomlagen N . Für ungerade N liegt eine Spiegel-, für gerade N eine Gleitspiegelsymmetrie vor. (b) Bandstruktur für verschiedene N . Die Pfeile markieren zwei Entartungen die für gerade N durch die Gleitspiegelung geschützt sind. (c) Bandlücke an \bar{X} (oben). DFT Eigenwerte des Inversionssymmetrieoperators für die höchsten besetzten und niedrigsten unbesetzten Zustände für ungerade N (unten). Adaptiert von [47] (a) & [35] (b-d).

Untergitter in sich selbst ab. Die Kombination aus einer Verschiebung um einen nicht primitiven Gittervektor (den Bruchteil eines Gittervektors, hier $\frac{1}{2}$) und einer Symmetriegruppe¹⁷ wird als nicht symmorph bezeichnet. Bei symmorphen Symmetrieoperationen fehlt hingegen die Verschiebung, sodass bei diesen (mindestens) ein Punkt unverändert bleibt.

Einige Beispiele für die physikalischen Konsequenzen aus diesem Symmetriewechsel sind in Abb. 4.3 dargestellt. So ändert sich beispielsweise die Anzahl der Kantenzustände in (a) von zwei auf vier Paare. Eine weitere Konsequenz zeigt sich, ebenfalls in der Bandstruktur, wenn man nicht symmorphen Symmetrien mit ihren symmorphen Partnern (d.h. der selben Symmetriegruppe ohne die Verschiebung) vergleicht. Wie in (b) durch die roten Pfeile markiert, existieren für den nicht symmorphen Fall ($N = 28$) zusätzliche Entartungen. Auch bei weiterer Reduktion von N bleibt die Entartung für alle geraden N erhalten, während sie für ungerade N bereits zwischen $N = 27$ und $N = 39$ verschwindet. Diese symmetriegeschützten Entartungen führen im hier vorliegenden Fall von SnTe sogar dazu, dass es im Bereich einzelner Atomlagen zwei Phasenübergänge zwischen topologischen und trivialem Zustand gibt. Diese durch die nicht symmorphen Symmetrie induzierte Zwangsbedingung führt dazu, dass die Bandlücke stets negativ, also invertiert und SnTe damit topologisch bleibt [rosa Datenpunkte in Abb. 4.3 (c)]. Für ungerade N hingegen wird die Entartung durch die zunehmende Wechselwirkung der Dirac-Kegel von Ober- und Unterseite aufgehoben. Das führt zu einer positiven Bandlücke und damit einer Umkehr der Invertierung bei SnTe-Filmen mit weniger als zwölf Lagen. Für die Monolage kommt es dann, aufgrund des Fehlens von Ober- und Unterseite, erneut zu einem Phasenübergang.

¹⁷ Auch Raumgruppe genannt. Def.: Menge aller Transformationen die ein Objekt unverändert lassen.

Veranschaulicht werden die zwei Phasenübergänge auch in Abb. 4.3 (c) im unteren Graphen, wo die Eigenwerte des Inversionssymmetrieoperators für die energetisch höchsten, besetzten und niedrigsten, unbesetzten Zustände gezeigt sind. Die Eigenwerte sind mit dem orbitalen Charakter der Bänder verknüpft und vertauschen daher ihre Ordnung falls es zu einer Bandinversion kommt. Die Dichtefunktionaltheorie (Density Functional Theory, DFT) ergibt nun, dass sich genau bei den selben Schichtdicken die Eigenwerte vertauschen. Die Bereiche mit unterschiedlicher Spiegel–Chern–Zahl n_M sind farblich hervorgehoben und repräsentieren für $n_M = 0$ einen trivialen und für $n_M = 2$ einen topologischen Zustand.

Im Allgemeinen führt das Hinzufügen weiterer Symmetrien zu einer weiteren Aufspaltung der AZ–Klassen in neue Unterklassen, wie in Kap. 2 gezeigt wurde. Daher ist es nicht überraschend, dass es für nicht symmorphie Symmetrien, im Vergleich zu ihren symmorphen Partnern, zu einer weiteren Differenzierung kommt. Interessanterweise kann es aber dazu kommen, dass symmorphie und nicht symmorphie Symmetrie nicht länger durch die selbe topologische Invariante klassifiziert werden, was bedeutet, dass sie unterschiedlichen topologischen Phasen angehören können [33, 34]¹⁸. Für $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ stellten Ozawa *et al.* fest, dass für eine gerade Anzahl von Atomlagen n_M immer verschwinden muss [48]. Hieraus folgt die Notwendigkeit für die nicht symmorphie Symmetrie eine andere topologische Invariante zu finden [47].

4.3 Kantenzustände in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$

Bevor die Modifikationen der STS–Spektren durch Stufenkanten auf der (001)–Oberfläche diskutiert werden, soll zunächst die prinzipiell zu erwartende Form des Spektrum anhand der in diesem Kapitel vorgestellten Bandstruktur gezeigt werden. Hierfür sind in Abb. 4.4 (a) & (b) mehrere charakteristische Punkte eingezeichnet: \bar{X} mit den zwei energetisch aufgespaltenen Dirac–Kegeln bei $E_D^{+,-}$, ihre symmetriegeschützten Schnittpunkte auf der $\bar{\Gamma X}$ Linie bei $\bar{\Lambda}_{1,2}$ und die Sattelpunkte bei $\bar{S}_{1,2}$. Letztere sollen nun genauer analysiert werden, da sich dort die Topologie der Fermifläche ändert.

In Abb. 4.4 (a) sind Isoenergiekonturen für verschiedene Energien E oberhalb der neuen Dirac–Punkte–Energie E_{D^*} gezeigt. Sie repräsentieren Schnittbilder der Fermifläche bei den jeweiligen Energien und wurden genauso wie die Bandstruktur in (b) gemäß Gl. 4.3 berechnet. Für hinreichend kleine E ist die Isoenergiekontur durch zwei getrennte Kreise gegeben, die jeweils einen

¹⁸Zur Vollständigen Betrachtung gehört neben den Symmetrien auch die Dimension d des physikalischen Systems. In Abb. 4.3 (c) ändert sich für den symmorphen Fall die topologische Invariante für den Übergang von $d = 2$ zu $d = 3$. Für den nicht symmorphen Fall ändert sich die Invariante nicht zwischen $d = 2$ und $d = 3$. Daher müssen zwei Bedingungen für eine triviale Phase erfüllt sein: symmorphie Symmetrie und ein zweidimensionales System. Somit lassen sich zwei Phasengrenzen finden, einmal eine mit konstanter (symmorpher) Symmetrie und als Funktion von d bzw. N und einmal mit konstantem d und in Abhängigkeit von der Symmetrie.

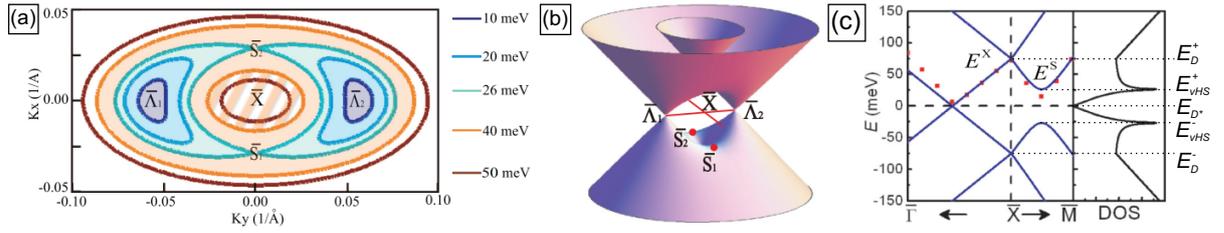


Abbildung 4.4 Topology der Fermifläche, Bandstruktur und STS-Spektrum von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$: (a) Linien konstanter Energie für verschiedene Werte oberhalb der neuen Dirac-Punkte. Schattierte Bereiche sind Elektronen-, schraffierte Lochtaschen. Bei der Sattelpunktsenergie E_S ändert sich die Topologie der Fermifläche, da aus zwei disjunkten Taschen ohne Loch eine zusammenhängende Tasche mit Loch wird. (b) Bandstruktur gemäß Gl. 4.3 mit charakteristischen Punkten: Neue Dirac-Punkte bei $\bar{\Lambda}_{1,2}$ entlang von $\bar{\Gamma}\bar{X}$; Sattelpunkte $\bar{S}_{1,2}$. (c) Schematisches Bild der Bandstruktur (links) und der Zustandsdichte (rechts) wie sie in einem STS-Spektrum abgebildet wird. Markante Punkte sind die Van Hove Singularitäten an den Sattel- und die Minima an den Dirac-Punkten. Adaptiert von [42].

$\bar{\Lambda}_{1,2}$ -Punkt umschließen. Mit zunehmender Energie ändert sich die Kontur hin zu einer deformierten Ellipse. Zunächst wird die Ellipse in Richtung \bar{X} gestaucht (lila Linie), ehe sich die konvexe in eine konkave Krümmung ändert (dunkelblau). Dieses Verhalten wird durch die Hybridisierung hervorgerufen um ein Kreuzen der Zustände zu vermeiden und durch den δ -Term in Gl. 4.3 beschrieben. Die Pseudoellipsen sind außerdem nicht länger um die $\bar{\Lambda}_{1,2}$ -Punkte zentriert, weil die jeweiligen Mittelpunkte ebenfalls in \bar{X} -Richtung schieben. Mit dem Erreichen der Sattelpunktsenergie bei $E_S = \delta$ berühren sich die beiden Konturen an den beiden Punkten $\bar{S}_{1,2}$ (hellblau), wobei \bar{X} nicht von den dazugehörigen Elektronentaschen (schattierter Bereich) eingeschlossen wird. Für noch größere Energien verschmelzen die beiden Taschen zu einer großen, um \bar{X} zentrierten Ellipse (orange). Außerdem entsteht eine zweite, kleinere Ellipse, welche ebenfalls um \bar{X} zentriert ist (schraffierter Bereich). Dabei handelt es sich um eine Lochtasche, die von dem unteren Teil des energetisch nach oben verschobenen Dirac-Kegels abstammt. Mit zunehmender Energie $E_S < E < E_D^+$ schrumpft daher die Lochtasche, während sich die Elektronentasche weiter ausdehnt. Für $E < E_{D^*}$ sehen die Konturen identisch aus, allerdings sind die Rollen von Loch- bzw. Elektronentaschen vertauscht.

Der Übergang von zwei getrennten Elektronentaschen hin zu einem geschlossenen Gebiet mit Loch(-tasche) darin entspricht einem topologischen Phasenübergang. Analog zu dem in Kap. 2 genannten Beispiel von Würfel und Donut ist es nicht möglich mittels einer stetigen Transformation erneut eine lochfreie Fermioberfläche zu erhalten. Diese Art von Phasenübergang wird als Lifshitz-Übergang bezeichnet. Im vorliegenden Fall ist er eng mit den Sattelpunkten in der Bandstruktur verknüpft. Da an diesen Punkten die Steigung verschwindet, besitzen die benachbarten Wellenvektoren in sehr guter Näherung die gleiche Energie. Das führt zu einer Singularität in der Zustandsdichte bei dieser Energie, die als Van Hove Singularität (Van Hove Singularity, vHS) bezeichnet wird [19, 42, 49].

Ein schematisches Bild der zu erwartenden Zustandsdichte in Abhängigkeit der Energie ist in

Abb. 4.4 (c) gegeben. Da die Oberflächenzustände energetisch in der Bandlücke des Volumenmaterials liegen, sind die Beiträge von Valenz- und Leitungsband hier minimal. Dadurch ergibt sich, dass die Energie E_{D^*} der neuen Dirac-Punkte mit dem globalen Minimum der DOS zusammenfällt, weil hier sowohl Volumen- als auch Oberflächenzustände ihr Minimum haben. Ausgehend von dieser Energie, die in der Rechnung dem Ferminiveau gleichgesetzt wurde, wächst die DOS zunächst linear an, bevor sie in der Nähe der vHS immer stärker ansteigt. Die vHS bildet dann ein lokales Maximum, von dem aus die DOS zunächst mit betragsmäßig wachsender Energie wieder abnimmt. Ursächlich hierfür ist die Annäherung an den oberen bzw. unteren Dirac-Punkt, welche somit lokale Minima in der DOS markieren. Sobald die Energie einen Dirac-Punkt passiert hat, steigt die DOS kontinuierlich an, zunächst langsam und dann immer stärker, sobald Valenz- und Leitungsband einsetzen.

Die wichtigsten Modifikationen, die an Kanten der (001)-Oberfläche beobachtet wurden, sind in Abb. 4.5 gezeigt. Neben den bisher diskutierten, topologischen 2D-Oberflächenzuständen entdeckten Sessi *et al.* zusätzlich noch topologisch geschützte 1D-Kantenzustände [39].

Der nicht triviale Ursprung wird aus (a) ersichtlich, wo in der linken Spalte die Topographie für drei verschiedene Sn-Konzentrationen x gezeigt sind. Die oberste Probe ist reines PbSe und somit trivial, die mittlere besitzt ein x nahe des Phasenübergangs und die unterste Probe ist mit $x \approx 0,33$ sicher topologisch nicht trivial. Für alle Proben sind zwei Stufenkanten zu sehen, eine Doppel- und eine Einfachstufe. Die rechte Spalte zeigt die simultan zur Topographie gemessene differentielle Leitfähigkeit bei der Energie des Dirac-Punkts. Für die triviale Probe ergibt sich eine weitestgehend homogene verteilte Leitfähigkeit, wobei es an den Stufenkanten wegen der endlichen Geschwindigkeit der Regelschleife zu einem schwach ausgeprägtem Bildartefakt kommt. Dieses Bild ändert sich drastisch sobald die Probe topologisch wird. Bereits für $x \approx 0,24$ wird die Probe, zumindest bei tiefen Temperaturen, topologisch. In der dI/dV -Karte ist deutlich eine erhöhte Intensität an der rechten Stufenkante zu sehen. Für die unterste Probe wird dieser Peak noch deutlicher, was insbesondere am Signal-zu-Rasch-Verhältnis im unterem Profil entlang der grauen Linie ersichtlich ist. Die Abwesenheit des Kantenzustands für PbSe zeigt den topologischen Charakter der Kantenzustände.

Um nun zu verstehen warum nur manche Stufenkanten einen Kantenzustand haben, ist in Abb. 4.5 (b) ein größerer Bildausschnitt gezeigt. Das abgebildete Spektrum (5) wurde auf einer Terasse aufgenommen und hat im wesentlichen die zuvor diskutierte Form. Nur der \bar{X} -Dirac-Punkt bei $E_{\bar{D}}$ ist wegen der Nähe zum Valenzband nicht eindeutig als Minimum zu erkennen. Da eine Einfachstufe eine Höhe von etwa 300 pm besitzt, kann dem Profil entnommen werden, dass bei 1 & 3 eine Einfach-, bei 2 eine Doppel- und bei 4 eine Vierfachstufe vorliegt. Anhand von Abb. 4.1 (a) wird ersichtlich, dass bei einer ungeraden Anzahl von fehlenden Atomlagen rechts und links von der Stufe Atome des selben Untergitters nebeneinander sitzen, während bei eine geraden

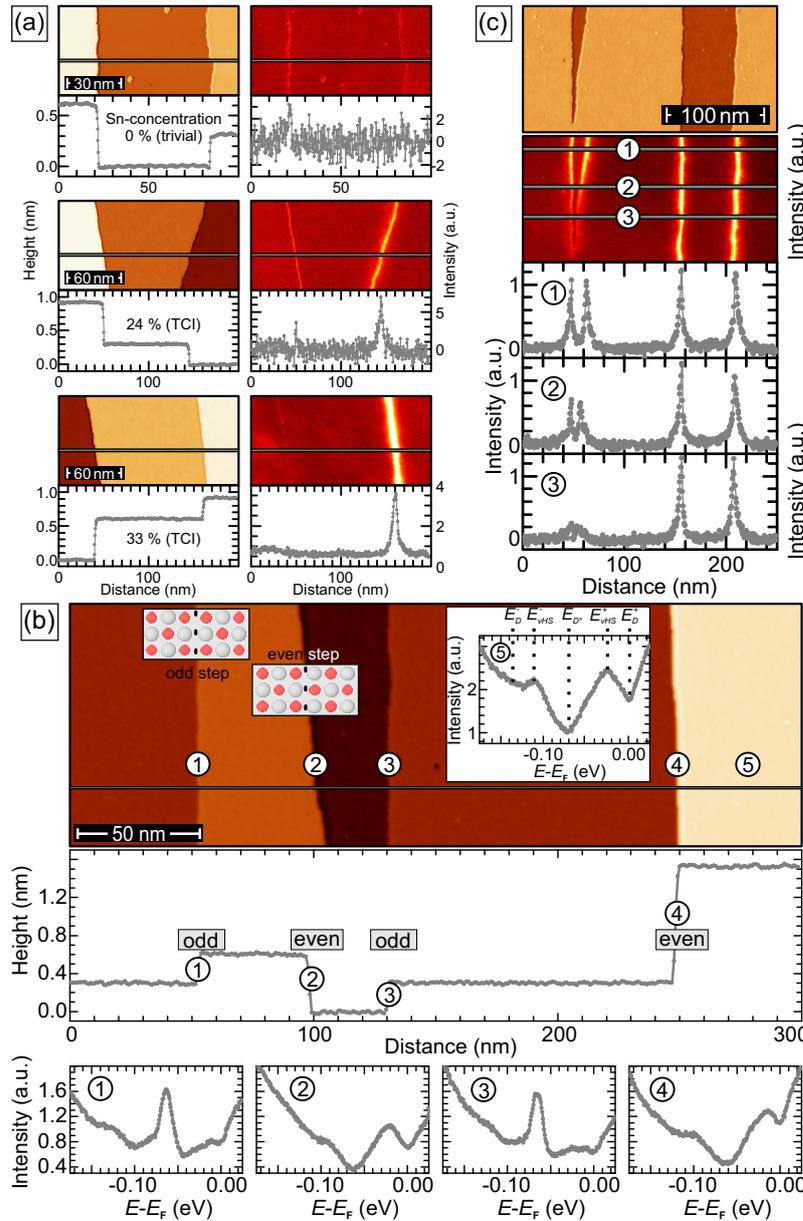


Abbildung 4.5 Kantenzustände von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$: (a) Topographie (links) und dazugehörige dI/dV -Karte bei $E = E_{D^*}$ (rechts), sowie Linienprofile entlang der grauen Balken. Gezeigt sind stets eine Einfach- und eine Doppelstufe für verschiedene Sn-Konzentrationen: trivial, nur bei hinreichend hohen Temperaturen trivial, topologisch. (b) Stufen verschiedener Höhe und dazugehörige Spektren einer topologischen Probe. Für Stufen mit einer ungeraden Anzahl an fehlenden Atomlagen (1 & 3) ist ein Peak bei $E = E_{D^*}$ zu sehen, falls die Probe topologisch ist. Gerade Stufen (2 & 4) unterscheiden sich spektroskopisch nicht von der Terrasse (5). Die Einschubbilder zeigen eine schematische Draufsicht auf ungerade (gerade) Stufen, wo wegen der Kristallstruktur die translationssymmetrie (nicht) gebrochen ist. (c) Topographie und dazugehörige dI/dV -Karte bei $E = E_{D^*}$, sowie Linienprofile entlang der grauen Balken für parallele und zusammenlaufende Kantenzustände. Sobald der Stufenabstand vergleichbar zur Ausdehnung der Kantenzustände (senkrecht zur Stufe) wird, nimmt die DOS mit kleiner werdendem Abstand stark ab. Adaptiert von [39].

Anzahl eine ganze Einheitszelle fehlt. In der Folge ist für den geraden Fall die Translationssymmetrie an der Stufenkante gegeben und für den ungeraden Fall gebrochen. Dies wird durch die modellhafte Darstellung von Atomen in den beiden kleinen Einschubbildern verdeutlicht.

Der Einfluss auf die elektronische Struktur ist in den Spektren 1-4 zu sehen. Während die ungeraden Stufenkanten alle durch einen Peak am Dirac-Punkt E_{D^*} gekennzeichnet sind¹⁹, kann bei den geraden Stufenkanten keine signifikante Änderung an der Zustandsdichte beobachtet

¹⁹ Tatsächlich tritt oft eine kleine, systematische Verschiebung zwischen Dirac-Punkt und globalem Minimum der DOS auf. Ursächlich hierfür können triviale Zustände, die mit den topologischen koexistieren oder Variationen des chemischen Potentials sein. Da die Sn-Atome zufällig räumlich verteilt sind und die Spektren für Terrasse und Stufenkante an unterschiedlichen Stellen gemessen werden, kann es zu Fluktuationen von einigen meV kommen. Eine ähnliche Beobachtung wurde auch bei Messungen im Magnetfeld des TI Sb_2Te_3 gemacht [50].

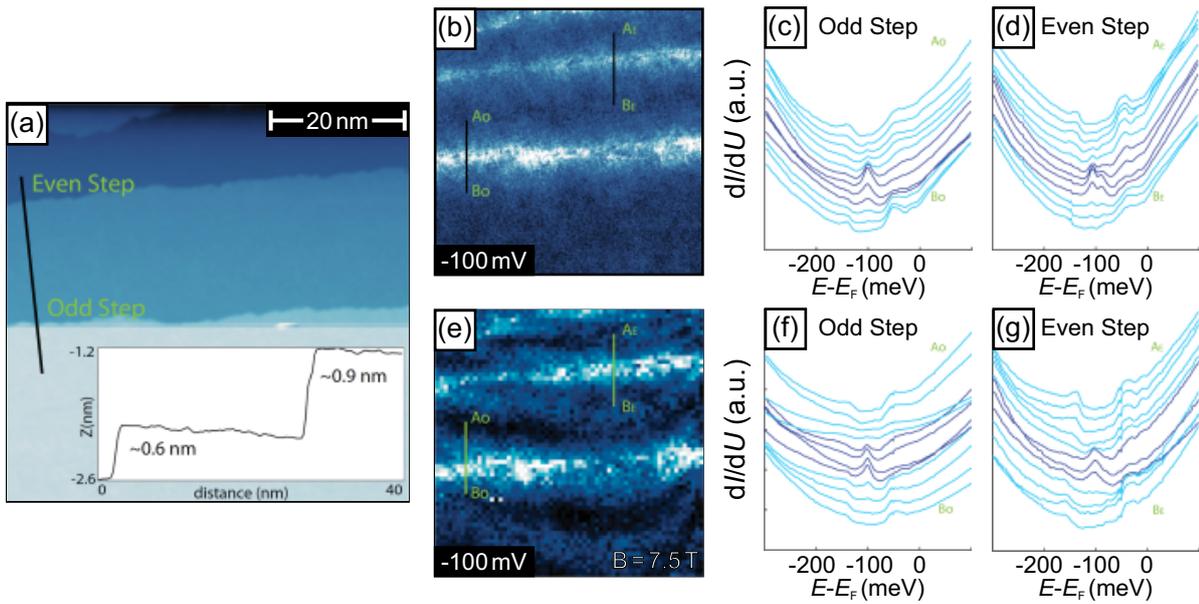


Abbildung 4.6 1D Kantenzustände an ungeraden Stufen: (a) Topographie der (001)-Oberfläche und Linienprofil über eine Doppel- und Dreifachstufe aus Ref.[51]. (b) Dazugehörige dI/dV -Karte bei $E = E_{D^*}$. (c,d) Wasserfalldiagramm der STS-Spektren entlang der in (b) eingezeichneten Linien. (e-g) Äquivalenter Datensatz zu (b-d), gemessen in einem Magnetfeld $B = 7,5$ T. Im Gegensatz zu Ref. [39] wurde auch an einigen geraden Stufen Zustände in Form eines Doppelpeak gefunden, der bei Anwesenheit eines Magnetfeldes zu einem Einfachpeak wird. Adaptiert von [51].

werden. Zum Ende dieses Kapitels werden verschiedene Erklärungsansätze präsentiert, die gegenwärtig Teil der wissenschaftlichen Diskussion sind.

Zunächst sollen noch einige weitere Eigenschaften der Kantenzustände diskutiert werden. Gemäß der Theorie topologisch kristalliner Isolatoren sollten diese Kantenzustände robust gegen Störungen sein, welche nicht die fundamentalen (Kristall-) Symmetrien brechen. Hierfür wurden Messungen bei erhöhten Temperaturen (80 K statt 5 K) und in Magnetfeldern (bis zu 11 T) senkrecht zur Oberfläche durchgeführt. Die Kantenzustände konnten in beiden Fällen nachgewiesen werden [39]. Andererseits zeigt Abb. 4.5 (c) vier ungerade Stufen, wobei die beiden linken Stufen sich im unteren Teil des Bildes annähern. Unterhalb der Topographie sind die dI/dV -Karte bei $E = E_{D^*}$, sowie dazugehörige Linienprofile gezeigt. Als räumliche Ausdehnung senkrecht zu einer Stufenkante wurden eine Länge von etwa 10 nm festgestellt, wie anhand der rechten beiden Stufen in den Linienprofilen ersichtlich wird. Sobald der Abstand der beiden linken Stufen einen vergleichbaren Wert erreicht, nimmt die Intensität stark ab. Je kleiner dabei der Abstand wird, umso stärker wird die DOS reduziert, bis schließlich gar keine zusätzlichen Zustände am Dirac-Punkt lokalisiert sind. Dieser Übergang wird in Kap. 5.3 detailliert untersucht.

Als Einstieg in die Diskussion zu dem topologischen Ursprung der Kantenzustände sind in Abb. 4.6 die experimentellen Ergebnisse von Iaiya *et al.* zusammengefasst [51]. In (a) ist die Topographie und ein Linienprofil gezeigt, woraus hervorgeht, dass die obere Stufe gerade und die untere ungerade ist. Daneben sind zwei äquivalente Datensätze gezeigt, einmal ohne (b-d) und einmal

mit einem Magnetfeld von $B = 7,5 \text{ T}$ (e-g). Die dI/dV -Karten bei $E = E_{D^*} = 100 \text{ meV}$ wurden an der selben Stelle wie die Topographie aus (a) aufgenommen. Gut erkennbar ist in beiden Fällen die erhöhte Intensität an der ungeraden Stufenkante, wie auch deren Fluktuation entlang der Kante. Die ungerade Kante zeigt nur eine geringe Intensitätserhöhung bei Nullfeld, wie sie beispielsweise durch Bildartefakte an Stufenkanten auch zu erwarten ist. An einigen Stellen nimmt die Intensität jedoch im Magnetfeld signifikant zu, insbesondere im rechten Bildteil. Ergänzend wurden daher Tunnelspektren entlang der eingezeichneten Linien aufgenommen (c,d,f,g). Für ungerade Stufen ergibt sich das bisher bekannte Bild: ein Kantenzustand (c), welcher robust gegen das Magnetfeld senkrecht zur Probenoberfläche ist (f). Für gerade Stufen ergibt sich jedoch ein differenzierteres Bild: An einigen Stellen kann ein Doppelpeak beobachtet werden (d), der unter Feldeinfluss in einen einzelnen Peak übergeht (g).

Zunächst analysieren Iaia *et al.* theoretisch eine ungerade Stufe und stellen fest, dass diese durch ein am Dirac-Punkt gepinntes Kramers-Paar²⁰ beschrieben wird. Der einfachste Ansatz dann eine gerade Stufe qualitativ zu beschreiben stellt die Überlagerung zweier ungerader Stufen dar. Weil zu jedem Kramers-Paar zwei Zustände gehören, führt dieses Modell zu vier Zuständen an einer geraden Stufe [vgl. Abb. 4.7 (a)]. Während die Zustände des selben Kramers-Paars nicht miteinander hybridisieren können, ist dies zwischen zwei Paaren durchaus möglich. In der Folge spalten die vier Zustände in zwei Paare auf, die somit zu einem Doppelpeak in der DOS führen. Die energetische Aufspaltung hängt dabei von den unbekanntem, mikroskopischen Details ab und variiert räumlich stark. Wenn die Aufspaltung groß genug wird, überlagern sich die Kramers-Zustände mit dem Valenz- und Leitungsband und die dazugehörigen Peaks sind gar nicht mehr beobachtbar. In Anwesenheit eines Magnetfeldes spalten die hybridisierten Paare in vier einzelne Zustände auf, wobei aus jedem Paar jeweils ein Zustand in Richtung des Valenz- und der andere zum Leitungsband hin verschoben wird. Dadurch wandern zunächst zwei der vier Zustände wieder in Richtung des Dirac-Punkts, sodass aus dem Doppel- ein Einfachpeak wird (die anderen beiden Peaks verschwinden im Valenz- bzw. Leitungsband).

Um den topologischen Ursprung zu verstehen betrachten Iaia *et al.* einen $k \cdot p$ -Hamiltonian, welcher ebenso wie der aus Gl. 4.3, auf Ref. [19] basiert. Aus Fits von ARPES-Daten können die Parameter für diesen Modell-Hamiltonian bestimmt werden [52]. Wird einer der Parameter mit Null genähert, so weist der Hamiltonian einer [010]-Stufe eine zusätzliche Teilchen-Loch-Symmetrie auf. Außerdem besitzt dieser Stufen-Hamiltonian zwei verschiedene m -Terme m und m' , welche jeweils mit einem Untergitter assoziiert sind [51].

In Abb. 4.7(b) sind die Draufsicht auf eine ungerade Stufe und das damit verbundene räumliche

²⁰ Als Kramers-Paar werden Zustände bezeichnet, die durch das Kramers-Theorem miteinander verknüpft sind. Das Theorem sagt für Systeme mit halbzahligen Spin und TRS eine geradzahlige Entartung voraus. Die Kramers-Partner besitzen deshalb jeweils entgegengesetzte Wellenvektoren und Spins, teilen sich jedoch den selben Energieeigenwert.

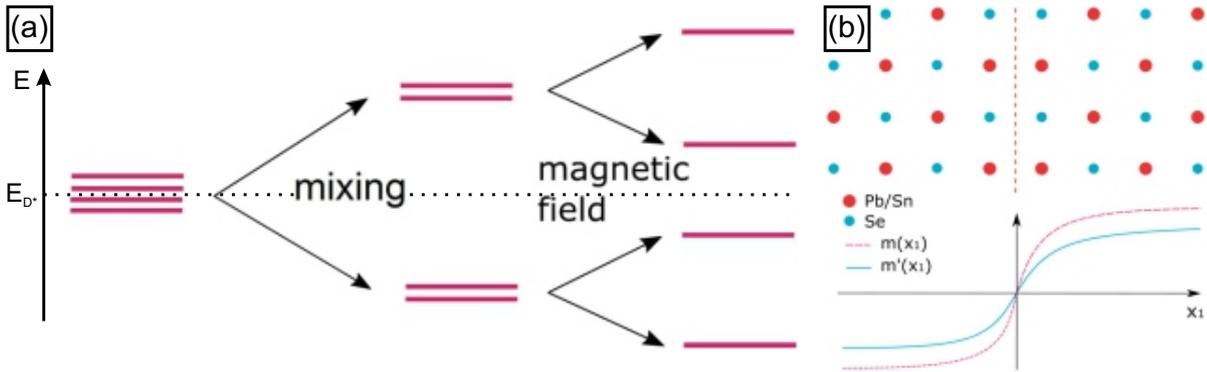


Abbildung 4.7 Iaia-Modell für Kantenzustände in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$: (a) Kramer's-Paare von einer geraden Stufe. Eine ungerade Stufe besitzt ein Kramer's-Paar, das aus zwei Zuständen besteht. Eine gerade Stufe entspricht der Überlagerung zweier ungerade und besitzt daher zwei Kramer's-Paare. Hybridisierung innerhalb eines Paares ist verboten, zwischen Paaren jedoch führt sie zu einer Aufspaltung in zwei Niveaus. Durch ein Magnetfeld spalten diese weiter auf, sodass am Dirac-Punkt aus einem Doppel- ein Einfachpeak wird. (b) Draufsicht auf eine ungerade Stufe und Profil der Masseterme des dazugehörigen Hamiltonian. Jedes Untergitter besitzt einen eigenen Masseterm m bzw. m' . Analog zu TI führt ein Vorzeichenwechsel zur Existenz topologisch geschützter Zustände. Adaptiert von [51].

Profil der m -Terme skizziert. An der Stufe kommt es zu einer Inversion von m und m' sowie zu einem allgemeinen Vorzeichenwechsel. Analog zu einem TI führt eine solche Domänenwand zur Existenz einer geschützten Mode, in diesem Fall in Form eines Kramer's-Paares von Jack-Rebbi Solitons. Durch die PHS müssen die beiden Moden bei $E = -E = 0 = E_{D^*}$ liegen, sodass sie am Dirac-Punkt lokalisiert sind.

Neben der gezeigten Konfiguration existiert noch eine zweite Art von ungerader Stufe, bei der der m -Term auf der linken Seite größer und rechts kleiner Null ist. Für eine weitere ungerade Stufe, die rechts in Abb. 4.7(b) eingefügt wird, tritt so ein Fall ein. Die beiden Stufen unterscheiden sich jedoch im Vorzeichen ihres topologischen Index ν , das einmal positiv und das andere mal negativ ist. Für einen verschwindend kleinen Abstand der Stufen überlagern sie sich und bilden eine gerade Stufe. Dabei addieren sich die topologischen Indizes, sodass $\nu = 2 + (-2) = 0$ folgt. Dementsprechend sind die Zustände an geraden Stufen trivial und können durch Wechselwirkungen aus der Bandlücke geschoben werden.

Der im Vergleich zu TIs unterschiedliche Betrag von $|\nu| = 2$ folgt aus der topologische Klassifizierung der Kantenzustände. Diese ergibt sich, wenn man neben der genäherten PHS auch die TRS und den eindimensionalen Charakter der Kantenzustände berücksichtigt. Die Kombination aus PHS und TRS führt zu einer chiralen Symmetrie und damit zu einer neuen Altland-Zirnbauer-Klasse: AIII. Für eindimensionale System wird diese Klasse durch eine gerade, ganzzahlige topologische Invariante $\nu \in 2\mathbb{Z}$ beschrieben. Der Betrag der ganzen Zahl \mathbb{Z} entspricht dabei der Anzahl an Kramer's-Paaren und ist für $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ gleich eins [51].

Die Analyse der Richtungsabhängigkeit der Kantenzustände von Rechciński *et al.* basiert ebenfalls auf dem Ansatz einer 1D topologischen Invarianten [53]. Da im Vergleich zu Iaia *et al.* das

Modell nur die Bandstruktur in der Umgebung eines \bar{X} -Punkts abbildet, muss für den zweiten eine separate Rechnung durchgeführt werden. Dadurch sind die beiden Punkte unabhängig voneinander und besitzen ihre jeweils eigene topologische Invariante. Ein weiterer Unterschied besteht auch darin, dass die von einer ungeraden Stufe getrennten Terrassen durch zwei eigenständige Hamiltonians H^A und H^B beschrieben werden. Diese sind jedoch nicht unabhängig voneinander, sondern können analog zu Ref. [51] durch Vertauschen der zu den jeweiligen Untergittern dazugehörigen Orbitalen bzw. Zuständen ineinander überführt werden. Die wichtigste Konsequenz daraus ist der bereits erwähnte Vorzeichenwechsel des m -Terms.

Rechciński *et al.* ziehen die beiden verschiedenen Hamiltonianoperatoren heran, um für die Terrassen unterschiedliche Terminierungen zu definieren. Stringenter ist jedoch von Phasen im Sinne eines zusammenhängendes Gebiet mit gleichen physikalischen Eigenschaften (hier insbesondere elektronischen) zu sprechen. Die gebrochene Translationssymmetrie an ungeraden Stufen wird durch eine strukturelle 180° -Phasenverschiebung bedingt, welche sich eins zu eins auch in der elektronischen Struktur und damit den elektronischen Eigenschaften niederschlägt. Im weiteren Verlauf dieser Arbeit wird daher die Terminologie „strukturelle Phase“ verwendet.

Um mit Hilfe dieser Hamiltonians ungerade Stufenkanten topologisch klassifizieren zu können, spielt die chirale Symmetrie eine Schlüsselrolle. Aufgrund der vorhandenen TRS kann diese nur in Kombination mit einer PHS des Hamiltonians realisiert werden. Prinzipiell erlauben die fundamentalen Symmetrien der (001)-Oberfläche von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ (TRS, M_y , M_x)²¹ mehr Terme für den Hamiltonian als Rechciński *et al.* in ihrem $k \cdot p$ -Modell berücksichtigen, insbesondere auch PHS-brechende. Ein Vergleich mit den Oberflächenzuständen aus Tight-Binding-Berechnungen rechtfertigt jedoch die Beschränkung auf vier Terme, welche alle der PHS genügen.

Unter diesen Annahmen bzw. Näherungen kann für H^A und H^B jeder geschlossene Pfad im zweidimensionalen k -Raum der Oberfläche durch eine Windungszahl ν^A bzw. ν^B , eine eindimensionale topologische Invariante, beschrieben werden [54, 55]. Wenn der Hamiltonian periodisch in k ist, dann gibt es zu jedem geschlossenen Pfad auch eine gerade Linie, welche topologisch äquivalent ist und daher durch die selbe Windungszahl beschrieben werden kann. Eine solche Linie wird durch einen Satz an Wellenvektoren definiert, die der Zwangsbedingung $\vec{k} \cdot \hat{e}_\parallel = \text{konst.}$ genügen²². In Abb. 4.8 (a) sind Beispiele für verschiedene Einheitsvektoren \hat{e}_\parallel und Werte der Konstanten gezeigt. Das Ergebnis von ν hängt davon ab, ob die Linie die Achse $\bar{\Gamma}\bar{X}$ zwischen

²¹ M_i sind Spiegelsymmetrien

²² Obwohl das verwendete Modell nicht periodisch in k ist, kann die Windungszahl trotzdem entlang einer solchen unendlichen Geraden berechnet werden, da erstens die Oberflächenzustände für große Werte von k mit den Volumenzuständen hybridisieren und zweitens keine topologisch nicht trivialen Zustände außerhalb der unmittelbaren Umgebung des \bar{X} -Punkts zu erwarten sind.

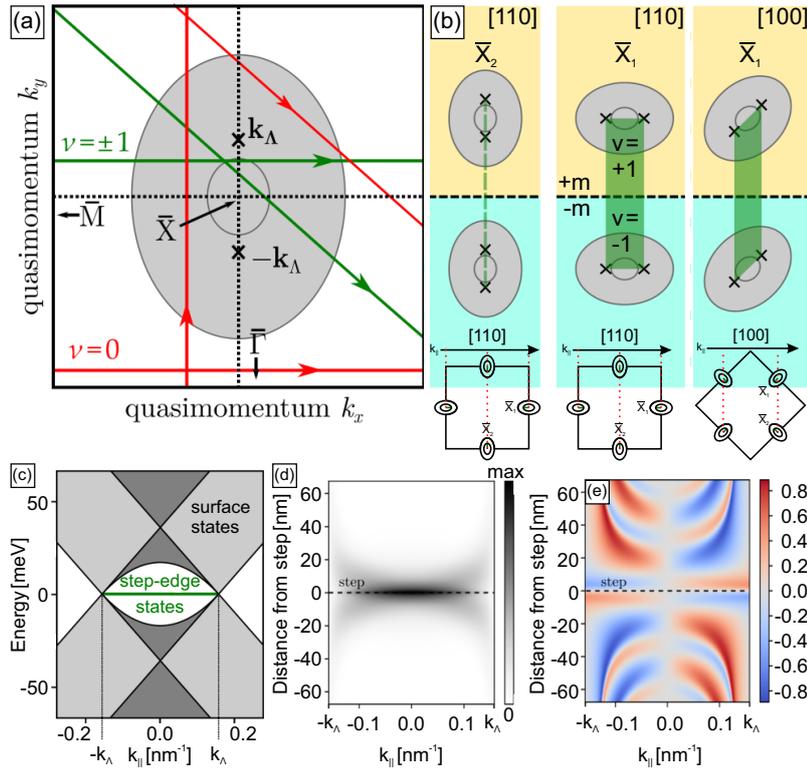


Abbildung 4.8 Richtungsabhängigkeit von Kantenzustände in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$: (a) Umgebung eines \bar{X} -Punkts mit Dirac-Punkten bei $\pm k_\Lambda$. Linien, die die Verbindungsstrecke ζ zwischen den beiden Dirac-Punkten kreuzen, haben eine Windungszahl $\nu = \pm 1$, alle anderen von $\nu = 0$. (b) Projektion von ζ auf verschiedene Stufenrichtungen: $[110]$ und $[100]$. Die Stufen verlaufen stets horizontal, während die Brillouin-Zone gedreht wird. Die vollständige Brillouin-Zone mit allen vier \bar{X} -Punkten ist in den unteren Teilabbildungen gezeigt. Für die Projektion auf die $[110]$ -Richtung sind die \bar{X} -Punkte inäquivalent, weil die beiden \bar{X}_2 -Punkte auf den selben und die beiden \bar{X}_1 -Punkte auf verschiedenen k_{\parallel} -Werten abgebildet werden. Für die $[100]$ -Richtung fallen die Projektionen paarweise zusammen. Diese Projektionen korrespondieren 1 zu 1 mit der von ζ auf die jeweilige Stufenrichtung, sodass in den grün markierten Bereichen eine nicht triviale Windungszahl auftritt. (c) Veranschaulichung der erlaubten k_{\parallel} -Werte der Kantenzustände. (d,e) Lokalisierung und Spinpolarisation der Kantenzustände senkrecht zur Stufe. Adaptiert von [53].

den beiden Dirac-Punkten bei $\pm k_\Lambda$ schneidet (grün) oder nicht (rot):

$$\nu = \begin{cases} -\text{sign}(m \cdot v_x'') & \text{für Schnittpunkt zwischen } \pm k_\Lambda, \\ 0 & \text{sonst.} \end{cases} \quad (4.4)$$

Da der Parameter v_x'' an einer Stufenkante seinen Wert nicht ändert, führt nur ein Vorzeichenwechsel von m zu einem Vorzeichenwechsel von ν . Genau dies geschieht bei einer ungeraden Stufe, sodass die selbe Linie im k -Raum für benachbarte Terrassen ein entgegengesetztes Vorzeichen ergibt. Der topologische Index der Stufe ergibt sich dann aus der Differenz $\Delta\nu = \nu^A - \nu^B = \pm 2$. Dementsprechend hat eine ungerade Stufe zwei topologisch geschützte Zustände.

Diese Zustände sind bezüglich ihrer Energie und ihres Wellenvektors lokalisiert, treten also nur

entartet bei einer Energie und bei bestimmten Werten von k auf. Die Entartung folgt, wie bei allen anderen Modellen auch, aus der PHS, welche die beiden Zustände bei $E = -E = 0 = E_{D^*}$ pinnt. Die Einschränkung auf bestimmte k -Werte folgt aus der Tatsache, dass $\nu \neq 0$ nur für Linien gilt, die die Verbindungsstrecke ζ zwischen $\pm k_\Lambda$ schneiden. Daher ist für einen beliebigen Wellenvektor die Projektion auf \hat{e}_\parallel , den Einheitsvektor parallel zur Stufe, entscheidend: $\vec{k} \cdot \hat{e}_\parallel = k_\parallel$. Da bei ungeraden Stufen die Translationssymmetrie nur entlang der Stufe nicht gebrochen ist, bleibt k_\parallel zugleich auch eine gute Quantenzahl zur Bandstrukturberechnung. Aus Gl. 4.4 ergibt sich dann, dass $\Delta\nu \neq 0$ nur für solche Werte von k_\parallel gegeben ist, die folgender Ungleichung genügen:

$$|k_\parallel| < \vec{k}_\Lambda \cdot \hat{e}_\parallel. \quad (4.5)$$

Für die beiden Stufenorientierungen [110] und [100] sind in Abb. 4.8 (b) in grün die Werte von k_\parallel hervorgehoben, die Ungl. 4.5 genügen. Diese beiden Richtungen zeichnen sich durch eine Besonderheit bei der Projektion der 1. Brillouin-Zone auf k_\parallel aus, die in dem unteren Teil der Abbildung veranschaulicht wird. Es werden zwei \bar{X} -Punkte auf den selben k_\parallel -Wert projiziert. Hierfür gibt es genau zwei Möglichkeiten: Entweder werden für eine der beiden Spiegellinien (M_x oder M_y) die beiden dazugehörigen \bar{X} -Punkte auf den selben k_\parallel -Wert abgebildet (bei der [110]-Richtung) oder es kommt zu einer paarweisen Projektion zwischen den Spiegellinien (bei der [100]-Richtung). In letzterem Fall gibt es zwei k_\parallel -Werte auf die jeweils ein \bar{X} -Punkt von M_x und ein \bar{X} -Punkt von M_y projiziert werden. Daher wurde in der Abb. 4.8(b) für die unterschiedlichen Spiegellinien ein Index für \bar{X} eingeführt.

Eine Stufe entlang der [110]-Richtung verläuft entlang der Spiegellinie $\overline{\Gamma\bar{X}}$. Deswegen werden die \bar{X} -Punkte einer der beiden Spiegellinien, markiert als \bar{X}_2 , auf den selben k_\parallel -Wert projiziert. Daher steht auch die Strecke $\overline{k_{\Lambda, \bar{X}_2}^- k_{\Lambda, \bar{X}_2}^+}$ senkrecht zur [110]-Richtung, sodass nicht einmal $k_\parallel = k_{\bar{X}_2}$ die Bedingung aus Ungl. 4.5 erfüllt. Außerdem ist ν nicht für Linien definiert, die mindestens einen der Dirac-Punkte schneiden. Für die anderen beiden \bar{X} -Punkte, markiert als \bar{X}_1 , liegt die Strecke $\overline{k_\Lambda^- k_\Lambda^+}$ parallel zur [110]-Richtung, sodass deren Projektion darauf maximal wird. Diese Konfiguration führt zu einer großen lokalen Zustandsdichte in STS-Spektren.

Bei einer Stufe in [100]-Richtung werden die vier \bar{X} -Punkte auf insgesamt zwei k_\parallel -Werte projiziert. Dabei werden auf den selben k -Wert jeweils ein \bar{X} -Punkt von M_x und M_y abgebildet, sodass es an den beiden k -Werten zu einer äquivalenten physikalischen Situation kommt. Die Verbindungsstrecke $\overline{k_\Lambda^+ k_\Lambda^-}$ zwischen den Dirac-Punkten schließt nun einen Winkel $\phi = \frac{\pi}{4}$ mit der Stufenorientierung ein. Dementsprechend nimmt der Wertebereich von k_\parallel ab, sodass prinzipiell eine niedrigeres Maximum in der lokale Zustandsdichte zu erwarten ist [vgl. Abb. 4.8 (c)].

In den weiteren Teilbildern sind in Abhängigkeit von k_\parallel (x -Achse) und vom senkrechten Abstand

zur Stufenkante (y -Achse) die Lokalisierung (d) und die Spinpolarisation (e) gezeigt. Präziser formuliert ist in (d) die Summe aus Betragsquadraten von orthogonalen Lösungen zu sehen. Die Kantenzustände sind für kleine Werte von k_{\parallel} , die also nahe bei $k_{\bar{x}}$ liegen, sehr stark lokalisiert. Generell zeigt sich, dass die Längenskala der Lokalisierung invers proportional zur Sattelpunktenergie ist. Andererseits zeigen die Berechnungen, dass für k_{\parallel} -Werte nahe der Dirac-Punkte eine signifikante Aufenthaltswahrscheinlichkeit auch für Distanzen d von bis zu 20 nm existiert. Für die unmittelbar an der Stufe lokalisierten Zustände zeigt die Spinpolarisation die typische Kopplung an den Wellenvektor. Wird die Bewegungsrichtung des Elektrons umgekehrt, so kommt es auch zu einer Invertierung des Spins. Da die Rechnung symmetrisch zur Achse $k_{\parallel} = 0$ ist, gilt dieser Befund für alle d . Interessanterweise kommt es zu Oszillationen der Polarisation, einerseits in Abhängigkeit von d und andererseits, bei hinreichend großem d , auch in Abhängigkeit von k_{\parallel} . Außerdem invertiert sich die Spinpolarisation, wenn das Vorzeichen von m geändert wird. Daher verlaufen Kantenzustände an der selben Stufe auf benachbarten Terrassen gegenläufig bezüglich ihrer Spinpolarisation. Weil der Kantenzustand nicht exakt auf der Stufe lokalisiert ist, sondern einige Nanometer neben ihr (siehe Abb. 4.5), lässt sich ihm eine eindeutig Spinpolarisation zuordnen.

Einen anderen Ansatz haben Brzezicki *et al.* verfolgt, die sich in ihrer Arbeit auf den symmorphen und nicht symmorphen Charakter der beteiligten Symmetrien konzentrieren. Dabei weist ihr Hamiltonian näherungsweise eine Spin-Rotations-Symmetrie auf, auf deren Grundlage sie eine dazugehörige topologische Invariante ermitteln. Mit Hilfe dieser beschreiben sie einen möglichen topologischen Ursprung für Kantenzustände allgemein in Abhängigkeit der Anzahl der Atomlagen auf den beiden Seiten einer Stufe [47]. Allerdings wurden weder die Spinpolarisation noch die maximale Dicke, respektive Anzahl an Atomlagen, bis zu der das Modell gültig bleibt, untersucht. Daher ist eine direkte Übertragung auf die Ergebnisse dieser Arbeit, deren Proben aufgrund ihrer Dicke in guter Näherung als halbbunendliche Kristalle betrachtet werden können, nicht möglich.

Zum Abschluss dieses Kapitels sollen noch einmal wesentliche Gemeinsamkeiten und Unterschiede zwischen den verschiedenen Modellen herausgestellt werden. Alle Modelle basieren direkt oder indirekt auf dem selben Hamiltonian aus Ref. [19] und weisen, nach entsprechenden Näherungen, eine PHS (und damit auch eine chirale Symmetrie) auf. Auch wenn sich dann der genaue Mechanismus und Formalismus zur topologischen Charakterisierung unterscheidet, so ist allen Modellen doch gemein, dass eben diese PHS entscheidend für die Energie der Kantenzustände ist. Da stets mindestens zwei, bzw. generell eine gerade Anzahl von Kantenzuständen vorhergesagt werden, führt die PHS dazu, dass die Zustände bei der Dirac-Punkt-Energie E_{D^*} entartet auftreten. Der fundamentale Unterschied liegt in der topologischen Beschreibung der Kantenzustände und somit in ihrem topologischen Ursprung. In Folge dessen besteht auch Uneinigkeit

über und insbesondere ob die Kantenzustände spinpolarisiert sind. Durch die Analyse von Wechselwirkungen zwischen den Kantenzuständen, aber auch mit magnetischen Momenten, können die in dieser Arbeit durchgeführten Experimente einen Beitrag zur Klärung dieser Debatte liefern.

5 Kopplung von 1D Kantenzuständen

Im vorherigen Kapitel wurden die bisherigen Ergebnisse zu (Pb,Sn)Se zusammengefasst und verschiedene Modelle zu dem topologischen Ursprung der Kantenzustände vorgestellt. Dabei hat sich eine Diskrepanz zwischen Ref. [39] und [51] bezüglich potenzieller Kantenzustände an geraden Kanten ergeben. Im ersten Teil dieser Arbeit wurden daher eine Vielzahl von geraden Stufen spektroskopisch auf Signaturen eines Kantenzustands bei E_{D^*} untersucht. Die dabei entdeckten, intrinsisch auftretenden Koppelungseffekte von 1D Kantenzuständen bilden den Hauptteil dieser Arbeit und dieses Kapitels. Zunächst beginnt dieses Kapitel jedoch mit einer kurzen Zusammenfassung wie im Allgemeinen die Rohdaten aufbereitet wurden. Im anschließenden Kapitel wird dann der Fokus auf die Wiederverwendung von (Pb,Sn)Se–Einmalspaltproben gelegt, um magnetische Adatome und deren extrinsische Wechselwirkungen zu untersuchen.

5.1 Datenaufbereitung

Topographie

Bei topographischen Daten wird zunächst ein Untergrund subtrahiert, um die Verkippung der Probenoberfläche relativ zur Scanebene zu kompensieren. Ursache dieser Verkippung können thermische bedingte Längenänderungen, verzögertes Reagieren der Piezoelemente auf die angelegte Spannung, eine endliche Präzision der horizontalen Ausrichtung bei der Befestigung der Probe oder ein Winkel zwischen Spalt- und Scanebene sein. Nach dieser Korrektur erscheinen alle Terrassen, die zur selben Atomlage gehören, in der selben Farbe. Außerdem werden die Stufenkantenhöhen in Linienprofilen einfacher ersichtlicher, da einerseits ein Mittelwert von der y -Achse abgezogen wird und andererseits die Verkippung der einzelnen Terrassen korrigiert wird. Die neue Achsenskala erleichtert das Ablesen der Stufenhöhen, weil sich an einer Stufe der Absolutwert der Regelschleife oft um weniger als 1% ändert. Dementsprechend macht sich in den Rohdaten die Stufe erst in den Nachkommastellen bemerkbar. Außerdem sind diese Höhenunterschiede oft auch klein gegenüber der Steigung der Verkippung, sodass nur wenige Datenpunkte um die Stufe herum zur Höhenbestimmung herangezogen werden können. In den bearbeiteten Daten hingegen sind die einzelnen Terrassen waagrecht und die Skala der y -Achsen ist von der selben Größenordnung wie die Stufenkantenhöhen.

Soweit möglich wird diese Korrektur mit Hilfe einer einzelnen Ebene durchgeführt, die von dem gesamten Bild subtrahiert wird. Definiert wird diese Ebene durch einen Bildausschnitt der physikalisch betrachtet auf der gleichen Höhe liegen muss, z.B durch eine oder mehrere Terrassen die zur gleichen Atomlage gehören. In einigen Fällen ist dies jedoch nicht möglich, sodass stattdessen linienweise ein Fit n -ter Ordnung ($n \leq 2$) abgezogen wurde.

Die weiteren Bearbeitungsschritte hängen von der Größe und Art des gezeigten Bildes ab. Für Übersichtsaufnahmen mit mehreren Stufenkanten wird der Kontrast an diesen erhöht, indem die Ableitung berechnet und mit dem Bild gemischt wird. Je größer dabei der Anteil der Ableitung ist, umso stärker werden die Stufen hervorgehoben. Allerdings nimmt mit zunehmender Ableitung auch die Breite der Stufen zu.

Auch atomar aufgelöste Bilder werden mit einer bearbeiteten Version des gleichen Bilds gemischt, um die Atome besser sichtbar zu machen. Hierfür eignet sich in der Regel eine schnelle Fouriertransformation (Fast Fourier Transform, FFT) und -filterung besser als die Ableitung. Das liegt zum einen daran, dass solche Aufnahmen sehr schnell gemacht werden, um Verzerrungen durch die Hystereseffekte der Piezos oder thermische Ausdehnungen zu minimieren. Mit steigender Messgeschwindigkeit nimmt dann jedoch auch das Rauschen zu. Zum anderen spielt die relativ geringe Korrugation eine wichtige Rolle, sodass beide Effekte oft dazu führen, dass die Ableitung selbst nur unzureichend zur Kontrastverstärkung ist. Stattdessen wird das gesamte Bild transformiert (2D FFT) und anschließend nur jene Bereiche für die inverse FFT berücksichtigt, die die charakteristischen Gitterpunkte beinhalten. Alle Frequenzkomponenten des Rauschens die außerhalb liegen werden somit aus dem Bild gefiltert. Dieses gefilterte Bild wird dann mit dem Ausgangsbild gemischt. Dadurch bleibt die ursprünglichen Informationen vollständig erhalten, aber die selektierten Frequenzkomponenten sind stärker gewichtet. Der Nachteil dieser Methodik zeigt sich an den Bildrändern, wo es, durch den Algorithmus bedingt, zu Verzerrungen kommt.

Spektroskopie

Informationen zur LDOS können sowohl mit ein- als auch ausgeschalteter Regelschleife erzielt werden. Im ersten Fall wird gleichzeitig zur Topographie eine Karte der differentiellen Leitfähigkeit aufgezeichnet, die die LDOS bei der angelegten Spannung zeigt. Solche dI/dV -Karten werden in dieser Arbeit in Form der Rohdaten gezeigt, da die physikalische Information ausschließlich in den relativen Intensitäten steckt. Durch Bildbearbeitung können jedoch Artefakte entstehen, die genau diese Verhältnisse verzerren.

Wenn die Regelschleife ausgeschaltet ist wird hingegen ein Spektrum über einen endlichen Spannungs- bzw Energiebereich aufgezeichnet. Wie in Kap. 3.2.2 bereits erläutert wurde, wird auf eine Normierung hier aufgrund der geringen Intervallbreite verzichtet. Für einzelne Spektren

wurde außerdem bei der Akquirierung darauf geachtet, dass die Messzeit pro Datenpunkt hinreichend groß im Vergleich zur Zeitkonstante des Lock-In-Verstärkers ist. Dadurch wird verhindert, dass die Daten systematische bezüglich der x-Achse verschoben sind. Für große Datensätze, bei denen mehrere tausend Spektren aufgenommen wurden, musste dieser Effekt in Kauf genommen werden, um die Messzeit an die Standzeit des LHe-Kyrostaten anzupassen. Da in diesen Fällen jedoch nur relative Energieunterschiede analysiert werden, fällt bei der Differenz dieser systematischer Fehler heraus.

Zum Glätten der Spektren wurde entweder ein FFT-Filter mit einer Fensterbreite von 3-5 Datenpunkten oder ein Polynom dritter Ordnung verwendet. Erstere Methode wurde für einzelne Spektren verwendet und basiert darauf, die höheren Frequenzkomponenten oberhalb eines Grenzwerts zu vernachlässigen. Die Verwendung eines Polynoms auf der anderen Seite reduziert die benötigte Rechenzeit bei der Bearbeitung großer Datensätze. Bei diesen wird neben dem direkt gemessenen dI/dV -Signal auch dessen zweite Ableitung numerisch bestimmt. Um hierbei das Rauschniveau hinreichend stark zu vermindern wurde beim Ableiten das dI/dV -, bzw. d^2I/d^2U -Signal durch Mittelung über mehrere Datenpunkte nochmals geglättet.

5.2 Kantenzustände an geraden Stufen

Zu Beginn dieser Arbeit wurden mehrere $\text{Pb}_{0.7}\text{Sn}_{0.3}\text{Se}$ -Proben gespalten und mittels eines x-y-Schiebetischs auf der gesamten Oberfläche nach geradzahlig Stufenkanten gesucht. Diese wurden dann durch einzelne Punktspektren oder in Form von mehreren Spektren entlang einer Linie (Linienpektroskopie) untersucht. Dabei wurden Parameter wie die Gesamtlänge, räumliche Auflösung entlang der Linie oder die Orientierung der Linie zur Stufenkante (entlang oder senkrecht dazu) variiert.

Die in Abb. 5.1 (a-d) gezeigten Ergebnisse sind repräsentativ für alle dabei gemessenen Datensätze. Zu sehen ist in (a) eine Topographie zweier Stufen die knapp unterhalb des Bildausschnitts zusammenlaufen. Aus der Farbkodierung geht hervor, dass die beiden Stufen unterschiedlich hoch sind. Aus dem oberen Bildteil wurde entlang der blauen Linie ein Linienprofil extrahiert, welches in (b) gezeigt ist. Auf der linken Seite beträgt die Stufenhöhe 3 \AA , was einer Atomlage entspricht. Auf der anderen Seite ist dieser Wert doppelt so groß, es handelt sich daher um eine Doppelstufe.

Die in (c) gezeigten Spektren sind an den entsprechend markierten Stellen in (a) aufgenommen worden. Die ungerade Stufe (blau) zeigt ein sehr ausgeprägtes Maximum bei etwa 120 meV , welches somit bei der Dirac-Energie liegt. An der geraden Stufe (rot) erkennt man bei annähernd der selben Energie ein deutliches Minimum, welches somit eine gute Approximation des Dirac-Punkts darstellt. Rechts und links des Minus sind zwei lokale Maxima zu erkennen.

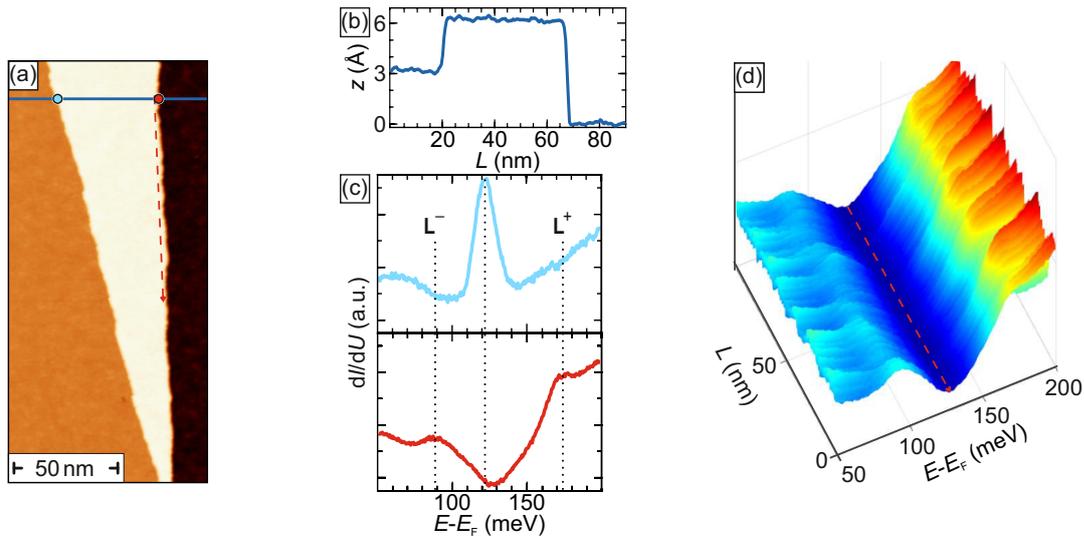


Abbildung 5.1 Analyse gerader Stufenkanten I: (a) Topographie von drei Terrassen, die durch eine ungerade bzw. eine gerade Stufenkante voneinander getrennt sind. (b) Aus dem Linienprofil geht hervor, dass links eine Einfach- und rechts eine Doppelstufe verläuft. (c) Spektren an den in (a) markierten Orten. (d) Linienspektroskopie entlang des roten Pfeils in (a). Wie in allen weiteren Daten auch sind keine Anzeichen für einen Kantenzustand an einer geraden Stufe erkennbar. Messparameter: $U = 120$ mV (a), $I = 200$ pA (a). Stabilisierungsparameter: $U = 200$ mV (c,d), $I = 200$ pA (c,d).

Diese sind auf die vHS zurückzuführen, sodass das Spektrum von seiner Form her identisch und somit ununterscheidbar von einem Spektrum auf einer atomar glatten Terrasse ist.

Entlang des in (a) eingezeichneten roten Pfeils wurde eine Linienspektroskopie durchgeführt. Deren Spektren sind in (d) als Wasserfallplot dargestellt. Gut zu erkennen sind leichte Variationen im Bereich der beiden vHS, nicht jedoch bei der Dirac-Energie. Bis zum heutigen Tag konnte in keiner der durchgeführten Messungen an einer geraden Stufenkante ein spektroskopischer Hinweis auf einen Kantenzustand in Form eines Einzel- oder Doppelpeaks gefunden werden. Im Rahmen dieser Arbeit können daher die Ergebnisse aus Ref. [51] nicht verifiziert werden.

Das verwendete Modell (vgl. Seite 42 in Kapitel 4.3) liefert zwar eine plausible Erklärung, warum nicht an jeder geraden Stufenkante ein Zustand beobachtbar sein muss, gleichzeitig fehlt aber auch eine belastbare Vorhersage, wann ein solcher Zustand auftreten kann bzw. sollte. Insbesondere die Frage ob es möglich ist, dass es unter bestimmten Bedingungen niemals Kantenzustände an gerade Stufen geben kann, bleibt unbeantwortet. Deswegen ist auch eine eindeutige Falsifizierung dieser Hypothese im Rahmen dieser Studie nicht möglich.

Ein weiteres Beispiel für eine untersuchte Stelle ist in Abb. 5.2(a) gezeigt. Zu sehen sind mehrere keilförmig zulaufende Stufenkanten, von denen es sich bei den meisten um ungerade Einfachstufen handelt. Dies wird durch die dI/dV -Karte in Abb. 5.2(b) veranschaulicht. Da diese bei der Dirac-Energie aufgenommen wurde, kommt es an ungeraden Stufen zu einer signifikanten Steigerung der Intensität. Lediglich bei der Stufe mit dem Spektrum (2) ist die Intensität nur geringfügig erhöht. Aus dem Linienprofil in (c) wird ersichtlich, dass es sich um eine Doppelstufe

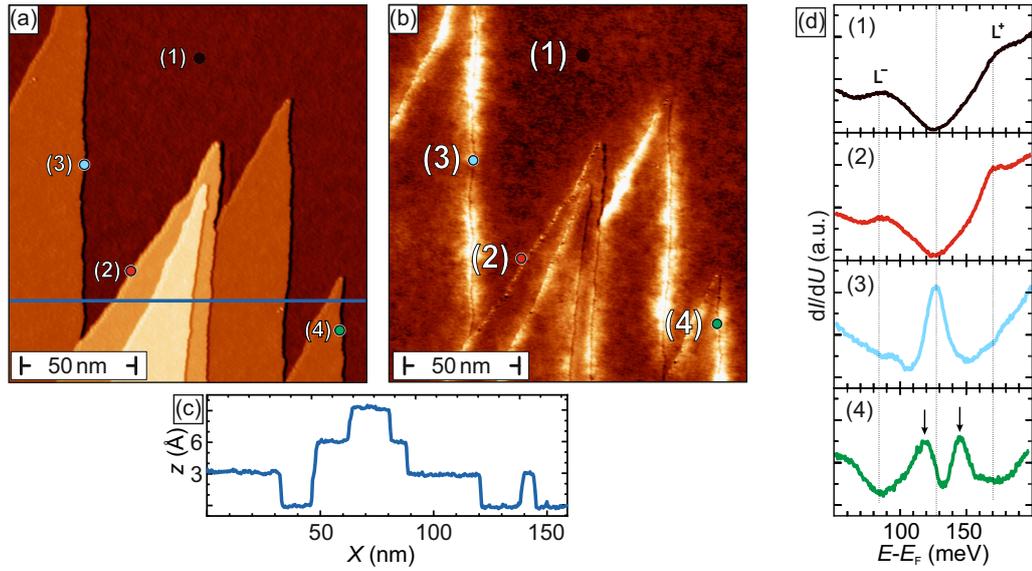


Abbildung 5.2 Analyse gerader Stufenkanten II: (a,b) Topographie mit mehreren zusammenlaufenden Stufen sowie simultan gemessene dI/dU -Karte der Dirac-Energie. (c) Aus Linienprofil und dI/dU -Karte geht hervor, dass nur die Stufe mit der Markierung (2) gerade ist. (d) Spektren an den in (a) und (b) markierten Positionen. Terrasse und gerade Stufe sind spektroskopisch ununterscheidbar, der Kantenzustand bei (3) energetisch leicht verschoben. Die ungerade Kante bei (4) zeigt unerwartet eine Aufspaltung in zwei Peaks. Messparameter: $U = 122$ mV (a, b); $I = 200$ pA (a, b). Stabilisierungsparameter: $U = 200$ mV (d); $I = 200$ pA (d).

handelt. Die minimal größere Intensität im Vergleich zu den umgebenden Terrassen resultiert aus einem Überschießen des Regelkreises an der Stufe und ist daher ein Messartefakt.

An verschiedenen Stellen wurden Punktspektroskopien durchgeführt, deren Ergebnisse in (d) gezeigt sind. Die Terrasse (1) zeigt sehr deutlich die charakteristische v -Form sowie die lokalen Maxima der vHS. Bei der Doppelstufe an Position (2) ändert sich, in Übereinstimmung mit den Ergebnissen aus Abb. 5.1 (d), nichts an dem Verlauf des Spektrums. Die Intensität in der dI/dV -Karte ist an Position (3) erheblich vermindert. Trotzdem zeigt sich eindeutig ein Peak, der dem Kantenzustand zuzuordnen ist. Eine genauere Betrachtung zeigt, dass sich das Maximum leicht zu höheren Energien verschoben hat, sodass sich bei der angelegten Spannung $U = 122$ mV die LDOS verringert. Ursächlich für diese lokale Schwankung des chemischen Potentials ist die zufällige Verteilung von Pb- und Sn-Atomen in dem entsprechenden Untergitter.

Obwohl die in Ref. [51] beobachteten Doppelpeaks an geraden Stufenkanten nicht reproduziert werden konnten, wurde überraschenderweise im Rahmen dieser Studie an ungeraden Stufenkanten gelegentlich ein Doppelpeak entdeckt. Eine dieser Stellen ist die Position (4), deren Spektrum zwei klar getrennte Peaks zeigt. Deren Intensität ist im Vergleich zum Einfachpeak verringert und ihre Aufspaltung beträgt etwa 25 meV.

Im folgenden Unterkapitel geht es darum zu klären, unter welchen Bedingungen dieser Doppelpeak entsteht und welche Gesetzmäßigkeiten diese Aufspaltung folgt. Außerdem soll ein einfaches Modell für dessen Beschreibung präsentiert werden.

5.3 Koppelungseffekte von Kantenzuständen an ungeraden Stufen

Die zuvor beobachtete Aufspaltung an ungeraden Stufen wurde zunächst an verschiedenen Proben und Stellen gemessen. Die Häufigkeit dieses Effekts schwankte dabei erheblich, sodass auf der Längenskala eines typischen STM-Übersichtsbildes von 250-500 nm oft gar keine und in anderen Fällen an mehreren Punkten eine Aufspaltung zu sehen war. Zudem variierten sowohl die Größe der Aufspaltung ΔE als auch das Intensitätsverhältnis zwischen den Peaks. Der Wert von ΔE bewegte sich dabei zwischen wenigen meV, (bei einem Peak mit einer Schulter) und mehreren 10 meV [wie z.B. bei Spektrum (4) in Abb. 5.1 (g)]. Neben einigen Spektren mit gleichgroßen Peaks wurde in der Mehrzahl der Fälle eine Asymmetrie festgestellt. Dabei kann sowohl der nieder- als auch der hochenergetische Peak größer sein.

Aus der Analyse einer Vielzahl von einzelnen Beobachtungen konnte dabei ein Zusammenhang zwischen der Topographie der Probe und dem Doppelpeak festgestellt werden. Die Korrelation zwischen Topographie und Aufspaltung besteht darin, dass die Aufspaltung nur dann zu sehen ist, wenn sich innerhalb von wenigen 10 nm eine weitere ungerade Stufe befindet. Bei geraden Stufen, wie z.B. in Abb. 5.1 (a), konnte kein Einfluss auf den Kantenzustand festgestellt werden. Es handelt sich somit um einen Effekt, der durch die Koppelung zweier Kantenzustände zustande kommt.

Stufen mit kleinem Abstand zueinander können prinzipiell auf drei verschiedene Topographien zurückgeführt werden. Die betreffenden Stufen können parallel verlaufen, in einem gemeinsamen Punkt enden oder sich gegenseitig schneiden und nach dem Schnittpunkt weiterlaufen. Alle drei Formen wurden in der vorliegenden Arbeit untersucht und werden in den folgenden Abschnitten vorgestellt.

5.3.1 Zusammenlaufende Stufenkanten

Die mit Abstand am häufigsten anzutreffende Konfiguration sind zusammenlaufende Stufen. Zwei Beispiele sind in den Abb. 5.4 & 5.5 gezeigt. Wie gleich diskutiert wird tritt die Koppelung nur für relativ kleine Stufe-Stufe-Abstände auf. Darum wurden für detaillierte Analysen Strukturen gesucht, bei denen die Stufen unter einem kleinen Winkel zueinander verlaufen. Typischerweise liegt der Winkel zwischen 5-10°, was zu stark in die Länge gestreckten, keilförmigen Strukturen führt, die zumindest abschnittsweise sogar parallele Segmente aufweisen.

In Abb. 5.4 (a) ist eine solche Topographie für eine keilförmige Insel auf einer ansonsten flachen Terrasse zu sehen. Die in (b) gezeigte dI/dV -Karte ist bei der Dirac-Energie aufgezeichnet worden und zeigt im unteren Bildteil eine ausgeprägte Intensität an den Stufen. Diese fluktuiert leicht und nimmt in der Bildmitte, markiert durch zwei Pfeile, abrupt signifikant ab. Im weiteren

Verlauf variiert die Intensität erneut, jedoch auf niedrigerem Niveau, bevor sie am oberen Ende der Keilstruktur (gestrichelte Ellipse) fast vollständig im Hintergrundsignal der umgebenden Terrasse verschwindet.

Da die Variationen der Intensität in der dI/dV -Karte alleine kein hinreichendes Kriterium für die Existenz des Doppelpeaks sind [siehe Spektrum (3) Abb. 5.2 (d)], wurde über dem gesamten Bildbereich von (a) eine Rastertunnelspektroskopie durchgeführt. Die Ergebnisse sind repräsentativ für eine Vielzahl solcher Datensätze, bei denen die laterale Auflösung üblicherweise etwa 5-20 nm entlang und 0.5-2 nm senkrecht zu den Stufen beträgt. Die gemessenen Spektren wurden dann in drei Klassen eingeteilt: Einfach-, Doppel- oder kein Peak.

Für die Klassifizierung spielte es dabei keine signifikante Rolle, ob sie händisch erfolgte oder ob die Spektren mit mehreren Gaussfunktionen angefitet wurden. In ersterem Fall wurden üblicherweise nur die Spektren mit wenigen Nanometern Abstand zu einer der Stufen ausgewertet. Die anderen Spektren wurden dann als peakfrei angenommen, da ihr Anteil am gesamten Datensatz deutlich größer als 50% ist und selbst die kleinsten Datensätze aus 1500-2500 Einzelspektren bestehen (Spitzenwerte: ≈ 25 -30k). Dieses Vorgehen wird außerdem durch die relativ starke Lokalisierung der Zustände an der Stufenkante gerechtfertigt [39], sodass in größeren Entfernungen zu den Stufen der Zustand bereits abgeklungen ist.

Für die Analyse mittels Fits wurden erst zwei, dann drei und ggf. nochmals vier Gaussfunktionen an den Energiebereich zwischen den vHS gefittet. Für die Klassifizierung ist dann ausschlaggebend, ob eine weitere Gaussfunktion noch zu einer signifikanten Verbesserung des Fits führt. Lediglich zwei Peaks werden für ein Spektrum auf einer Terrasse benötigt, um eine passable Anpassung zu erhalten. Deren Positionen stimmen dann ungefähr mit denen der vHS überein. Zusehen ist dies in Abb. 5.3 (a), wo die einzelnen Gaussfunktionen in blau und ihre Summe in rot dargestellt werden. Die Messdaten wurden nur im Energiebereich 75-175 meV angepasst, weil außerhalb von diesem Bereich das Spektrum durch die Beiträge von Valenz- und Leitungsband dominiert wird. Diese sind jedoch nicht von Interesse, sodass ihre Berücksichtigung keinen Mehrwert an Information bietet, gleichzeitig aber die Anpassung der Daten erschwert.

Im Teilbild (b) ist ein Spektrum mit einem Kantenzustand gezeigt. Hier steigert sich die Qualität des Fits erheblich, wenn eine dritte Gaussfunktion berücksichtigt wird. Die niederenergetische vHS ist nur noch schwer erkennbar und wird durch eine Gaussfunktion angepasst, deren Maximum nicht mehr im abgebildetem Energiebereich liegt. Für einen Doppelpeak ergibt sich analog durch eine vierte Gaussfunktion erneut eine signifikante Verbesserung des Fits. Weil aber bei sinnvoller Wahl der Startwerte und Intervallgrenzen zusätzliche Freiheitsgrade beim Anpassen eines Fits immer zu einem besserem Ergebnis führen, ist die Unterscheidung, welche Verbesserungen signifikant sind, entscheidend.

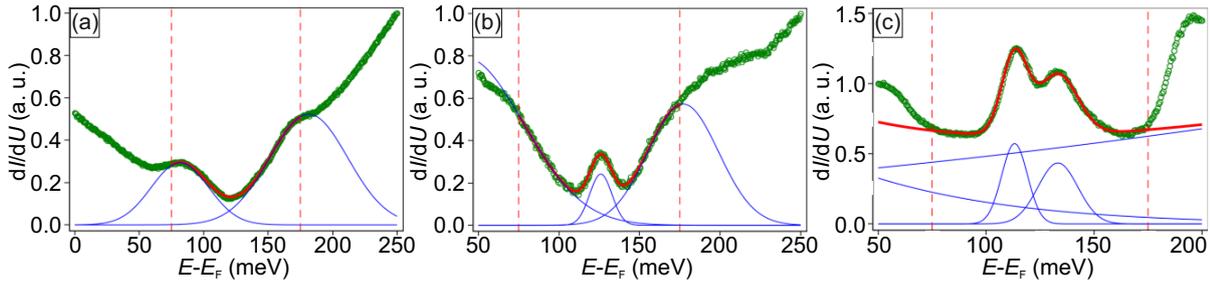


Abbildung 5.3 Klassifizierung durch Anpassung mehrerer Gaussfunktionen: (a) Ein Spektrum über einer Terrasse (grün) wird um den Dirac-Punkt herum gut durch zwei Gaussfunktionen (blau) approximiert. Nur die Daten zwischen den vertikalen, gestrichelten Linien werden für den Gesamtfit (rot) angepasst. (b) Ein analoges Vorgehen für bei einem Kantenzustand nur dann zu einem vergleichbar gutem Ergebnis, wenn zusätzlich ein dritter Peak in der Nähe der Dirac-Energie berücksichtigt wird. (c) Bei einem Doppelpeak ergibt sich durch die Berücksichtigung zweier weiterer Gaussfunktionen nahe des Dirac-Punkts die beste Anpassung. Die ersten beiden Peaks, welche zuvor grob bei den vHS lagen, können dann jedoch weit außerhalb des Messbereichs liegen und große FWHM aufweisen. Gut zu erkennen ist, wie die Anpassung des Fits an die Daten außerhalb des gestrichelten Bereichs erheblich schlechter ist. Stabilisierungsparameter $U = 250$ mV (a,b), $U = 50$ mV (c), $I = 200$ pA (a-c).

Eine pauschale Antwort auf diese Frage ist nicht trivial, da zwar statistische Kriterien zur Bewertung der Güte eines Fits, wie z.B. das Bestimmtheitsmaß R^2 , verfügbar sind, jedoch eine objektive Festlegung von Grenzwerten kaum möglich ist. Die Intervallgrenzen von Fitparametern und in ungünstigen Fällen sogar deren Startwerte, können zu unterschiedlichen Fits und verschiedenen R^2 führen. Da die Standardalgorithmen beim Fitten stets die Abweichung maximal minimieren möchten, nutzen sie jeden verfügbaren Parameter, selbst wenn dies nur zu einer sehr kleinen Verbesserung von R^2 führt. Deswegen wird, z.B. auch beim Fitten eines Terrassenspektrums mit mehr als zwei Gaussfunktionen, niemals eine der Amplituden Null gesetzt. Dies macht das Prozedere mit mehreren Fits und unterschiedlich vielen Gaussfunktionen nötig.

Im Folgenden sollen daher nun einige physikalische Überlegungen zur Einschränkung der drei Parameter einer Gaussfunktion [Amplitude, Position, Halbwertsbreite (Full Width at Half Maximum, FWHM)] präsentiert werden. Die Amplitude kann für jede Gaussfunktion auf den Bereich 0-10 begrenzt werden, da dies dem Wertebereich des Signalausgangs des Lock-In-Verstärkers entspricht.

Die Position der Peaks kann gut beschränkt und die dazugehörigen Startwerte relativ genau geschätzt werden. Die ersten beiden Peaks liegen relativ nahe bei den vHS und somit weit entfernt vom Dirac-Punkt. Generell weicht ihre Lage aber umso stärker von der der vHS ab, je dominanter der Kantenzustand im Spektrum ist. Besonders gut ist dies für den Doppelpeak in Abb. 5.3(c) zu sehen. Hier ist bei $E \approx 175$ meV kein lokales Maximum sondern ein lokales Minimum beobachtbar. In solchen Fällen kann es vorkommen, dass die ersten beiden Peaks beim Anpassen weit aus dem Messbereich geschoben und mit einer relativ hohen FWHM modelliert werden. Sie stellen dann zwar keine sinnvolle Approximation der vHS mehr dar, aber sie beschreiben im-

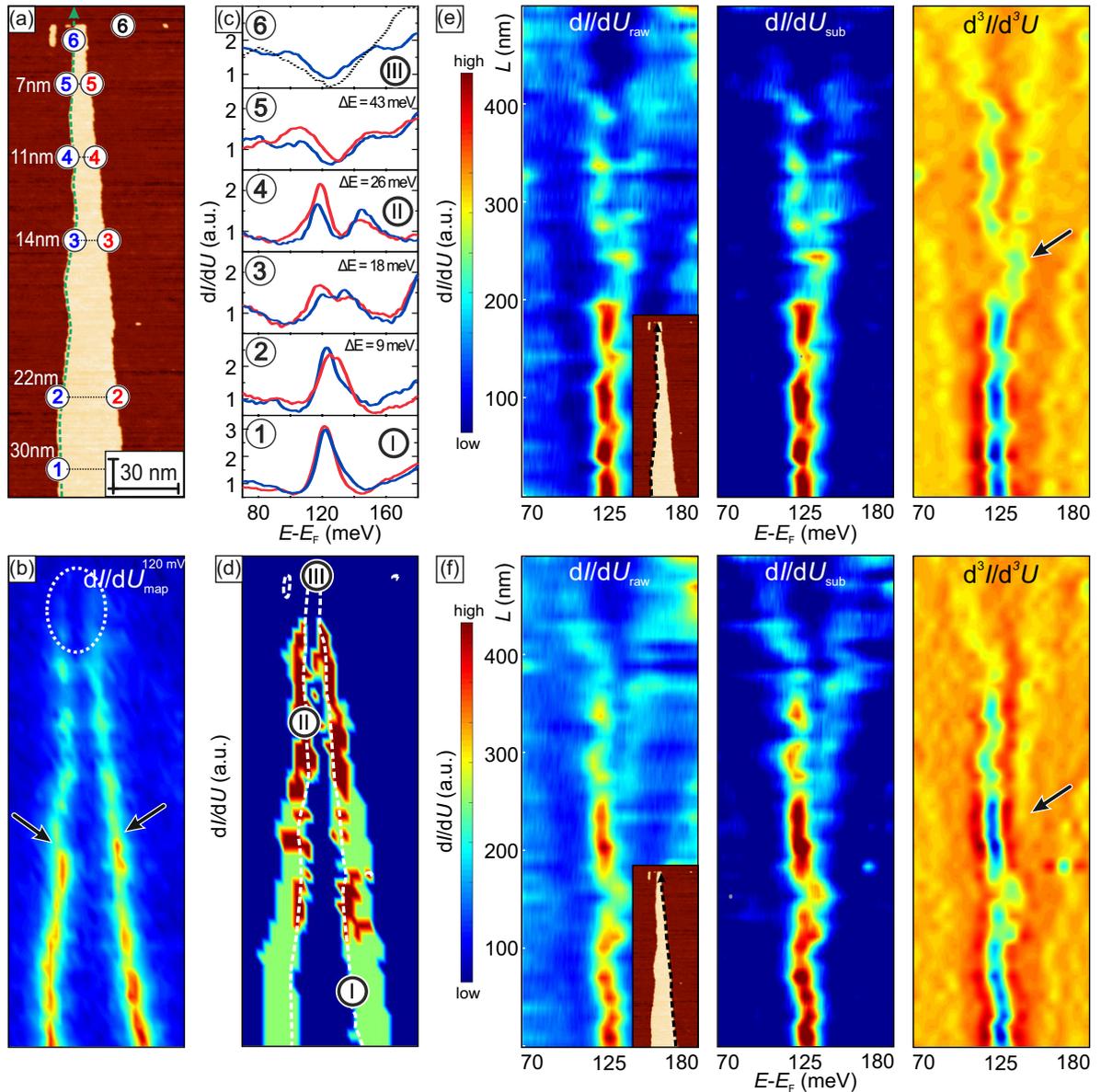


Abbildung 5.4 Kopplungseffekte an zusammenlaufenden, inselartigen Stufenkanten: (a) Topographie von zusammenlaufenden, ungeraden Stufen, die eine Insel bilden. (b) Simultan gemessene dI/dV -Karte bei der Dirac-Energie. (c) Ausgewählte Spektren der linken (blau) und rechten (rot) Stufe bei verschiedenen Abständen. Die Spektren (I), (II) & (III) sind exemplarisch für die in (d) farbcodierten Bereiche. (d) Über das gesamte Bild in (a) wurde eine Rastertunnelspektroskopie durchgeführt. Jedem Spektrum wurde eine Klasse und jeder Klasse eine Farbe zugeordnet: Einfachpeak grün, Doppelpack rot und kein Peak blau. (e,f) Spektren entlang der linken (e) und rechten (f) Stufe, entnommen aus dem selben STS-Datensatz (vgl. gestrichelte Linien in den Einschubbildern). Gezeigt sind links die Rohdaten, in der Mitte die Rohdaten abzüglich eines Hintergrunds und rechts die zweite Ableitung des Spektrums, also d^3I/d^3U . Messparameter: $U = 120$ mV, $I = 200$ pA (a,b). Stabilisierungsparameter $U = 200$ mV, $I = 200$ pA (c,e,f).

mer noch den Hintergrund des Kantenzustands akzeptabel (im relevanten Energiebereich). Die Peaks drei und vier sind hingegen immer auf einen Bereich um die Dirac-Energie beschränkt, weil diese den Kantenzustand beschreiben.

Die FWHM kann für die ersten beiden Gaussfunktionen nur grob begrenzt werden, da einerseits geringe Werte zum Anpassen der relativ schmalen vHS benötigt werden, zum anderen aber auch große Werte um einen flachen Hintergrund zu erzeugen, auf dem ein dominanter Kantenzustand sitzt [vgl. Abb 5.3 (a) und (c)]. Die Gaussfunktionen drei und vier hingegen können grob mit der FWHM eines ungekoppelten Kantenzustands abgeschätzt werden. Nach unten sind sie dann durch die experimentelle Energieauflösung begrenzt, nach oben ergibt sich jedoch ein Ermessensspielraum, welcher erneut R^2 beeinflusst. Ursächlich hierfür ist die Tatsache, dass die FWHM mit der Lebensdauer des Zustands verknüpft ist und sich diese wiederum in Folge der real auftretenden Wechselwirkungen ändern kann.

Beide Methoden zur Analyse weisen daher Schwächen bei der Objektivität auf. Nichtsdestotrotz liefern beide Ansätze gleichwertige Ergebnisse, was daran liegt, dass für die wesentlichen Aussagen in dieser Arbeit die Entscheidung, ob ein bestimmtes Spektrum nun einen oder zwei Peaks zeigt, statistisch gesehen nicht relevant ist. Aufgrund der Eingangs erwähnten großen Anzahl an Spektren innerhalb jedes einzelnen Datensatzes, sowie der zweistelligen Anzahl an Datensätzen insgesamt, sind die hier nun folgenden Befunde statistisch gesehen robust. Außerdem ist es gerade im Übergangsbereich von einem zu zwei Peaks plausibel anzunehmen, dass eine eindeutige Zuordnung nicht immer möglich ist. So kann im Rahmen der experimentellen Auflösung ein einzelner, verbreiteter Gausspeak äquivalent zu zwei theoretisch zu erwartenden Peaks mit sehr geringer Aufspaltung sein.

Aus der Rastertunnelspektroskopie über die Topographie von Abb. 5.4 (a) sind in (c) für fünf verschiedene Stufe–Stufe–Abstände d im Bereich von 7–30 nm die dazugehörigen Spektren der linken (blau) und rechten Stufe (rot) dargestellt. Für $d > 30$ nm (1) zeigen beide Stufen einen ausgeprägten Einfachpeak bei der Dirac–Energie. Die FWHM dieser Kantenzustände stimmt mit der von freistehenden, ungeraden Stufenkanten überein. Die Verbreiterung eines Zustands ist mittels der Heisenbergschen Unschärferelation direkt mit der Lebensdauer verknüpft. Die Lebensdauer eines Zustands wiederum wird i. d. R. durch Wechselwirkungen modifiziert. Deshalb kann aus der gleichbleibenden FWHM geschlossen werden, dass für $d > 30$ nm zwei entkoppelte Kantenzustände vorliegen.

Bei dem nächst kleineren Abstand von $d = 22$ nm (2) weisen die Spektren bereits signifikante Abweichungen in ihrer Form auf. An der linken Stufe ist ein relativ schmaler Peak mit einer kleinen Schulter bei höheren Energien zu sehen. Das rote, an der rechten Einfach–Stufenkante gemessene Spektrum zeigt eine merkliche Verbreiterung gegenüber (1), welche von der Breite her sowohl den schmalen blauen Peak als auch dessen Schulter umfasst. Außerdem ist eine leichte Aufspaltung des Peaks in zwei Maxima erkennbar. Diese Form deutet, genauso wie die Schulter, darauf hin, dass es sich nicht mehr um einen einzelnen, sondern um zwei Peaks handelt. Die Aufspaltung ΔE beträgt etwa 9 meV und ist damit vergleichbar zur FWHM eines ungestörten

Kantenzustands. Da die Spektren aus einem experimentellen Datensatz stammen, sind sie nicht nur bei den gleichen Parametern stabilisiert worden, sondern sind zusätzlich mit der selben Spitze gemessen worden. Hieraus ergibt sich, dass der Einfluss der Spitze auf die STS-Spektren minimal ist, sodass die größtmögliche Vergleichbarkeit der Intensitäten vorliegt. Diese nimmt Zwischen (1) und (2) deutlich ab, ein Trend der sich mit kleiner werdendem d weiter fortsetzt. Für $d = 14$ nm hat sich die Intensität in etwa halbiert und stagniert auf diesem Niveau bei weiterer Verringerung des Abstands. Eine adäquate Beschreibung der Spektren durch einen verbreiterten Einzelpeak ist aufgrund der reduzierten Intensität in der Mitte nicht mehr möglich. Andererseits sind die beiden Peaks noch nicht sehr stark ausgeprägt, was in diesem Bereich erwartungsgemäß zu großen Schwierigkeiten bei der Klassifizierung führt. Zum einen sind verschiedene Formen beobachtbar, von denen viele für einen möglichen zweiten Peak lediglich schwache lokale Maxima zeigen. Zum anderen erschwert vor allem die grobe laterale Auflösung entlang der Stufen die Beobachtung einer kontinuierlichen, räumlichen Entwicklung der Peaks. Somit fehlt eine wichtige Möglichkeit, um ein ambivalentes Maximum ggf. doch eindeutig einem Peak zuzuordnen zu können. Dort wo die Aufspaltung hinreichend klar erkennbar ist, ergibt sich mit $\Delta E = 18$ meV bereits eine deutlich stärkere Koppelung der Kantenzustände.

In dem Bereich um $d = 10$ nm (4) herum ist der Doppelpeak am deutlichsten zu sehen. Einerseits ist die Höhe der beiden Peaks im Vergleich zum Untergrund dominant und andererseits die Aufspaltung mit etwa 30 meV so groß, dass die Überlappung der Peaks nicht mehr stört. Die Form des Doppelpeaks lässt sich entsprechend der relativen Intensitäten in drei Gruppen einteilen: Symmetrisch (Abb. 5.4(c), Spektrum (4), blauer Graph), asymmetrisch mit höherer Intensität bei niedrigerer Energie (Abb. 5.4(c), Spektrum (4), roter Graph) und asymmetrisch mit höherer Intensität bei größerer Energie. Obwohl jeder Doppelpeak eindeutig zugeordnet werden kann, konnte keine kohärente Zuordnung entlang der jeweiligen Stufenkante beobachtet werden. Längs jeder Stufe variiert die Gruppenzugehörigkeit der Spektren innerhalb weniger Nanometer. So ist in Abb. 5.4(e) beispielsweise bei $L = 300$ nm die Intensität der beiden Peaks in etwa gleich hoch. Bei etwas geringerem L ist jedoch zunächst der Peak bei höherer Energie intensiver und noch ein kurzes Stück davor ist der Peak bei niedrigeren Energien intensiver. An der selben Stufe koexistieren somit symmetrische und asymmetrische Intensitätsverteilungen. Es konnte daher weder entlang noch zwischen den beiden Stufen ein eindeutiges Muster bezüglich dieser Gruppen festgestellt werden.

Eine weitere Reduzierung auf $d = 7$ nm (5) führt zu einer weiteren Aufspaltung von $\Delta E = 43$ meV. Auch der Untergrund, besonders oberhalb des Dirac-Punkts, gewinnt an Intensität. Dadurch nimmt die relative Peakhöhe im Vergleich zum Untergrund ab. Gleichzeitig steigt die FWHM, wodurch die Peaks schwerer zu identifizieren sind (beispielsweise der blaue Peak bei $E \approx 110$ meV).

Aufgrund dieser Trends sind für $d \lesssim 5$ nm (6) keine Peaks mehr eindeutig erkennbar. Das Spektrum nähert sich in (5) und (6) insgesamt mehr und mehr dem der Terrasse an (gestrichelte Linie). Zu sehen ist, dass die LDOS am Dirac-Punkt zwar noch minimal größer als auf der Terrasse ist, im Vergleich zu (1) jedoch dramatisch reduziert. Über alle Spektren hinweg lässt sich anhand der LDOS am Dirac-Punkt der Verlauf aus Abb. 5.4(b) verstehen. Die ungekoppelten Kantenzustände besitzen die größte Intensität, welche sich mit dem Einsetzen der Kopplung schlagartig reduziert (Pfeile). In der Folge schieben die Peaks in Richtung der vHS, wodurch sie sich immer weniger am Dirac-Punkt überlagern. Sobald der Abstand sehr klein wird (gestrichelte Ellipse) konvergiert aufgrund der Aufspaltung das Spektrum der Stufenkanten schließlich gegen das der Terrasse.

Prototypisch Spektren für die drei Klassen Einfach-, Doppel- oder kein Peak sind in (c) mit (I),(II) und (III) markiert. In Abb. 5.4(d) sind alle Spektren des Datensatzes den jeweiligen Klassen zugeordnet und mit den Farben grün (I), rot (II) und blau (III) kodiert worden. Die gestrichelten weißen Linien dienen als Orientierungshilfe für den Verlauf der Stufenkanten. Gut zu sehen ist, dass im unteren Bildausschnitt nur Einzelpeaks beobachtet wurden, was auf eine vernachlässigbare Kopplung schließen lässt. Daran schließt sich ein Übergangsbereich an, in dem den Spektren sowohl Einzel- als auch Doppelpicks zugeordnet werden. Dieser Bereich beginnt ab etwa 20 – 25 nm und weist sowohl zwischen den Stufen, also auch innerhalb der Stufen (Spektren senkrecht zur Stufe für die $d = konst.$ gilt) rote-grün gemischte Bereiche auf. Bei diesen Abständen führt die einsetzende Kopplung zu ersten Modifikationen der Peakform, die jedoch nur teilweise eindeutig einen Doppelpick repräsentieren. Sobald der Abstand einen kritischen Wert von $d \lesssim 15$ nm unterschreitet, ist die Koppelung so stark, dass beiden Stufen durchgehend zwei Peaks zugeordnet werden können. Aufgrund der weiter zunehmenden Aufspaltung ist anzunehmen, dass die Koppelung ebenfalls immer stärker wird. Ab $d \lesssim 5$ nm wird ein Regime der starken Kopplung erreicht, was dazu führt, dass die Peaks, wie zuvor beschrieben, in die vHS geschoben werden.

Um das zuvor beschriebene Verhalten besser zu visualisieren, sind in Abb. 5.4(e) & (f) die einzelnen Spektren entlang der linken bzw. rechten Stufe gezeigt. Hierfür wurden aus der Rastertunnelspektroskopie jeweils zwei Spektren pro Linie entnommen, welche an den beiden Stufen lokalisiert sind. Diese sind der Abbildung in y -Richtung gegen die Energie auf x -Achse aufgetragen. Die dI/dV -Werte sind von niedrigen Werten in blau hin zu hohen Werten in rot farbcodiert. In der linken Spalte sind, bis auf eine Reduktion des Energiebereichs, die Rohdaten gezeigt. Gut zu sehen sind jeweils die hohe Intensität im unteren Bildteil an der Dirac-Energie, welche von dem Einzelpeak stammt. Auch die sehr niedrige Intensität ganz oben im Bild bei der selben Energie ist klar erkennbar.

Verzerrt wird die Darstellung jedoch durch die teilweise recht ausgeprägte Intensität am Rand, ein Effekt, der bei Betrachtung des gesamten aufgenommenen Energiefensters noch stärker ist. Um diesen Beitrag, der vom Hintergrund herrührt, zu unterdrücken, wurde in den mittleren Spalten ein parabolischer Untergrund abgezogen. Nun lässt sich im Bereich zwischen etwa 250-400 nm eine Art energetische Verbreiterung (e) bzw. zwei Peaks (f) erkennen.

Das Regime der starken Kopplung und das Einsetzen des zweiten Peaks werden jedoch erst in der rechten Spalte deutlich, wo die zweiten Ableitung des dI/dV -Signals, also das d^3I/d^3U -Signal, gezeigt ist. Dies entspricht der Krümmung des dI/dV -Signals, sodass ein einzelner Peak durch eine rechts-links-rechts Krümmung beschrieben wird. Im unteren Bildausschnitt ist anhand des Einzelpeaks zu sehen, dass das in dem gewählten Farbschema einer rot-blau-rot Abfolge entspricht. Dementsprechend ist für einen Doppelpeak eine R-B-R-(R)-B-R Sequenz zu erwarten, wobei die Rechtskrümmung zum Ende des ersten und die zum Anfang des zweiten Peaks kontinuierlich ineinander übergehen und deshalb als ein einzelner Bereich auftreten.

An der mit einem Pfeil markierten Stelle in (f) kommt es schlagartig zum Einsetzen der zuvor beschriebenen Krümmungssequenz. Allerdings ist die Linkskrümmung der Doppelpeaks schwächer als die des Einfachpeaks, was dazu führt dass die Doppelpeaks grün und nicht blau erscheinen. Interessanterweise scheint es, dass der zweite Peak einfach einsetzt und sich nicht kontinuierlich durch eine Aufspaltung entwickelt, wie es die beiden linken Spalten suggerieren. Im Fall der linken Stufe in (e) ist das dazugehörige Spektrum, respektive Stelle, ebenfalls mit einem Pfeil markiert. Zum einen ist hier die Krümmung weniger stark ausgeprägt, was an der Farbskala erkennbar ist, zum anderen setzt der zweite Peak etwas weiter oben (in y -Richtung) ein. Vor allem scheint es sich jedoch um einen kontinuierlichen Übergang zu handeln. Dem ist aber entgegenzustellen, dass der Einzelpeak unterhalb des Pfeils auf einem etwa 50 nm langen Stück einen Versatz hin zu höheren Energien aufweist. Dadurch ist eine eindeutige Beurteilung nicht möglich. Dieser Versatz korreliert räumlich mit einer Beule der linken Stufe unterhalb 14 nm-Markierung in Abb. 5.4(a).

In der mittleren Spalte ist der Doppelpeak im Bereich $L > 400$ nm in (e) gar nicht mehr und in (f) nur sehr verschmiert zu erkennen. In der zweiten Ableitung kann der linksgekrümmte, grün-gelbe Bereich hingegen noch gut nachvollzogen werden. Die Aufspaltung nimmt auch für $L > 400$ nm weiter zu, allerdings erfährt nun auch der niederenergetische Peak eine (Rot-) Verschiebung. Bis dahin weicht er kaum von der Dirac-Energie ab und fluktuiert lediglich etwas. Für mittlere und schwache Kopplungen resultierte die Aufspaltung also vornehmlich aus der zunehmenden Blauverschiebung des hochenergetischen Peaks.

Ein weiteres Beispiel für zusammenlaufende Stufenkanten, die sich unter einem kleinen Winkel schneiden, ist in Abb. 5.5(a) gezeigt. Zwei Unterschiede zum zuvor diskutierten Fall sind zum einen die Stufenrichtung (fallend statt steigend) und zum anderen die Tatsache, dass im unteren

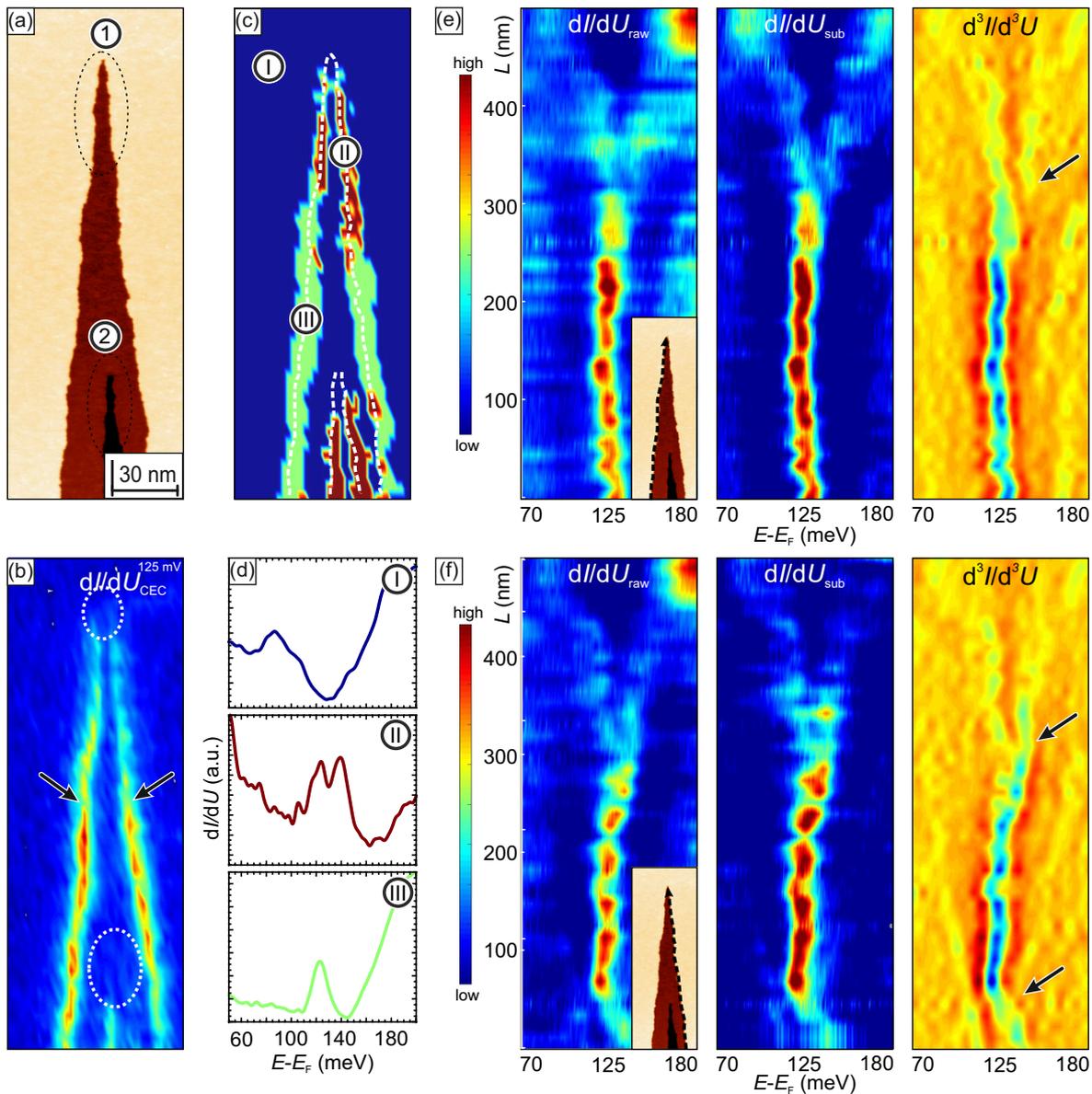


Abbildung 5.5 Kopplungseffekte an zusammenlaufenden Fehlstelleninseln: (a) Topographie von zusammenlaufenden, ungeraden Stufen, die zwei Fehlstelleninseln bilden. (b) Schnitt konstanter Energie aus den STS-Spektren bei der Dirac-Energie. (c) Über das gesamte Bild in (a) wurde eine Rastertunnelspektroskopie durchgeführt. Jedem Spektrum wurde eine Klasse und jeder Klasse eine Farbe zugeordnet: Einfachpeak grün, Doppelpeak rot und kein Peak blau. (d) Exemplarische Spektren für die in (c) farbcodierten Bereiche. (e,f) Spektren entlang der linken (e) und rechten (f) Stufe, entnommen aus dem selben STS-Datensatz. Gezeigt sind links die Rohdaten, in der Mitte die Rohdaten abzüglich eines Hintergrunds und rechts die zweite Ableitung des Spektrums, also d^3I/d^3U . Messparameter: $U = 200$ mV, $I = 10$ pA (a). Stabilisierungsparameter: $U = 200$ mV, $I = 200$ pA (d-f).

Bildteil nochmals zwei Stufen zusammenlaufen. Es handelt sich somit um eine Fehlstelleninsel, womit die Frage geklärt werden kann, ob die Stufenrichtung (hoch, runter) einen Einfluss auf den Doppelpeak hat. Die Kombination aus zwei ineinander liegenden Fehlstelleninseln ist besonders interessant, weil an jeder ungeraden Stufe die Translationssymmetrie durch das Vertauschen des

Pb/Sn- mit dem Se-Untergitter gebrochen ist. Diese führt dann, wie in Kap. 4 diskutiert, dazu, dass zwei benachbarte Terrassen durch zwei unterschiedliche Hamiltonians beschrieben werden und damit zu verschiedenen strukturellen Phasen gehören. Von Interesse ist nun, ob es eine Rolle spielt welche Phase ober- und welche unterhalb der Stufe liegt.

Da diese Phasen in einem STM-Bild makroskopisch ununterscheidbar sind, kann nur für zusammenhängende Bilder eindeutig festgestellt werden, welche Terrassen die selbe strukturelle Phase haben. Für die mittlere Terrasse ist genau das der Fall, da diese bei der Fehlstelleninsel (1) unterhalb und bei der Fehlstelleninsel (2) oberhalb der jeweiligen Stufe liegt. Allerdings liegt die zweite Insel nicht mittig in der ersten, sodass der Abstand rechts unten sehr klein wird. Auf der zuvor festgestellten, relevanten Längenskala von etwa 25 nm sind teilweise drei Stufenkanten zu finden. Dieser Umstand verkompliziert die Auswertung der Spektren für die untere Stufe, sodass sich die Analyse im Folgenden für diese auf die reine Klassifizierung beschränkt.

Anstelle einer dI/dV -Karte, welche nicht gemessen wurde, ist in (b) ein Schnitt konstanter Energie aus den STS-Spektren gezeigt. Hierfür wurde an jedem Punkt des Rasters der dI/dV -Wert des Spektrums bei $U = 125$ mV ausgelesen und graphisch aufgetragen. Die wesentlichen Unterschiede zu einer dI/dV -Karte sind die geringere laterale Auflösung und die etwas höhere Stabilisierungsspannung von $U = 200$ mV. Zu sehen sind auch in diesem Fall die selben Charakteristika wie in Abb. 5.4(b). In den Bereichen wo die Fehlstelleninseln zusammenlaufen nimmt die Intensität drastisch ab und ist kaum von der auf der Terrasse unterscheidbar. Oberhalb von Insel (2) beträgt der Abstand der oberen beiden Stufen mehr als 30 nm. In diesem Bereich besitzen beide Kantenzustände eine hohe Intensität, die mit abnehmendem Abstand schlagartig merklich abnimmt (Pfeile) und dann immer schwächer wird.

Analog zur vorherigen Auswertung wurde auch für diese Stelle jedes STS-Spektrum in eine der drei Klassen eingeteilt. Die entsprechend farbcodierte Darstellung ist in Abb. 5.5(c) und exemplarische Spektren für Einzel-, Doppel und kein Peak in (d) gezeigt. Die geänderte Richtung der Stufe (fallend statt steigend) hat keinen Einfluss auf das Auftreten der Aufspaltung. Der Bereich, in dem zwar die Stufenkanten noch existieren, ihnen jedoch keine Peaks mehr zugeordnet werden können, ist für beide Fehlstelleninseln kleiner als in Abb. 5.4(a). Wahrscheinlich liegt dies jedoch daran, dass dort die Stufen am oberen Ende der Insel ein Stück quasi parallel verlaufen, während sie in Abb. 5.5(a) direkt zusammenlaufen. Aus geometrischer Sicht ist es dem größeren Winkel dann auch zuzuschreiben, dass sowohl der Bereich mit nur Doppelpicks, als auch der Übergangsbereich wo z.T. Einfachpeaks beobachtet werden, kleiner sind.

Für die untere Fehlstelleninsel schließt sich an den peakfreien Bereich direkt der mit Doppelpicks an beiden Stufen an. Hieraus kann die wichtige Schlussfolgerung gezogen werden, dass die Aufspaltung nicht davon abhängt welche strukturelle Phase oberhalb und welche unterhalb der Stufe sitzt. Die Aufspaltung ist somit nicht von der genauen atomaren Besetzung der Untergit-

ter abhängig, sondern ihr müssen andere Mechanismen zugrunde liegen. Für einen topologisch geschützten Zustand erscheint die Feststellung, dass er unabhängig von der exakten atomaren Konfiguration der Oberfläche ist, trivial. Wie aber in Kap. 4.3 gezeigt wurde, sind sich alle theoretischen Modelle einig, dass der Kantenzustand bei der Dirac-Energie entartet ist. Die Aufspaltung in zwei Peaks weist aber darauf hin, dass die fundamentale Symmetrie, die die Entartung garantiert hat, gestört bzw. gebrochen ist. Unter diesen Umständen ist dieser experimentelle Befund als nicht trivial einzustufen.

Für die beiden Stufen der Fehlstelleninsel (1) sind in Abb. 5.5 (e,f) analog zu Abb. 5.4 (e,f) die Rohdaten und prozessierten Spektren entlang der Kanten gezeigt. Die Rohdaten links zeigen auch hier gut die hohe Intensität des Einzelpeaks sowie seine Abwesenheit dort, wo die Stufen konvergieren. In der Mitte wurde wieder der Hintergrund abgezogen und rechts die zweite Ableitung gebildet. An der linken Stufe (e) ist im d^3I/d^3U -Signal gut erkennbar wie erneut der zweite Peak abrupt einsetzt (Pfeil). Auf der selben Höhe setzt auch bei der rechten Stufe (f) der zweite Peak ein. Wie bereits in Abb. 5.4 (e) kommt es auch hier unmittelbar zuvor zu einer Blauverschiebung des Einzelpeaks. Auch in diesem Fall ändert sich in der Topographie der Stufenabstand an der entsprechenden Stelle aufgrund einer Beule und wird zwischenzeitlich größer statt kleiner. Allerdings tritt diesmal die Verschiebung bei der nicht deformierten Stufe auf. Außer der Stufenkantenform kommt für die lokale Schwankung des chemischen Potentials auch die zufällige Verteilung von Pb- und Sn-Atomen als Ursache in Frage. Jedoch sorgt dieser Effekt typischerweise für eine Variation der Dirac-Energie um etwa ± 5 meV. Hier weicht die Position des Peaks aber um 12-15 meV von der mittleren Lage des Dirac-Punkts ab, sodass dieser Mechanismus als alleinige Erklärung unwahrscheinlich ist.

Die Spektren der rechten Kante zeigen erwartungsgemäß aufgrund des kleinen Abstands zur Fehlstelleninsel (2) auch im unteren Bereich ein Doppelpeak. Auf Höhe des unteren Pfeils findet ein diskontinuierlicher Übergang zu einem Einfachpeak statt, wie er auch zuvor bei der linken Stufe beobachtet wurde. Dieser Übergang findet jedoch nicht erst dort statt, wo die untere Fehlstelleninsel verschwindet, sondern ungefähr auf der selben Höhe auf der auch die Intensität der linken Stufe von Fehlstelleninsel (2) in (b) verschwindet. Während der Abstand zur äußeren Stufe etwa 15 nm beträgt, ist er innerhalb der Fehlstelleninsel in diesem Bereich um den Faktor 2-5 kleiner. In Abb. 5.4 (c) ist die Aufspaltung für $d = 7$ nm etwa zweieinhalb mal so groß wie bei $d = 14$ nm. Geht man in erster Näherung davon aus, dass die Aufspaltung proportional zur Kopplung ist, dann führt die wesentlich stärkere Kopplung innerhalb der inneren Fehlstelleninsel zur Auslöschung des Kantenzustands, sodass es für die äußeren Kanten keinen Partner zum Koppeln mehr gibt.

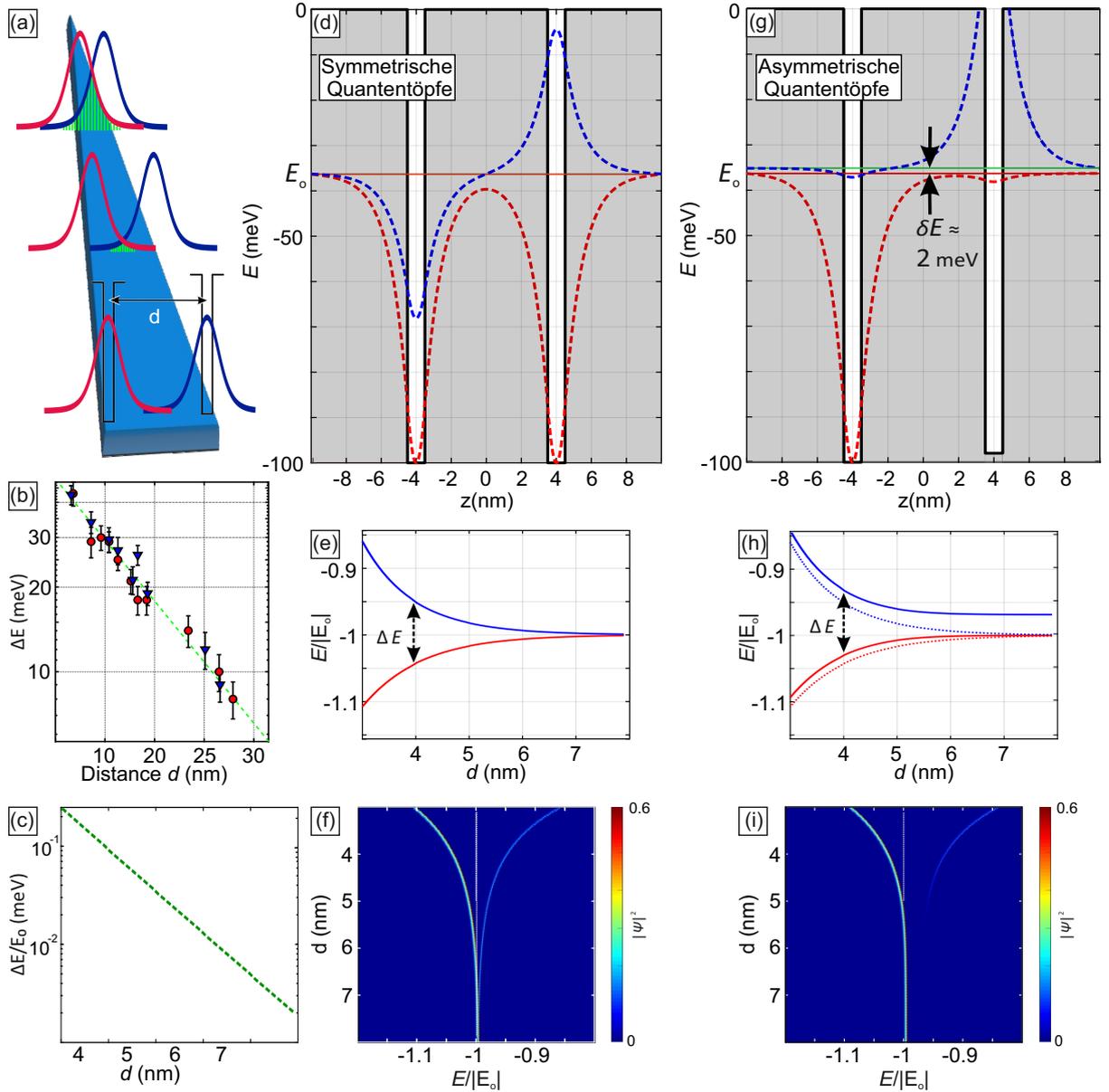


Abbildung 5.6 Doppelquantentopf-Modell: (a) Skizze einer zusammenlaufenden Stufe mit Quantentöpfen im Abstand d . Die Lokalisierung an den Stufen wird durch Gaussfunktionen approximiert und deren Überlappung grün hervorgehoben. (b) Halblogarithmische Auftragung der experimentellen Aufspaltung gegen d mit linearer Anpassung. (c) Gleiche Auftragung der zu erwartenden Aufspaltung basierend auf der Energiedifferenz zwischen anti- und symmetrischer Lösung des Doppelquantentopf-Grundzustands. (d) Skizze zweier identischer Quantentöpfe sowie der anti- (blau) und symmetrischen Lösung (rot). (e) Abstandsabhängige Aufspaltung der beiden Lösungen. (f) Implementierung des Betragsquadrats der Wellenfunktionen in den Graph (e). Anmerkung: Die Intensität der antisymmetrischen Lösung ist aus Gründen der Übersichtlichkeit mit 2 multipliziert worden. (g-i) Analoge Darstellungen zu (d-f) für einen asymmetrischen Doppeltopf. Die Tiefe des rechten Topfs weicht um 2 meV ab. Die Entartung des Grundzustands wird aufgehoben und die Aufspaltung erfolgt nicht mehr um die Dirac-Energie herum. Alle Ergebnisse von (c-i) wurden durch numerisches Lösen der SGL unter folgenden Modellparametern erhalten: Quantentopftiefe $V_0 = 100$ meV, -breite $B = 1$ nm.

Doppelquantentopmodell

Um die Koppelung und ihre Ursache qualitativ besser zu verstehen wird nun ein einfaches Modell herangezogen. Ausgangspunkt ist die experimentelle Beobachtung, dass die Kantenzustände senkrecht zur Stufe stark räumlich beschränkt sind [39]. Eine generelle Grundvoraussetzung für die räumliche Lokalisierung von elektronischen Zuständen ist, dass in irgendeiner Form ein Einschlusspotenzial vorliegt. In diesem Modell wird der Kantenzustand einer einzelnen, freistehenden Stufe durch einen Quantentopf lokalisiert. Zusammenlaufende Stufen entsprechen dann einem Doppelquantentopf mit veränderlichem Abstand d .

Eine Skizze dieses Modells ist in Abb. 5.6 (a) gezeigt, in der zusätzlich noch die räumliche Beschränkung durch Gaussprofile skizziert ist. Die Überschneidung der beiden Gaussfunktionen ist dann ein Maß für die Kopplung, respektive die Größe der Energieaufspaltung. Eine Möglichkeit dieses Problem zu lösen besteht in der Berechnung des Integrals über dem Produkt zweier Gaussfunktionen mit d als freiem Parameter. Die Berechnung der Gesamtfläche und vor allem deren Abhängigkeit von d lässt sich jedoch näherungsweise auch einfacher bestimmen. Da die Überschneidung zunächst noch weit von den Maxima der Peaks entfernt stattfindet, können die Gaussprofile durch ihr exponentielles Abklingverhalten genähert werden. Außerdem ist für zwei identische Gausskurven, die lediglich gegeneinander verschoben sind, die Überlappung symmetrisch bezüglich des Schnittpunkts. Das bedeutet, dass die Flächen rechts und links vom Schnittpunkt gleich groß sind und vor allem dem selben funktionalen Zusammenhang (von d) genügen.

In Abb. 5.6 (a) ist die linke Gaussglocke rot und die rechte blau, sodass die Überlappung links des Schnittpunkts durch den exponentielle ansteigenden Teil der blauen und rechts von dem exponentiell fallenden Teil der roten Kurve begrenzt wird. Mit abnehmendem d wird die blaue Kurve in Richtung der roten verschoben, wodurch der Schnittpunkt auf der roten Kurve weiter wandert. Da die rote Kurve in diesem Bereich durch eine e-Funktion genähert ist, steigt auch die von ihr begrenzte Fläche rechts des Schnittpunkts und damit auch die Gesamtfläche exponentiell an.

Trägt man die experimentell gemessenen Aufspaltungen, wie in Abb. 5.6 (b), halblogarithmisch gegen d auf, so lassen sich die Werte gut durch eine lineare Anpassung beschreiben. Die Aufspaltung nimmt somit tatsächlich wie in dem einfachen Modell der Überlappung zweier Gaussfunktionen exponentiell mit d ab. Der selbe funktionale Zusammenhang ergibt sich, wenn die Energien des Grundzustandes im symmetrischen Doppelquantentopf betrachtet werden. Symmetrisch bedeutet hier, dass die beiden Töpfe die selbe Breite B und Tiefe V_0 haben und ist nicht zu Verwechseln mit der Symmetrie der Wellenfunktionen. Der Grundzustand dieser gekoppelten Quantentöpfe besitzt nämlich, analog zum H_2 -Molekül, eine symmetrische und eine

antisymmetrische Lösung. Deren energetischer Abstand, normiert auf die Grundzustandsenergie E_0 , ist in (c) dargestellt und folgt ebenfalls einem exponentiellen Zusammenhang als Funktion von d .

Eine Skizze für die symmetrischen Quantentöpfe, zusammen mit den Wellenfunktionen des Grundzustands, ist in Abb. 5.6 (d) zu sehen. Die rote Lösung ist bezüglich der Mitte des Doppeltopfs bei $z = 0$ spiegelsymmetrisch und entspricht daher der symmetrischen Lösung des H_2 -Moleküls. Bei der blauen Wellenfunktion muss anschließend an eine Spiegelung noch das Vorzeichen invertiert werden, sodass es sich um die antisymmetrische Lösung handelt. Ein allgemeineres Kriterium stellt die Anzahl an Nulldurchgängen dar, denn eine antisymmetrische Lösung weist eine ungerade und eine symmetrische Lösung eine gerade Anzahl auf.

Um die Energieaufspaltung zwischen symmetrischer und antisymmetrischer Lösung zu bestimmen wurde die SGL für den Doppelquantentopf für $B = 1 \text{ nm}$ & $V_0 = 100 \text{ meV}$ numerisch gelöst. Bereits für den endlichen Einfachquantentopf werden die Energieniveaus durch eine transzendente Gleichung beschrieben, sodass keine analytische Lösung mehr angegeben werden kann. Die auf E_0 normierten Ergebnisse für symmetrische und antisymmetrische Wellenfunktion sind in Abb. 5.6 (e) gezeigt. Bei großen Abständen sind die beiden Zustände energetisch entartet und spalten sich auf, sobald d und B ungefähr die selbe Größenordnung erreichen.

Die Entartung lässt sich qualitativ anhand des (unendlichen) Einfachpotenzialtopfs verstehen. Bei diesem wird der Grundzustand durch eine symmetrische Lösung mit maximaler Aufenthaltswahrscheinlichkeit W in der Mitte und der erste angeregte Zustand durch eine antisymmetrische Lösung mit verschwindendem W in der Mitte beschrieben. Ferner besitzt die Wellenfunktion der Grundmode keine Nulldurchgänge, was sie von allen höheren Energieniveaus unterscheidet. Bei einer isolierten Betrachtung der beiden Quantentöpfe zeigt sich, dass sowohl die rote als auch die blaue Lösung jeweils in beiden Töpfen ein Maximum in der Mitte besitzen. Außerdem besitzt keine der Lösungen eine Nullstelle innerhalb eines Quantentopfes, sodass beide qualitativ den Grundzustand beschreiben.

Im Grenzfall $\lim_{d \rightarrow 0}$ entsteht ein Einfachpotenzialtopf mit doppelter Breite und den zuvor beschriebenen Formen für den Grund- und den ersten angeregten Zustand. Während zuvor der Nulldurchgang der blauen Kurve quasi im unendlichen lag und vernachlässigbar war, ist diese Näherung nun nicht mehr gültig. Stattdessen muss der asymmetrische Charakter berücksichtigt werden, was bedeutet, dass die blaue Kurve dem ersten angeregten Zustand entspricht. Dementsprechend steigt in (e) die Energie der asymmetrischen Lösung an. Die rote Kurve hingegen beschreibt weiterhin den Grundzustand, da mit abnehmendem d auch das Maximum bei $z = 0$ kleiner wird. Dieses entsteht durch die exponentielle Dämpfung der Wellenfunktion im klassisch verbotenen Bereich, welcher gerade die Dicke d besitzt. Die beiden roten Minima wandern zusammen und bilden ein einziges Extremum in der Mitte des doppelt so breiten Einfachquan-

tentopfs, was seinem Grundzustand entspricht. Da die Energie E_0 bei einem Quantentopf von B abhängt und mit dieser sinkt, fällt in (e) die rote Kurve für kleiner werdende d -Werte.

Die Aufspaltung lässt sich also als im Kontext eines kontinuierlichen Übergangs von einem Doppel- in einen Einfachquantentopf verstehen. In zwei weit entfernten Quantentöpfen stellt die global antisymmetrische Lösung lokal eine zweite symmetrische Lösung des Problems dar, welche mit der global symmetrischen energetisch entartet ist. Für kleine Abstände ist die antisymmetrische Lösung auch lokal antisymmetrisch, was dazu führt dass sie sich der Energie des ersten angeregten Zustand eines Einfachpotenzialtopfs annähert.

Anzumerken ist noch, dass im Einfachpotenzialtopf ($L = 2 \cdot B$) der Abstand der ersten beiden Energieniveaus mit $\frac{1}{L^2}$ skaliert. Diese Energiedifferenz ist jedoch nicht die hier gemessene Aufspaltung, sondern entspricht deren Grenzwert für $\lim_{d \rightarrow 0}$. Die hier gemessene Aufspaltung beschreibt, wie schnell sich der entartete Grundzustand diesem Grenzwert annähert. Maßgeblich hierfür ist das exponentielle Abklingen der Wellenfunktionen in dem klassisch verbotenen Bereich. Neben der energetischen Lage von symmetrischer und antisymmetrischer Lösung spielt auch die Besetzungswahrscheinlichkeit des Zustands eine entscheidende Rolle im Experiment. Deswegen ist in Abb. 5.6 (f) das Betragsquadrat der beiden Wellenfunktionen gezeigt. Dabei ist einerseits zu beachten, dass die Energie umso tiefer liegt je negativer der Wert auf der x-Achse ist. Andererseits wurde die antisymmetrische Lösung mit einem Faktor 2 multipliziert um den Kontrast zu erhöhen. Trotzdem weist die symmetrisch Lösung eine höhere Intensität auf.

Dieser Graph ist direkt vergleichbar mit dem d^3I/d^3U -Signal aus den Abb. 5.4 (e,f) & 5.5 (e,f). Während in den Berechnungen die Aufspaltung näherungsweise symmetrisch um E_0 erfolgt, also sowohl die energetisch tiefer liegende, symmetrische Lösung als auch die hochenergetische antisymmetrische Lösung sich verschieben, ergibt sich im Experiment ein anderes Bild. Hier verschiebt zunächst nur die hochenergetische Mode, während die Mode bei niedrigen Energien in guter Näherung bei $E_0 = E_{D^*}$ verweilt. Erst bei sehr kleinen Abständen schiebt auch diese etwas hin zu niedrigeren Energien. Außerdem tritt die Aufspaltung in der Berechnung kontinuierlich auf, wohingegen der zweite Peak im Experiment schlagartig ab einem kritischen d -Wert beobachtbar war.

Um diese Beobachtungen in Einklang mit dem Modell zu bringen wurden weitere Berechnungen mit einem leicht asymmetrischen Doppelquantentopf durchgeführt. Die Asymmetrie bestand dabei in einer kleinen Abweichung der Tiefe des rechten Topfs von $\Delta V_0 = 2 \text{ meV}$. Bereits diese relativ kleine Störung hat signifikante Konsequenzen für die Lösungen des Doppeltopfs. Es finden sich zwar weiterhin eine symmetrische und eine antisymmetrische Lösung (die blaue Kurve hat einen Nulldurchgang und die rote nicht), aber deren Entartung bei größeren Abständen ist aufgehoben.

In der Skizze in Abb. 5.6 (g) beträgt die Energiedifferenz $\delta E \approx 2 \text{ meV}$. Außerdem zeigt sich nun, dass die rote Kurve fast vollständig im linken und die blaue fast ganz im rechten Topf lokalisiert ist. Wird der jeweils kleinere Anteil vernachlässigt, dann kann die rote Kurve wie zuvor qualitativ als Grundzustand des linken und die blaue als Grundzustand des rechten Topfs aufgefasst werden (Grenzfall $\lim_{d \rightarrow \infty}$). Da der rechte Topf etwas flacher ist, liegt einerseits der Referenzpunkt für den Grundzustand um ΔV_0 weiter oben. Andererseits liegt der Grundzustand selbst aber aufgrund des geringeren Werts von V_0 bei einer geringfügig höheren Energie. Deshalb entspricht δE für große d auch nicht exakt ΔV_0 . Durch diese Verschiebung kommt es auch zu keiner Resonanz im jeweils anderen Quantentopf. In (d) wird z.B. ein Zustand bei E_0 aus dem linken Quantentopf durch den intrinsisch vorhandenen Zustand des rechten Topfs verstärkt. In (g) hingegen wird lediglich dessen Dämpfung unterdrückt, da die Wellenfunktion in einen Bereich mit $E > V$ eintritt.

Der zweite Grenzfall $\lim_{d \rightarrow 0}$ lässt sich nun jedoch nur noch näherungsweise durch einen Einfachpotenzialtopf beschreiben. Es ist weiterhin davon auszugehen, dass die antisymmetrische Wellenfunktion in der Energie ansteigt, da sie sich der Energie des ersten angeregten Zustands annähern sollte. Allerdings sind die Details der Aufspaltung durch die Abweichungen von dem idealen Einfachpotenzialverlauf gegeben.

Deshalb sind in Abb. 5.6 (h) die Ergebnisse der numerischen Berechnung für die Aufspaltung analog zu (e) für den asymmetrischen Doppeltopf gezeigt (durchgezogene Linien). Um die Unterschiede zu (e) hervorzuheben sind die Lösungen des symmetrischen Topfs ebenfalls gestrichelt eingezeichnet. Die Verstimmung der Quantentöpfe führt zu zwei markanten Änderungen. Erstens ist die Entartung bei höheren d -Werten aufgehoben und die antisymmetrische Lösung energetisch nach oben verschoben. Zweites liegt die symmetrische Wellenfunktion nun deutlich länger bei E_0 , was insgesamt dazu führt, dass die Aufspaltung nicht einmal mehr näherungsweise symmetrisch um E_0 ist.

Der Graph mit der Aufenthaltswahrscheinlichkeit in Abb. 5.6 (i) zeigt dieses Verhalten ebenfalls. Besonders interessant ist nun aber, dass der antisymmetrische Ast erst ab dem Unterschreiten eines kritischen Abstands eine signifikant von Null verschiedene Amplitude hat. Somit lassen sich beide Phänomene, die asymmetrische Aufspaltung bezüglich der Dirac-Energie und das abrupte Einsetzen des zweiten Peaks, erklären.

Mögliche Ursachen für die Asymmetrie der Quantentöpfe könnten lokale Schwankungen der Pb-Sn-Stöchiometrie, die unterschiedlichen kristallographischen Richtungen oder die Rauigkeit der Stufen sein. Die letzten beiden Effekte sind nicht vollständig unabhängig voneinander, da atomar glatte Stufenkanten nur in Hochsymmetrierichtungen möglich sind. Je nachdem wie kompatibel eine Richtung also mit der Kristallstruktur ist gibt es bereits eine gewisse intrinsisch notwendige Rauigkeit einer Stufe.

5.3.2 Parallellaufende Stufenkanten

Neben dem Doppelquantentopf-Modell gibt es noch weitere Möglichkeiten die Aufspaltung oder deren Asymmetrie bezüglich der Dirac-Energie zu erklären. Besonders die Asymmetrie fehlt dabei in den verschiedenen Modellen zunächst oft, wie z.B. bei der Aufspaltung eines entarteten Zustands in ein quantenmechanisches Zwei-Niveau-System, bei zwei gekoppelter Oszillatoren oder bei dem zuvor verwendeten Doppelquantentopf. Erst durch die grundlegende Annahme eines leicht asymmetrischen Problems wird sie beobachtbar. So reproduzieren beispielsweise auch die gekoppelten Oszillatoren die Asymmetrie erst, wenn die Eigenfrequenzen verstimmt werden. Im Gegensatz zu dem Doppelquantentopf mangelt es diesem Modell jedoch an physikalischen Anknüpfungspunkten und somit an Aussage- und Vorhersagekraft. Dabei geht es um Fragen wie: Welche Bedeutung haben die Massen und Federkonstanten der einzelnen Oszillatoren? Was verursacht deren Kopplung? Warum ist zu erwarten das ein rein klassisches Modell die quantenmechanische Natur von elektronischen Zuständen gut beschreiben soll?

Eine vernünftige physikalische Interpretation für die Asymmetrie liefert hingegen eine weitere Erklärungsansatz, der auf der Arbeit von Smoluchowski zur Anisotropie der Austrittsarbeit von Metallen aus dem Jahr 1941 basiert [56]. Grundgedanke dieser Arbeit ist zum einem die Feststellung, dass unterschiedliche Oberflächenorientierungen aufgrund der Kristallstruktur zu verschiedenen Stufendichten führen. Zum anderen ergibt sich durch Änderungen der Ladungsdichte an der Oberfläche (an den Stufen) ein Polarisierung, die wiederum die Austrittsarbeit modifiziert. Der erste Effekt entspricht einfach dem 3D-Analogen zur Rauigkeit von Stufen unterschiedlicher Orientierung. Der zweite Effekt hingegen führt zu einem Überschuss an positiven Ladungen auf der oberen und zu einer negativen Raumladung auf der unteren Terrasse unmittelbar an der Stufe.

Da bei zwei benachbarten Stufen diese Ladungen miteinander wechselwirken können, sollte ein struktureller Effekt auf die Asymmetrie feststellbar sein. Die Ursache liegt darin, dass bei zwei aufeinander fallenden (steigenden) Stufen die Ladungsreihenfolge $p-n-p-n$ ($n-p-n-p$) entsteht, während eine Kombination aus fallender und steigender Stufe zu einer $p-n-n-p$ bzw. $n-p-p-n$ Reihenfolge führt. Die beiden mittleren Ladungen gehören zu verschiedenen Stufen und wechselwirken im ersten Fall attraktiv und im zweiten repulsiv, sodass eine Rot- bzw. Blauverschiebung zu erwarten ist.

Eine Topographie mit zwei benachbarten Stufen der selben Richtung ist in Form der bisher untersuchten, keilförmigen Strukturen extrem selten zu finden. Häufiger als beispielsweise eine Doppelstufe die in zwei Einfachstufen aufspaltet, sind Stufen mit annähernd paralleler Orientierung zu finden. Hier besteht das Problem viel mehr darin, dass sich deren Abstand oft nur wenig über viele 100 nm ändert. Die überwiegende Mehrheit der Stufen besitzt jedoch einen zu großen

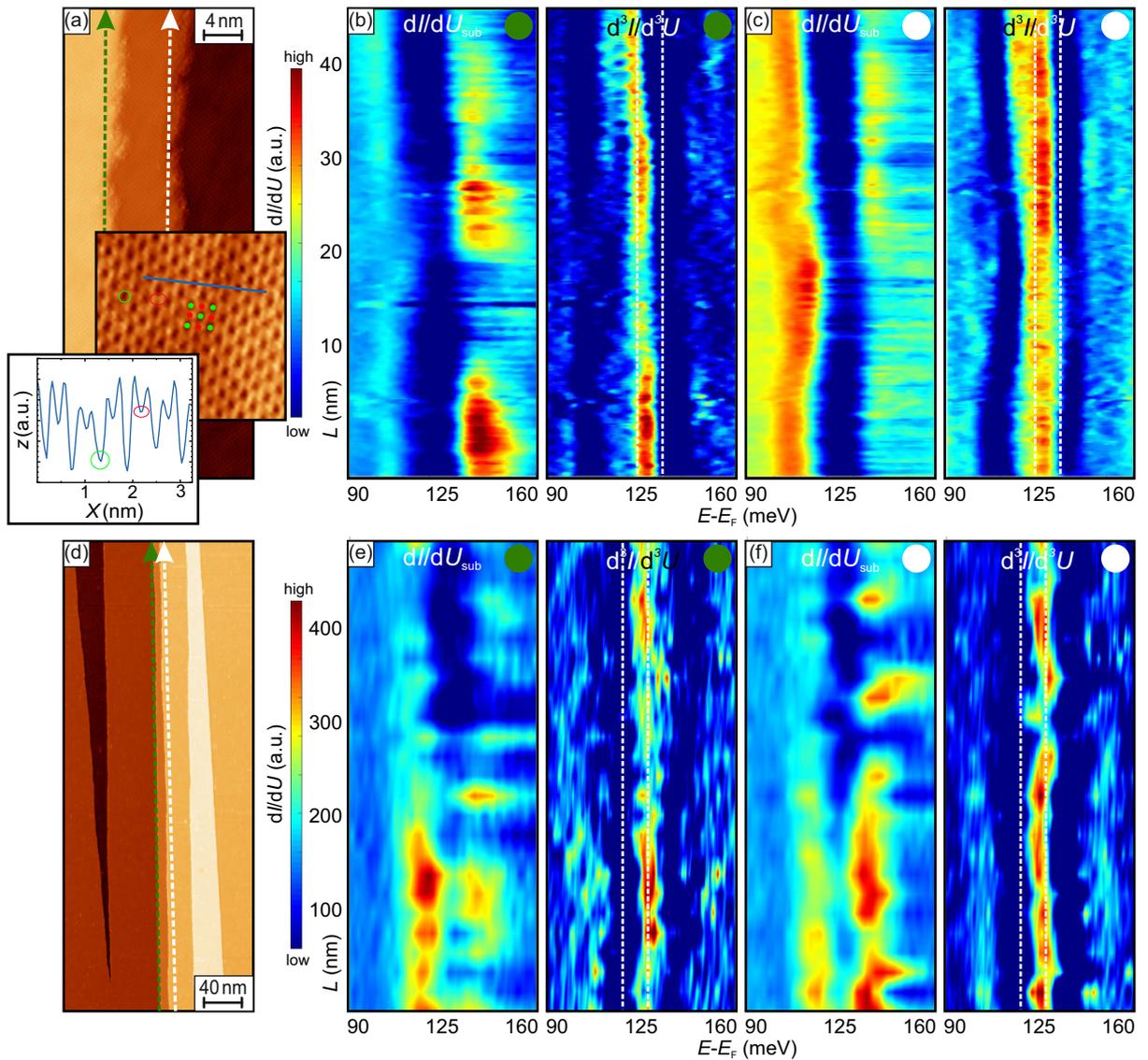


Abbildung 5.7 Parallel verlaufende, gekoppelte Stufenkanten: (a) $(20 \times 50) \text{ nm}^2$ Topographie zweier paralleler Stufen mit atomarer Auflösung $[(5 \times 5) \text{ nm}^2]$ und dazugehörigem Linienprofil als Teilbilder. (b) dI/dV - und d^3I/d^3U -Signal aus STS-Messungen entlang der linken Stufe. Von den Rohdaten wurde der Hintergrund abgezogen. Die dunklen Bereiche im d^3I/d^3U -Signal korrespondieren mit den Peaks. Die weiß gestrichelten Linien markieren den Energiebereich des Dirac-Punkts, basierend auf Referenzspektren an umgebenden, freistehenden Einfachstufen. (c) Analoge Darstellung zu (b) für die rechte Stufe, basierend auf dem selben Datensatz. (d) Weitere Topographie paralleler Stufen, aber mit zehnfach größerer Skala. (e,d) Äquivalente Darstellung der rechten und linken Stufe zu (b,c). Die Aufspaltung des Doppelpeaks unterscheidet sich, da oben der Peak bei höheren und unten der Peak bei niedrigeren Energien nahe der Dirac-Energie verbleibt. Messparameter: $U = 500 \text{ mV}$, $I = 50 \text{ pA}$ (a); $U = 180 \text{ mV}$, $I = 10 \text{ pA}$ (d). Stabilisierungsparameter: $U = 200 \text{ mV}$, $I = 250 \text{ pA}$ (b,c); $U = 180 \text{ mV}$, $I = 200 \text{ pA}$ (e,f).

Abstand um den Doppelpeak beobachten zu können und unter denen mit passendem Abstand muss es sich bei beiden um eine ungerade Stufenhöhe handeln.

In Abb. 5.7 sind zwei Stellen gezeigt, die all diesen Restriktionen genügen. Die beiden gezeigten Topographien unterscheiden sich in der Längenskala um eine Größenordnung. Mit einer Größe

von $(20 \times 50) \text{ nm}^2$ ergibt sich mit 512×512 Bildpunkten für den kleineren der beiden in Abb. 5.7(a) eine Auflösung im Bereich von \AA . Das Teilbild in der rechten oberen Ecke zeigt einen $(5 \times 5) \text{ nm}^2$ großen Ausschnitt der rechten Terrasse mit atomarer Auflösung. Interessanterweise sind die Atome nicht als Erhöhungen zu sehen, sondern als Vertiefungen. Dieser Effekt lässt sich durch ein kegelförmiges Orbital am vordersten Atom der Spitze erklären. Dieser unkonventionellen Terminierung ist wahrscheinlich auch die unsaubere Erscheinung der Stufenkanten geschuldet.

Obwohl die Spitze demnach anscheinend suboptimal für STM-Messungen ist, sind im Linienprofil trotz verhältnismäßig großer Tunnelspannung und relativ geringem Tunnelstrom beide Untergitter erkennbar. Die bereits diskutierten Minima sind gut als globale Minima (grün) zu sehen, werden jedoch nicht von einem einzelnen Maximum sondern von einem Doppelpeak voneinander getrennt. Zwischen diesen Doppelpeaks liegt jeweils nochmal ein lokales Minimum (rot). Eine Analyse der Abstände zeigt, dass sowohl die grünen als auch die roten Extrema jeweils eine Periodizität von etwa 6 \AA haben. Außerdem ist der mittlere Abstand zwischen roten und grünen Minima im Rahmen des Fehlers halb so groß, was einerseits bedeutet dass das Linienprofil entlang der $\langle 100 \rangle$ -Richtung orientiert ist und andererseits nur erklärbar ist, wenn sowohl das Se als auch das Pb/Sn-Untergitter zu sehen ist.

Linienpektroskopien für die linke und rechte Stufenkante sind in Abb. 5.7(b) und (c) dargestellt. Zu sehen sind jeweils das dI/dV -Signal nach Subtraktion des Untergrunds und dessen zweite Ableitung. Dabei ist an der rechten Stufe tendenziell der hoch- und an der linken Stufe der niederenergetische Peak besser im dI/dV -Signal zu sehen.

Im Gegensatz zu den zusammenlaufenden Stufen bleibt der Abstand und deswegen auch die Kopplung zwischen den Kantenzuständen nahezu konstant. Folglich variiert die Position der beiden Peaks, bis auf leichte Schwankungen, nicht. Daher ist es in diesen Datensätzen unmöglich direkt mit der selben Spitze die Lage des Dirac-Punkts auch für entkoppelte Stufen zu messen. Diese Referenzmessungen wurden deshalb vor bzw. nach den STS-Messungen durchgeführt. Wie die Position der Doppelpeaks auch, schwankt die Dirac-Energie leicht durch zufällige Variationen der lokalen Pb-Konzentration. Die weiß gestrichelten Linien in den d^3I/d^3U -Karten begrenzen deshalb den Bereich, in dem der Dirac-Punkt zu erwarten ist. Ausgangspunkt dieser Schätzung ist die mittlere gemessene Energie des Dirac-Punkts plus bzw. minus der beobachteten energetischen Streuung.

Die so ermittelten Energien passen gut zur den Bereichen mit maximaler Krümmung in den d^3I/d^3U -Signalen, wie es zumindest teilweise für zusammenlaufende Stufen bei sehr kleinen Abständen ($d \lesssim 6 \text{ nm}$) ebenfalls der Fall ist. Die Minima hingegen liegen bei den beiden Energien der Doppelpeaks und weisen erneut eine Asymmetrie bezüglich des Dirac-Punkts auf. Diesmal ist jedoch der niederenergetische Peak stärker verschoben, was der im Smoluchowski-

Modell zu erwartenden Rotverschiebung entspricht. Um diese Beobachtung zu verifizieren ist in Abb. 5.7(d) eine weitere Topographie mit näherungsweise paralleler Stufenkanten gezeigt. Aufgrund der um eine Größenordnung größeren Längenskala ist der genaue Abstand beider Stufen nur schwer direkt erkennbar. Die Übersichtsaufnahme wurde jedoch deshalb gewählt, um auf die problematische Nähe der zusammenlaufenden Insel auf der rechten Terrasse eingehen zu können. Ihr Abstand zu der weiß markierten Stufe beträgt an der engsten Stelle lediglich etwa 25 nm, sodass eine Kopplung gerade so noch ausgeschlossen bzw. vernachlässigt werden kann. Der Abstand zwischen den beiden parallelen Stufen ist mit etwa 9 nm etwas größer als in (a). Wegen der Ergebnisse der zusammenlaufenden Stufen ist daher damit zu rechnen, dass der Dirac-Punkt bei etwas niedrigeren Energien als die maximale Krümmung des d^3I/d^3U -Signals liegen sollte.

Die Teilbilder (e) und (f) zeigen die analogen Daten zu (b) und (c) auch für diese Kantenzustände. Auch hier sind, wenn auch weniger ausgeprägt, der nieder- bzw. hochenergetische Peak an jeweils einer der Stufen etwas besser zu erkennen. Auch für diesen Datensatz wurde anhand von Referenzmessungen an Einfachstufen in der Umgebung die Lage des Dirac-Punkts bestimmt und markiert. Im Vergleich zu (b) und (c) liegt nun tatsächlich der Dirac-Punkt bei etwas niedrigeren Energien als das Maximum im d^3I/d^3U -Signal. Dadurch überlappt dieser Energiebereich partiell mit dem linken Minimum, sodass in diesen Fall eindeutig der Peak bei niedrigen Energien kaum verschoben ist. Damit ist das entgegengesetzte Verhalten zu (a-c) zu sehen, was im Widerspruch zu dem Smoluchowski-Modell steht und diese Erklärung daher ausgeschlossen werden kann.

5.3.3 Kreuzende Stufenkanten

Bei den bisher untersuchten Strukturen ist der Winkel zwischen den beiden Stufen entweder sehr klein, etwa $5-10^\circ$, oder sogar verschwindend gering, wie bei den zuletzt untersuchten, parallelen Stufenkanten. Da die Orientierung der Stufe auch die Richtung des Impulsvektors des 1D-Kantenzustandes vorgibt, ist in allen bisher untersuchten Beispielen die Wechselwirkung durch annähernd parallele Wellenvektoren gegeben. Andererseits ist bekannt, dass der Wellenvektor in der Festkörperphysik eine besondere Rolle einnimmt, wie beispielsweise seine entgegengesetzten Anordnung bei Cooper-Paaren oder auch seine typische Kopplung an den Eigendrehimpuls in topologischen Materialien. Insbesondere die Spin-Impuls-Kopplung (Spin-Momentum Locking, SML) ist für die eindimensionalen, topologischen Kantenzustände experimentell weder bestätigt noch widerlegt worden. Um den Wellenvektor zu variieren werden deshalb im Folgenden nun Stufenkanten mit größeren Öffnungswinkeln untersucht.

Die Topographie einer typischen Spaltprobe weist auf der Längenskala eines STM-Bildes (0.1-1 μm) eine Vielzahl von Terrassen unterschiedlicher Breite auf. Diese werden überwiegend von

Einfach- und Doppelstufen voneinander getrennt. Einige dieser Stufen laufen zusammen, die anderen sind in der Regel grob in die selbe Richtung orientiert. Das Verhältnis zwischen eher parallelen und zusammenlaufenden Stufen variiert dabei auf Längenskalen von einiger μm und insbesondere in der Nähe größerer Bruchstellen. Bei solchen sind oft mehrere Stufen, zum Teil einige Nanometer hoch, in nahezu senkrechter Richtung zu der parallelen Orientierung zu beobachten. Diese verlaufen meistens geradlinig und über sämtliche Terrassen hinweg, was auf eine Spaltung entlang einer Hochsymmetrierichtung hindeutet.

In der näheren Umgebung solcher Buchstellen ist daher die Wahrscheinlichkeit besonders hoch, Stufen mit großem Winkel zueinander oder sogar senkrechter Orientierung zu finden. Im Gegensatz zu den bisherigen Fällen kreuzen sich jedoch die Stufen, sodass sich die Wechselwirkung auf einen kleinen räumlichen Bereich um den Kreuzungspunkt herum beschränkt. Aus geometrischer Sicht ist zu erwarten, dass der Bereich mit Doppelpeak umso kleiner ausfällt, je näher der Winkel bei 90° liegt, weil bei gleichem Abstand vom Schnittpunkt der Stufe–Stufe–Abstand bei einem rechten Winkel am größten ist. Außer den Abstand beeinflusst der Winkel auch die Projektion der Wellenvektoren aufeinander, also den parallelen Anteil des Impulses. Die zu klärende Frage lautet daher, ob der Einfluss auf die Wechselwirkung nur von dem Abstand oder von beiden Faktoren abhängt und gegebenenfalls welcher Beitrag dominiert.

Exemplarisch für verschiedene Winkel sind in Abb. 5.8 die Ergebnisse für etwa 90° (a) und 60° (d) gezeigt. Wie es für eine senkrechte Orientierung zu erwarten ist, verlaufen die Stufen entlang von Hochsymmetrierichtungen $\langle 100 \rangle$ und $\langle 010 \rangle$. Für den 60° -Winkel verläuft nur eine der Stufen parallel zu einer Hochsymmetrierichtung. Die Schnitte konstanter Energie sind aus STS-Messungen extrahiert und in den Teilbildern (b) und (e) gezeigt. Das dI/dV -Signal visualisiert die Intensitätsverteilung bei der jeweiligen Dirac-Energie von $E_{\text{DP}^*} = 130 \text{ meV}$ (b) bzw. $E_{\text{DP}^*} = 85 \text{ meV}$ (e). Die farbcodierten Spektren entlang der beiden Stufen sowie das entsprechende d^3I/d^3U -Signal sind in (c) und (f) dargestellt.

Am Kreuzungspunkt selbst ist in den dI/dV -Karten in keinem der beiden Fälle eine nennenswerte Intensität festzustellen, die sich signifikant von den Terrassen abhebt. In diesem Bereich existiert also kein Kantenzustand. Die Größe der Teilbilder in (b) und (e) unterscheidet sich um einen Faktor 2.5, sodass der Eindruck entstehen kann, dass der peakfreie Bereich im Fall des 90° -Winkels größer ist. Tatsächlich ist jedoch das Gegenteil der Fall, weil die weiß gestrichelten Kreise in beiden Bildern exakt einen Durchmesser 10 nm besitzen. In Abb. 5.8(b) markiert dieser Kreis sehr gut die vier Punkte ab denen die Intensität der jeweiligen Kantenzustände fast vollständig verschwunden ist. Bei dem 60° -Winkel in (e) hingegen besitzt ein äquivalenter Kreis (lila) einen etwas größeren Durchmesser von etwas weniger als 14 nm.

Die Abwesenheit des Kantenzustands ist auch in dem Linienprofil in (c) für beide Stufen sehr gut zu erkennen. Interessanterweise zeigt sich weder direkt im dI/dV -Signal noch in dessen

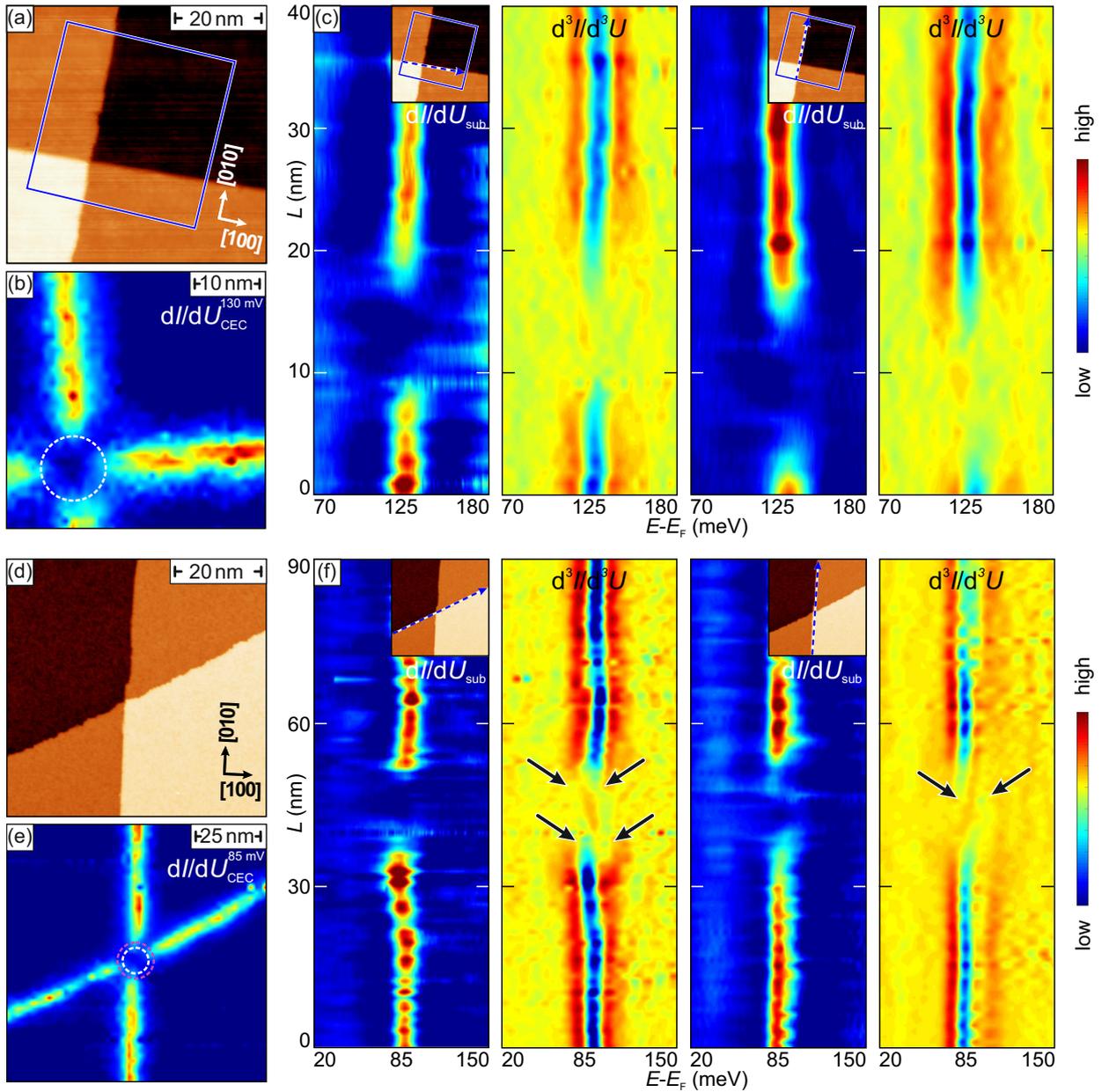


Abbildung 5.8 Kreuzende Stufenkanten: (a) Topographie zweier rechtwinkliger, sich kreuzender Stufen. (b) Schnittbild konstanter Energie über den in (a) markierten Bereich, extrahiert aus STS-Messungen bei $E_{D^*} = 130$ meV. Am Kreuzungspunkt verschwindet die Intensität in einem 10 nm Durchmesser fast vollständig. (c) dI/dV - und d^3I/d^3U -Signal entlang der jeweiligen Stufen. Es gibt keine Anzeichen für eine Aufspaltung, stattdessen geht die Peakintensität kontinuierlich zurück. (d) STM-Topographie für sich kreuzende Stufen unter einem Winkel von etwa 60° . (e) Schnittbild konstanter Energie, extrahiert aus STS-Messungen bei $E_{D^*} = 85$ meV. Der weiße Kreis ist von identischer Abmessung wie in (b), der Bereich stark verminderter Intensität ist jedoch mit etwa 14 nm Durchmesser größer (lila). (f) dI/dV - und d^3I/d^3U -Signal entlang der jeweiligen Stufen. In der höheren Ableitung ist an den markierten Stellen ein zweites Minimum, also ein Doppelpeak zu erkennen. Messparameter: $U = 120$ meV, $I = 200$ pA (a); $U = 1$ V, $I = 10$ pA (d). Stabilisierungsparameter: $U = 200$ mV, $I = 200$ pA (c); $U = 185$ mV, $I = 200$ pA (f).

Ableitung irgendein Hinweis auf eine Aufspaltung der Kantenzustände. Selbst in dem besonders sensitivem d^3I/d^3U -Bild gibt es kein Anzeichen, dass sich ein zweites Minimum ausbildet. Stattdessen gleicht sich die Intensitätsverteilung immer weiter an das Hintergrundrauschen an, je kleiner der Abstand zum Kreuzungspunkt wird.

Bei der $\langle 100 \rangle$ -Stufe lässt sich unterhalb des Schnittpunkts außerdem eine leichte Tendenz zu einer Rechtskrümmung, hin zu höheren Energien, feststellen (vgl. Abb. 5.8(c), zwei Spalten). Diese Krümmung betrifft sowohl das Minimum (türkisfarbener Bereich) als auch die beiden benachbarten Maxima (orange Bereiche). Bei den zuvor vorgestellten, zusammenlaufenden Stufen konnte zumindest an jeweils einer Kante ein ähnliches Verhalten beobachtet werden. Kurz bevor es zur Aufspaltung kommt, verschiebt sich der Peak erheblich zu höheren Energien. Möglicherweise handelt es sich also auch hier um ein erstes Anzeichen für eine Aufspaltung, allerdings wird die Intensität des Peaks sehr schnell gedämpft. Dadurch kommt es gar nicht erst zu einer signifikanten energetischen Verschiebung des Peaks und auch nicht zu der anschließenden Aufspaltung.

Bei der Betrachtung des dI/dV -Signals ergibt sich für den kleineren Winkel in (f) zunächst ein ähnliches Bild. Die Zustände an den Kanten verblassen je näher sie dem Kreuzungspunkt kommen und kein Doppelpeak ist erkennbar. Unterschiede zeigen sich jedoch deutlich in der d^3I/d^3U -Darstellung, wo am Übergang zum Schnittpunkt ein zweites, relativ schwaches Minimum an drei der vier Stufen zu sehen ist. Die entsprechenden Stellen sind mit Pfeilen markiert. Die vierte Stufe weist hingegen keinen Doppelpeak auf, zeigt jedoch eine Rechtskrümmung. Im Gegensatz zu der leichten Krümmung im 90° -Fall kommt es hier auch zu einer signifikanten Energieverschiebung. In der Gesamtheit der Daten erscheint es daher plausibel dieses Phänomen als das Einsetzen einer durch Koppelung induzierten Aufspaltung zu interpretieren.

Aus der vollständigen Unterdrückung der Aufspaltung im Falle senkrecht kreuzender Stufenkanten lässt sich schließen, dass durch den eingeschlossenen Winkel nicht nur der geometrische Aspekt des Stufe-Stufe-Abstands beeinflusst wird, sondern dass die quantenmechanische Natur der Kantenzustände eine Rolle spielt. Das beobachtete Verhalten lässt sich erklären, wenn die Zustände durch orthogonale Wellenfunktionen beschrieben werden, sodass $\langle \Phi_n | \Phi_m \rangle = \delta_{n,m}$ gilt. Hier steht $\delta_{n,m}$ für das Kronecke-Delta, welches nur für identische Indizes einen von Null verschiedenen Wert annimmt. Im Falle quantenmechanischer Zustände symbolisieren die Indizes verschiedene Quantenzahlen, von denen folglich mindestens eine von der Stufenorientierung abhängig sein muss.

Zwei Möglichkeiten diese Bedingung zu erfüllen sind entweder der Spin oder der orbitale Charakter, auf dem bei atomaren Zuständen der Bahndrehimpuls basiert. Die Abhängigkeit des Spin von der Ausbreitungsrichtung ist für D -dimensionale topologische Materialien durch die charakteristische Spin-Impuls-Kopplung ihrer $D - 1$ -dimensionalen Oberflächenzustände tri-

vialerweise sofort ersichtlich. Ob dieser Zusammenhang jedoch auch für topologische Zustände in $D - 2$ gilt, ist theoretisch umstritten, wie in Kap. 4 gezeigt wurde. Auf experimenteller Seite existiert lediglich die Beobachtung von Sessi *et al.*, dass eine Fourier-Transformation keine Anzeichen für Rückstreuung aufweist. Dieses Ergebnis ist zwar für spinpolarisierte Kantenzustände zu erwarten, belegt jedoch deren Existenz aufgrund möglicher anderer Erklärungen nicht. Vielmehr handelt es sich hierbei um ein Widerspruchsexperiment, bei dem die Annahme einer Spinpolarisation widerlegt werden kann, sofern es zu Rückstreuung kommt. Aufgrund dieser unklaren Datengrundlage wird der Spin hier als Erklärungsansatz verworfen. Zunächst sind noch weitere Untersuchungen nötig um die Rolle der Spin-Bahn-Kopplung auf topologische Zustände der Ordnung $D - d$ ($D \geq d \geq 2$) zu untersuchen.

Im Gegensatz dazu ist der orbitale Charakter der Dirac-Kegel intensiv theoretisch untersucht worden und liefert eine fundierte Argumentationsgrundlage. Zunächst sei noch angemerkt, dass hier nicht direkt vom Bahndrehimpuls gesprochen wird, weil diese Größe zwar für die lokalisierten Energieniveaus eines Atoms eine gute Quantenzahl ist, nicht aber für die delokalisierten Zustände im Valenz- oder Leitungsband eines Festkörpers. Für senkrechte Stufenorientierung liegen die beteiligten Kantenzustände folglich auch orthogonal im k -Raum. Daher ist der eine Kantenzustand um den \bar{X} - und der andere um den \bar{Y} -Punkt lokalisiert. An beiden Punkten tragen die Se-Orbitale zum energetisch nach oben verschobenem und die Sn-Orbitale zu dem energetisch nach unten verschobenem Dirac-Kegel bei. Durch die 90° -Symmetrie des Gitters auf der einen Seite und der 180° -Antisymmetrie der beteiligten p -Orbitale auf der anderen Seite wird ersichtlich, dass an den Punkten \bar{X} und \bar{Y} verschiedene p -Orbitale, nämlich p_x und p_y , an der Ausbildung der Dirac-Kegel beteiligt sind. Es ergibt sich daher automatisch, dass die entsprechenden Kantenzustände auf orthogonalen Orbitalen basieren.

Nachdem nun festgestellt wurde, dass die Kantenzustände entlang der $\langle 100 \rangle$ - und $\langle 010 \rangle$ -Richtung auf orthogonalen Orbitalen basieren und (möglicherweise) daher nicht miteinander wechselwirken können, stellt sich die Frage, warum am Kreuzungspunkt trotzdem die Kantenzustände verschwinden. Da dieses Phänomen offenbar nicht an die wechselwirkungsinduzierte Aufspaltung gekoppelt ist, liegt es nahe, dass ein hierarchisch höherstehendes Prinzip dahinter steckt. Da es sich hier um topologisch geschützte Oberflächenzustände handelt, erscheint es wahrscheinlich, dass in diesem Fall die zugrundeliegenden Symmetrien eine Schlüsselrolle einnehmen. Auch wenn alle theoretischen Beschreibungen ein PHS des Hamiltonians annehmen um den 1D-Zustand topologisch zu charakterisieren, so wird die Physik in diesem Materialsystem doch von der Kristallsymmetrie bestimmt. Hinsichtlich der Kristallstruktur liegt am Kreuzungspunkt auch tatsächlich eine fundamental andere Situation vor. An diesem Punkt existiert eine Doppelstufe, die durch Translationsinvarianz in alle Richtungen beschrieben wird, während an den umgebenden Einfachstufen diese Invarianz nur entlang der Kanten gegeben ist. Unter der Voraussetzung

das dieser Unterschied zu distinkten topologischen Klassen führt, kann der Schnittpunkt als 0D-Defekt eines 1D topologischen Zustands aufgefasst werden. Folglich existiert an diesem Punkt kein topologischer Schutz für den Kantenzustand, sodass dieser verschwindet. Um die These der Schlüsselrolle der kristallinen Symmetrie auf die Existenz des Kantenzustands weiter zu analysieren, werden in dem folgenden Kapitel einige intrinsische Kristalldefekte untersucht.

5.4 Kristalldefekte

Jede Störung der Translationssymmetrie eines Kristallgitters führt zur Bildung eines Defekts. Dabei kann mittels der räumlichen Ausdehnung eine Klassifizierung vorgenommen werden, je nachdem in wie viele Richtungen die kristalline Ordnung gebrochen ist. Dreidimensionale Defekte wie Einschlüsse von Fremdmaterial oder Poren sind ausschließlich im Inneren eines Kristalls zu finden. Zweidimensionale Defekte hingegen sind allgegenwärtig in STM-Messungen, weil nur Oberflächen experimentell zugänglich sind und eben diese 2D-Defekte der 3D-Kristallstruktur repräsentieren. Ohne es explizit zu erwähnen wurde in dieser Arbeit bisher sogar fast ausschließlich die elektronische Struktur von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ an Defekten untersucht, da Stufenkanten wiederum 1D-Störstellen der 2D-Oberfläche sind. In diesem Abschnitt sollen nun weitere 1D-Defekte, wie Versetzungslinien oder Schraubversetzungen untersucht werden. Konventionelle, punktförmige Gitterdefekte wie Leer- oder Substitutionsstellen sind hier nicht von Interesse, weil deren räumliche Ausdehnung klein im Vergleich zu den relevanten Längenskalen des Kantenzustands ist. Zum einen sind die Zustände selbst bereits auf mehrere Nanometer um die Stufen lokalisiert und zum anderen setzt deren Kopplung bereits bei etwa 25 nm ein. Für einzelne Defekte ist daher nur ein kleiner oder gar komplett vernachlässigbarer Einfluss auf die elektronische Struktur zu erwarten, solange die Defektdichte hinreichend niedrig ist.

Vergrabener Liniendefekt

Der erste hier untersuchte Liniendefekt ist in Abb. 5.9(a) als schwache Erhebung zu sehen, kreuzt eine ungerade Stufe und verschwindet dann. Es ist unklar ob der Defekt an diesem Punkt seinen Ursprung hat oder ob er an dieser Stelle abknickt und weiter unterhalb der Oberfläche fortläuft, sodass er nicht mehr zu sehen ist. Es handelt sich auf jeden Fall um einen atypischen 1D-Defekt, weil Versetzungslinien normalerweise bis zur Oberfläche gehen oder geschlossene Pfade bilden. Beides kann hier jedoch nicht beobachtet werden, sodass eine genaue Klassifizierung nicht möglich ist. Besonders auffällig ist aber die Tatsache, dass die Versetzungslinie nicht entlang einer Hochsymmetrierichtung zu verlaufen scheint, wie es bei der geradlinigen, glatten Stufe, die sie kreuzt, der Fall ist. Insbesondere Schraubversetzungen, die im zweiten Teil dieses Abschnitts

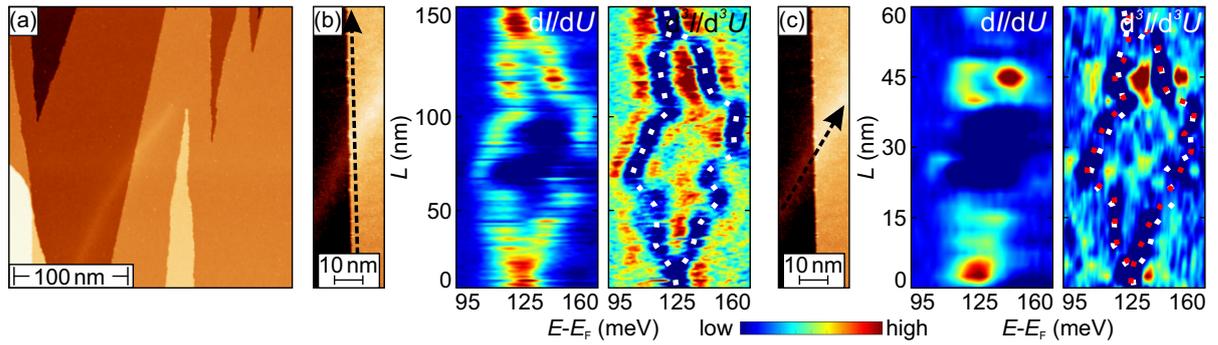


Abbildung 5.9 Kreuzung zwischen ungerader Stufenkante und Versetzungslinie: (a) Übersichtsaufnahme einer Spaltprobe mit einer schwachen Erhebung, die durch eine Versetzungslinie verursacht wird. (b) V.l.n.r.: Großaufnahme des mittels STS abgerasterten Bereichs, dI/dV - und d^3I/d^3U -Signal entlang der Stufenkante. Die weiß gestrichelte Linie veranschaulicht die Lage der Minima respektive der beiden Peaks. (c) Analoge Darstellung zu (b) für die Versetzungslinie. Die rote Linie veranschaulicht die Lage der Minima respektive der beiden Peaks, die weiße Linie gibt den reskalierten Verlauf aus (b) wieder. Die beiden Verläufe stimmen qualitativ sehr gut überein, quantitativ klingt die Aufspaltung bei der Versetzungslinie jedoch 2.5-fach schneller ab. Messparameter: $U = 775$ mV, $I = 10$ pA (a); $U = 200$ mV, $I = 20$ pA (b). Stabilisierungsparameter: $U = 175$ mV, $I = 200$ pA.

gezeigt und in nennenswerter Häufigkeit beobachtbar sind, weisen solche eine Orientierung entlang einer Hochsymmetrierichtung auf.

Ein möglicher Erklärungsansatz ist, dass es um eine Schraubversetzung mit einem wenigen Nanometer dickem Film darauf handelt. In diesem Fall würde die halbunendliche Stufenkante der Schraubversetzung nur als schwache Erhebung zu sehen sein und ihre Orientierung kann von der des sichtbaren Films abweichen, wenn das Wachstum zwischen Film und Substrat nur teilweise oder gar nicht epitaktisch ist. Allerdings bleibt die Frage offen, warum der Verlauf gekrümmt und nicht geradlinig ist.

Unabhängig von seiner genauen Beschaffenheit handelt es sich bei diesem Defekt um ein Unikat, dessen Beobachtung einen Glücksfall darstellt. Bedauerlicherweise fand seine Entdeckung zu einem frühen Zeitpunkt dieser Studie statt, zu dem die genauen Mechanismen, die zur Aufspaltung des Kantenzustands führen, noch unklar waren. Daher ist der Bereich für die STS-Messungen, der in Abb. 5.9 (b) zu sehen ist, mit 25 nm um die Stufe herum sehr knapp bemessen. Die STS-Spektren entlang der Stufenkante zeigen am unteren und oberen Bildrand ein ausgeprägtes Maximum bei $E_{D^*} = 125$ meV, welches um den Kreuzungspunkt herum o-förmig in ein Doppelpeakfeature aufspaltet. Die Aufspaltung erreicht dabei Werte von mehr als 40 meV und ist damit genauso groß wie bei den zusammenlaufenden Stufenkanten mit einem Abstand von $d \approx 7$ nm. Die weiß gestrichelte Linie im d^3I/d^3U -Signal veranschaulicht nochmals die Lage und den Verlauf der Doppelpeaks.

Durch den gekrümmten Verlauf der Versetzungslinie kann der spektroskopische Verlauf über etwa 60 nm verfolgt werden, wobei der maximale Abstand zur Stufenkante jedoch bei höchstens etwa 10 nm liegt. Dieser Wert wird bei $L = 0$ nm und $L = 60$ nm erreicht, wo die Spektren

jeweils einen Peak bei der Dirac-Energie zeigen. Da der Kantenzustand der Stufe in dieser Distanz bereits abgeklungen sein sollte, weist diese Intensität somit darauf hin, dass auch die Versetzungslinie einen Zustand besitzt. Die Spektren selbst zeigen außerdem bei $L = 15$ nm und $L = 45$ nm zwei Intensitätsmaxima, weisen also einen Doppelpeak auf. Um jedoch die Bereiche dazwischen zu analysieren, ist ein Blick auf das d^3I/d^3U -Signal nötig.

Dort sind über den gesamten Bereich $5 \text{ nm} \lesssim L \lesssim 55 \text{ nm}$ hinweg zwei Minima zu erkennen. Analog zu Abb. 5.9(b) sind die Positionen der Doppelpeaks durch eine rot gestrichelte Linie markiert. Zusätzlich sind noch die weiß gestrichelten Linien aus Abb. 5.9(b), unter Berücksichtigung der Skalierung, eingezeichnet. Jede Linie für sich alleine beschreibt die Variation der energetischen Lage eines Peaks entlang der Stufenkante (weiß) bzw. Versetzungslinie (rot). Der horizontale Abstand zwischen zwei gleichfarbigen Linien entspricht daher der energetischen Aufspaltung. Für beide Peaks stimmt der Verlauf der beiden jeweiligen Linien gut überein. Lediglich bei dem höherenergetischen Peak gibt es kleinere Abweichungen. Die Aufspaltung des Doppelpeaks verhält sich also entlang der Stufe und der Versetzungslinie gleich, jedoch nimmt sie bei letzterer signifikant schneller ab, da die Abbildung um den Faktor 2.5 angepasst wurde.

Der prinzipielle Verlauf der Spektren ist also für Stufenkante und Versetzungslinie nahezu identisch, auch wenn für letztere die Aufspaltung schneller mit zunehmendem Abstand vom Kreuzungspunkt abklingt. Dieser Befund unterstützt die zuvor aufgestellte These, dass auch die Versetzung einen Kantenzustand hat. Es gilt jedoch zu bedenken, dass Stufenkanten ab einem Abstand von etwa 25 nm zueinander bereits Einfluss auf die Ergebnisse der Tunnelspektroskopie nehmen. Daher wäre zur vollständigen Verifizierung dieser Hypothese Spektren mit Peak bei der Dirac-Energie in mindestens diesem Abstand zur nächsten ungeraden Stufenkante nötig. Im Rahmen dieser Arbeit liegt jedoch kein Spektrum an einer passenden Stelle vor.

Aber selbst aus dem Fehlen eines Peaks könnte nicht zwingend auf die Abwesenheit eines Kantenzustands geschlossen werden, da einerseits unbekannt ist, wie tief die Versetzungslinie unter der Oberfläche sitzt, und andererseits wie groß dessen Abklinglänge ist. Sessi *et al.* geben für die Ausdehnung des Kantenzustands senkrecht zur Oberfläche 4 Gitterlängen, also etwa 2.5 nm an [39]. Da dieser Wert eine Größenordnung kleiner ist als der in dieser Arbeit ermittelte kritische Stufe-Stufe-Abstand, ist es möglich, dass zwar auf den Terrassen über der Versetzung kein Zustand zu sehen ist, es jedoch sehr wohl zu einer Aufspaltung an der Stufenkante kommt. Dies setzt allerdings voraus, dass die kritischen Werte für lateralen und vertikalen Abstand vergleichbar sind, es also zu keiner Entkoppelung der Oberflächen- und Volumenzustände kommt.

Sofern die Wellenfunktion im Volumen nicht zusätzlich unterdrückt wird (z.B. aufgrund der offensichtlich vorhandenen Gitterverzerrung), können ab einem vertikalen Abstand von etwa 5 nm die Zustände von Stufe und Versetzung überlappen. Dieser Wert ist vergleichbar mit den 7 nm, welche bei dem Vergleich der Aufspaltung mit zusammenlaufenden Stufen ermittelt wurden.

Zusammengefasst kann daher festgestellt werden, dass die Daten die Existenz eines Kantenzustands entlang der Versetzungslinie nahelegen. Als Kopplungsmechanismus kommen, mangels Kenntnis über den vertikalen Abstand zwischen Stufenkante und Versetzung, sowohl die direkte Überlappung der Wellenfunktionen als auch ein indirekter Einfluss analog den zusammenlaufenden Stufenkanten in Frage.

Schraubversetzung

Bei der zweiten Art von eindimensionalen Defekten, die im Rahmen dieser Arbeit untersucht werden, handelt es sich um Schraubversetzungen. Im Gegensatz zu dem vorherigen Fall kann diese Defektart durch den kontinuierlichen Übergang zweier Terrassen ineinander eindeutig identifiziert werden.

Um deren Struktur zu erklären wird zunächst die Entstehung von Stufenkanten durch eine Versetzung betrachtet. Eine unendlich ausgedehnte Normalenebene zur Oberfläche durchstößt vier Seiten eines kubischen, endlichen Gitters. Wenn der Kristall entlang dieser Ebene gespalten wird und die Bruchstücke senkrecht zur Oberfläche um einen primitiven Gittervektor verschoben werden, dann entsteht auf der Ober- und Unterseite jeweils eine atomare Stufe. Bei Schraubversetzungen handelt es sich um eine halbunendliche Normalenebene, deren Ausdehnung in der Oberflächenebene beschränkt ist. Die Begrenzungslinie verläuft durch den Kristall und senkrecht zur Oberfläche. In Folge dessen schneidet die Ebene auch nur drei Seiten des Kristalls. Analog zur gerade zuvor beschriebenen Entstehung von Stufenkanten werden an der Schnittfläche die Atome senkrecht zur Oberfläche gegeneinander verschoben, sodass eine Stufe entsteht. In dem Bereich um die Begrenzungslinie herum führt der Übergang von „Stufe vorhanden“ hin zu „keine Stufe vorhanden“ dazu, dass die Atome nicht mehr auf den idealen Gitterplätzen sitzen können. Stattdessen winden sie sich schraubenartig entlang der Begrenzungslinie und verbinden so die gegeneinander versetzten Atomlagen auf beiden Seiten der Stufenkante miteinander. Dort wo die Ebene nicht existiert bildet die Oberfläche hingegen ein zusammenhängendes, translationsinvariantes Gitter. Anschaulich beschrieben entsteht so ein Defekt dann, wenn ein Kristall von der Seite her (auf seiner gesamten Höhe) eingeschnitten wird und die beiden Teile senkrecht zur Schnittlinie gegeneinander verschoben werden.

Die eindimensionale Defektlinie verläuft bei Schraubversetzungen also senkrecht zur Oberfläche. In Abb. 5.10 (a) ist die Topographie um den Schnittpunkt dieser Linie mit der Oberfläche gezeigt. Nach unten verläuft, entlang einer Hochsymmetrierichtung, die Schmittebene und dementsprechend die Stufenkante. Im oberen Bildteil ist anhand der Farbe ein kontinuierlicher Übergang zwischen der oben (gelb) und der unten (rot) gelegenen Terrasse zu sehen. Daneben wurde aus den STS-Messungen über diesen Bereich ein Schnittbild konstanter Energie rekonstruiert, wel-

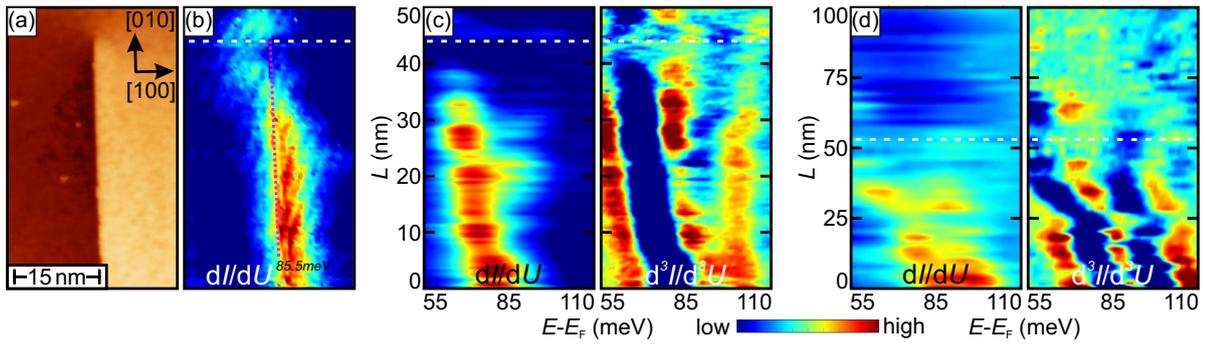


Abbildung 5.10 Aufspaltung des Kantenzustands an Schraubversetzungen: (a) Topographie von einer intrinsischen Schraubversetzung. Die Schnittfläche entlang der sich eine Stufenkante ausbildet verläuft atomar glatt und in Hochsymmetrierichtung. (b) Schnittbild konstanter Energie, extrahiert aus STS-Messungen bei $E_{D^*} = 85$ meV. Die lila Linie zeigt den Verlauf der Stufe, die weiße Linie markiert deren Ende. (c,d) Links dI/dU - und rechts d^3I/d^3U -Signal der in (a) gezeigten Schraubversetzung. Die weißen Linien markieren den Ursprung der Schraubversetzung. Messparameter: $U = 1$ V, $I = 10$ pA (a). Stabilisierungsparameter: $U = 125$ mV, $I = 200$ pA.

ches die Intensitätsverteilung bei der Dirac-Energie zeigt (b). Während die weiße, horizontale Linie hier und in allen weiteren Teilbildern den Ursprung der Schraubversetzung markiert, veranschaulicht die vertikale, lila Linie den Verlauf der Stufenkante. Bis zu einem Abstand von etwa 20 nm zeigt sich das gewohnte Bild, eine ausgeprägtes Intensitätsmaximum mit leichten Fluktuationen wenige Nanometer neben der Stufe auf der oberen Terrasse. Anschließend nimmt die Intensität über eine Länge von etwa 10 nm signifikant ab. Unterschreitet der Abstand einen kritischen Wert von $d_c \approx 10$ nm so ist dann ein Intensitätsmaximum auf der untenliegenden Terrasse zu sehen, das sichelförmig um den Ursprung der Schraubversetzung verläuft und sogar im oberen Teil des Bildes über die Stufe hinaus reicht.

Dieses Verhalten korreliert auch mit der Höhe der Stufenkante. Im unteren Teil des Bildes beträgt diese etwa 2.5 \AA und liegt damit etwas unterhalb der ungestörten Stufe (3 \AA). Auf einer Länge von 5 nm ändert sich die Höhe um etwa 0.1 \AA und nimmt ab d_c mit 1 \AA pro 5 nm ungefähr eine Größenordnung schneller ab. In diesem Bereich ist also die Deformation des Gitters durch die Defektlinie am ausgeprägtesten. Gleichzeitig wechselt in diesem Bereich der Kantenzustand von der oberen auf die untere Terrasse. Zum jetzigen Zeitpunkt sind Schraubversetzungen der einzige bekannte Fall in dem der Kantenzustand auf der unteren Terrasse bzw. in direkten Umgebung des Schnittpunktes zwischen Versetzungslinie und Oberfläche sogar gänzlich ohne Stufenkante beobachtet werden kann. Am oberen Rand des Bildes beträgt dann der Höhenunterschied zwischen ganz links und ganz rechts immer noch etwa 0.5 \AA , es liegt also noch kein ungestörtes Gitter vor.

Um dieses Phänomen genauer zu untersuchen wurden zwei STS-Messungen durchgeführt. In Abb. 5.10(c) wird über den Bildausschnitt von (a) gemessen, in (d) wurde ein etwas größerer Bereich gewählt, der auch das Gebiet oberhalb des Ursprungs abdeckt. Die einzelnen Daten-

punkte werden dabei in Reihen senkrecht zur Stufenkante aufgenommen. In der linken Spalte ist jeweils das dI/dV - und in der rechten Spalte das d^3I/d^3U -Signal gezeigt. Auf den ersten Blick erscheinen die Daten inkonsistent, da in (c) fast durchgehend nur ein einzelner Peak im dI/dV -Signal zu sehen ist, in (d) jedoch ein Doppelpeak. Tatsächlich lassen sich, ähnlich wie bei zusammenlaufenden Stufen, in einem gewissen Übergangsbereich die Spektren einer Reihe teilweise beiden Klassen zuordnen. Tendenziell ist dabei der Doppelpeak besser zu erkennen, wenn der Abstand zur Stufenkante kleiner ist. Diese Unterschiede können bereits bei einem Abstand von 1-2 nm zwischen zwei Datenpunkten auftreten. Da die Daten aus zwei unterschiedlichen Messungen stammen und die laterale Auflösung jeweils auf etwa 0.5 nm begrenzt ist, kommt es zu einer kleinen Verschiebung bezüglich des Abstands zur Stufenkante zwischen (c) und (d), die von vergleichbarer Größe ist. Außerdem sind die Grenzwerte des Farbcodes für beide Bilder nicht einheitlich. Bei gleicher Auftragung würde die Intensität der Doppelpeaks erheblich geringer ausfallen. Auch entspricht deren Abstand in etwa der Breite des Einzelpeaks in (c), es handelt sich also dort um einem bereits stark verbreiterten Peak. Die Teilbilder zeigen also Spektren in einem leicht unterschiedlichen Abstand senkrecht zur Stufe, was dazu führt, dass bei (c) der Doppelpeak gerade noch nicht und in (d) gerade so bereits erkennbar ist.

Die Hypothese eines gerade so als solchen noch nicht identifizierbaren Doppelpeaks in (b) bestätigt sich im d^3I/d^3U -Signal. Spätestens ab $L > 10$ nm ist eindeutig ein zweites Minimum zu erkennen, darunter wird dieses durch ein auffallend breites Maximum überlagert. In (d) ist, wie bereits im ursprünglichen Signal, über die gesamte Länge hinweg klar ein Doppelpeak bzw. Doppelminimum zu sehen. Die kleinste ermittelte Aufspaltung liegt in beiden Messungen bei etwa 18 meV und nimmt für $d < 25$ nm bis zu einem Wert von etwa 28 meV zu. In beiden Datensätzen kommt es in diesem Bereich außerdem zu einer stärkeren Blauverschiebung des niederenergetischen Peaks (im Vergleich zur Rotverschiebung des hochenergetischen Peaks), was zu einer leichten Asymmetrie der Aufspaltung führt. Mit dem Erreichen des Ursprungs der Stufe verschwindet der hochenergetische Peak, während der andere noch wenige Nanometer darüber hinaus existiert. Ab $L \gtrsim 75$ nm ist in (d) nur noch ein schwaches Minimum zwischen 100-110 meV zu sehen, welches zum Teil bereits vorher auch schon beobachtbar ist. Da dieses Minimum jedoch am Rand der Spektren sitzt, kann nicht abschließend festgestellt werden, ob es sich um einen Peak oder den ansteigenden Hintergrund z.B durch die vHS handelt. Aufgrund der im direkten Vergleich zum Doppelpeak relativ niedrigen Intensität und der relativ weiten Entfernung zur Dirac-Energie wird der mit diesem Minimum assoziierte Peak bzw. Anstieg hier dem Hintergrund zugeordnet. Die parallele Verschiebung des Doppelpeaks im unteren Bildteil für $L \lesssim 20$ nm ist wahrscheinlich auf die Schwankung des chemischen Potentials durch Fluktuationen der Sn-Konzentration zurückzuführen. Mit einer Verschiebung von $\Delta E \approx 5$ meV liegt sie somit sowohl bezüglich der Längen- als auch Energieskala in dem zu erwartenden Bereich.

Auch in diesem Fall stellt sich die Frage nach dem Mechanismus der zur Aufspaltung des Kantenzustands führt. Als Möglichkeit kommt auch hier, falls die Versetzungslinie im inneren einen Zustand besitzt, eine direkte Kopplung in Frage. Zu dem Zeitpunkt dieser Arbeit liegen zwar keine Kenntnisse über spezifische Berechnungen oder Vorhersagen für $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ bezüglich eines eindimensionalen, topologisch geschützten Zustands entlang der Ursprungslinie von Schraubversetzungen vor. Prinzipiell wird diese Möglichkeit auch nicht durch die in dieser Arbeit festgestellte Unterdrückung der Wechselwirkung 1D-Kantenzustände bei senkrechter Orientierung ausgeschlossen. Die STM-Daten zeigen deutlich, dass die Oberfläche um den Ursprung der Stufenkante herum in alle Raumrichtungen verkippt und die Gittersymmetrie entsprechend gestört ist. Dementsprechend kann aufgrund des Gitters nicht zwingend von einer vollständigen Orthogonalität bezüglich der beteiligten Wellenvektoren bzw. Orbitalen ausgegangen werden.

Alternativ kann es sich, analog zu den zusammenlaufenden Stufenkanten, auch um eine langreichweitige Wechselwirkung handeln, die durch die Änderung der kristallinen Struktur auf Längenskalen der Kohärenzlänge der Elektronen induziert wird. In diesem Fall sind die fundamentalen (Kristall-) Symmetrien und deren Brechung für die Aufspaltung von entscheidender Bedeutung, wie sie auch am Kreuzungspunkt von ungeraden Stufen auftreten. Die Gesamtheit der Daten aus früheren Experimenten hat dabei gezeigt, dass ein eindimensionaler, topologisch geschützter Zustand an Kanten entsteht, wenn diese die Translationssymmetrie des Gitters brechen. In diesem Kapitel wurde nun die Aufhebung dieses Symmetriebruchs untersucht. Zunächst wurde der einfachste Fall herangezogen, zwei Stufen die jeweils für sich genommen die Translationssymmetrie brechen und dann deren Abstand sukzessive verringert. Noch anschaulicher ist der Übergang von Symmetriebruch zu Nicht-Brech bei den zuletzt vorgestellten Schraubversetzungen. Sobald der Bruch der Translationssymmetrie aufgehoben wurde, konnte eine Aufspaltung des Kantenzustands beobachtet werden. Diese Aufspaltung deutet darauf hin, dass sich eine Bandlücke öffnet und der topologische Schutz nicht mehr oder nur noch partiell gegeben ist.

Nach meinem Kenntnisstand handelt es sich bei den hier vorliegenden Daten um den ersten und einzigen Fall, wo sich die Bandlücke nicht durch das Brechen einer Symmetrie öffnet, sondern durch die Aufhebung eines Symmetriebruchs. Bisher konnte der Zusammenhang zwischen Symmetrie, Oberflächenzustand und Bandlücke folgendermaßen beschrieben werden²³: Durch eine Symmetrie wird in dem topologisch nicht trivialem Material eine ungerade Anzahl an Bandinversionen an Hochsymmetriepunkten garantiert. Durch die Stetigkeitsbedingungen an den Grenzflächen zu nicht trivialen Materialien (Isolatoren) muss es zum Schließen der Bandlücke durch Oberflächenzustände kommen. Die gegebene Symmetrie garantiert somit Oberflächenzuständen. Sobald die zugrunde liegende Symmetrie gebrochen wird, kann die Stetigkeitsbedin-

²³Spezialfälle wie den den QHE und alle weiteren Fälle, wo der topologische Schutz auch ohne fundamentale Symmetrie (TRS, PHS, C) existiert ausgenommen.

gung auch im Volumen des „nicht trivialen“ Materials erfüllt werden, sodass die Garantie von Oberflächenzuständen entfällt. Der Symmetriebruch führt daher zu dem Verschwinden der Oberflächenzustände und dem Öffnen einer Bandlücke. Bei (Pb,Sn)Se existiert der Kantenzustand nun dort wo die (Translations-) Symmetrie gebrochen ist. Der Symmetriebruch garantiert nun die Existenz eines Oberflächenzustands. Gleichzeitig gibt es Situationen, in denen der Oberflächenzustand sich aufspaltet, d. h. es kommt zur Öffnung der Bandlücke. Diese Aufspaltungen korrelieren mit solchen Topographien, bei denen der Bruch der Translationssymmetrie rückgängig gemacht bzw. aufgehoben wird. Somit führt die Aufhebung des Symmetriebruchs nun zu dem Verschwinden der Oberflächenzustände und dem Öffnen der Bandlücke. Zusammengefasst ergibt sich also, dass die Rollen der Symmetrie bzw. der Symmetriebrechung bezüglich der Existenz von topologisch geschützten Zuständen vertauscht sind.

6 Dotierung von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ mittels Adatome

Nachdem im vorherigen Kapitel die intrinsischen Wechselwirkungen zwischen Kantenzuständen und der Einfluss von Defekten auf diese untersucht wurden, soll in dem nun folgendem Kapitel die Einkopplung extrinsischer Freiheitsgrade betrachtet werden. Eine naheliegende Wahl hierfür sind statische elektrische oder magnetische Felder, da diese relativ einfach experimentell realisiert werden können.

Bei elektrischen Felder steht oftmals in erster Linie der Ladungstransport im Mittelpunkt des Interesses, was jedoch eine kontrollierte Kontaktierung der Probe voraussetzt. Diese ist bei 1D-Kanälen, die nicht am Rand einer Probe sitzen, eine große Herausforderung, weil in diesem Fall nur eine Stufenkante zum Transport beitragen soll. Dementsprechend muss sichergestellt werden, dass zwei relativ weit auseinanderliegende Kontakte mit signifikanter lateraler Ausdehnung an die selbe Stufe angebracht werden und diese zwischendurch nicht unterbrochen wird. Eine Möglichkeit wäre die Nutzung eines Mehrfach-Spitzen-STM, welches jedoch nicht zur Verfügung steht. Selbst wenn das Problem der Kontaktierung gelöst ist, so muss anschließend noch sichergestellt werden, dass der Ladungstransport des 1D-Kanals dominant gegenüber dem unerwünschten Ladungstransport im Volumen des Materials ist. Diese Herausforderungen übersteigen den Rahmen dieser Arbeit deutlich, weswegen stattdessen der Einfluss von Magnetismus auf die Kantenzustände untersucht wird.

Externe Magnetfelder wurden bereits sehr früh an $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ angelegt und die Robustheit der Kantenzustände in diesen gezeigt [39, 51]. Warum soll dann überhaupt noch die Wechselwirkung mit magnetischen Adatomen untersucht werden? Ein Vergleich zwischen der mittleren Magnetfeldstärke von externen Felder und dem Molekularfeld des intrinsischen Magnetismus zeigt, dass die Feldstärken bei letzterem um Größenordnungen höher sein können. Ein direkter Vergleich der beiden Größen ist jedoch grenzwertig bis unangemessen, weil das Molekularfeld eine rein phänomenologische Hilfsgröße ist. Es liegen fundamental unterschiedliche Wechselwirkungsmechanismen zugrunde, was der eigentliche Grund ist, warum durch intrinsischen Magnetismus neue Phänomene auftreten können.

Tatsächlich ist auch die Aufgabe, einen intrinsisch magnetischen (topologischen) Isolator zu finden, aus Sicht der Kristallzucht, nicht trivial. Daher wurde hier ein einfacherer Ansatz gewählt, bei dem Materialien mit magnetischem Moment auf die Oberfläche aufgedampft werden. Durch vielfältige Wechselwirkungsmechanismen wie z.B Hybridisierung mit Oberflächen- und/oder Vo-

luminenzuständen oder durch Interaktion mit den Leitungselektronen (Ruderman Kittel Kasuya Yosida interaction, RKKY–Wechselwirkung) können so magnetische Freiheitsgrade induziert werden.

Dieser Ansatz bietet noch einen weiteren Vorteil, nämlich den, dass die Adatome in den meisten Fällen nicht isoelektronisch sind, also entweder Elektronen an den Kristall abgeben oder von diesem aufnehmen. Dadurch können sie benutzt werden, um das Fermi-niveau zu verschieben. Dabei kann die Lage des Fermi-niveaus durch die Bedeckung mit Adatomen viel präziser kontrolliert werden, als dies beispielsweise durch Variation der Prozessparameter (u. a. Temperatur, Druck, Stöchiometrie) während des Kristallwachstums möglich wäre. Intrinsische Defekte können ebenfalls zu Dotierung führen und durch Änderung der Prozessparameter kann deren Formationsenergie kontrolliert werden, jedoch ergibt sich oft entweder eine p- oder n-Dotierung des ganzen Kristalls oder die Kompensation findet nur in einem Teilbereich statt.

Der Dotierungseffekt der Adatome ist für sich genommen bereits interessant, weil wegen des Pauli-Prinzips nur Elektronen in unmittelbarer energetischen Nähe zum Fermi-niveau am Transport bzw. allgemein an Wechselwirkungen teilnehmen können. Durch die Dotierung ist es möglich, das Fermi-niveau in die Nähe des Dirac-Punktes zu verschieben, sodass auch die Dirac-Elektronen zu den Wechselwirkungen beitragen können. Durch deren Spin-Impuls-Kopplung und der daraus folgenden Robustheit gegen elastische Rückstreuung ergibt sich die Möglichkeit für einen langreichweitigen, spinpolarisierten Ladungstransport. Solche Systeme aus korrelierten Dirac-Elektronen könnten zu neuartigen Wechselwirkungen und Phasen führen.

Diese Überlegungen beziehen sich zunächst auf die 2D-Oberflächenzustände, da deren SML theoretisch gut verstanden, beschrieben und experimentell nachgewiesen wurde (vgl. Kap. 4). Darüber hinaus bieten Experimente mit magnetischen Adatome jedoch auch die Möglichkeit, einen Beitrag zur kontrovers diskutierten Frage, ob bei den 1D-Kantenzustände eine Spinpolarisation vorliegt oder nicht, zu liefern. Deshalb beschäftigt sich diese Arbeit auch mit dem Aufdampfen von Adatomen. Da diese Art von Experimenten jedoch Einmalversuche sind, bei denen jede neue Probenpräparation eine neue Spaltprobe benötigt, geht es zunächst schwerpunktmäßig um die Möglichkeit, $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ durch Standard-Ultrahochvakuum-Methoden (Ultra-High Vacuum, UHV) *in situ* zu reinigen.

6.1 *In situ* Reinigung von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$

Typische Ansätze, um kristalline Proben im UHV zu reinigen, sind entweder kombinierte Zyklen aus Ionenzerstäubung und Tempern, Tempern unter einer Reaktivgasatmosphäre oder sogenannte Hoch-Temperatur-Pulse (high temperature flash, HTF). Letztere Methode wird vor allem bei Metallen mit sehr hohem Schmelzpunkt angewandt, weil durch Elektronenbeschuss in einem

Hochspannungsfeld bereits bei moderaten Leistungen von einigen 100 Watt Temperaturen von über 2000°C erreicht werden können. Dadurch können selbst Oxidschichten entfernt werden, deren Bindung zum Substrat zu stark für die Ionenzerstäubung ist. Bei Halbleitern kann wegen des höheren Widerstands und der in der Regel deutlich niedrigeren Schmelzpunkte auch direkt ein Strom durch die Probe geschickt werden. Auf HTF wurde im Rahmen dieser Arbeit verzichtet, da $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ keine stark gebundene Oxidschicht besitzt, sodass diese durch andere Prozesse erfolgreich entfernt werden konnte.

Eine Möglichkeit hierfür stellt die Exposition in einer Reaktivgasatmosphäre dar, wobei es oft zur Überwindung der Aktivierungsenergie nötig ist, dass das Substrat trotzdem noch geheizt werden muss. Wenn das Substrat nicht geheizt werden soll oder die benötigten Temperaturen nicht aushält, so kann alternativ das Reaktionsgas auch beim Einlassen in die UHV-Kammer gecrackt und in Radikalform direkt der Oberfläche zugeführt werden. Oft wird die Probe nur noch moderat getempert, um die ungleichmäßige Materialabtragung zu kompensieren und eine glatte Oberfläche zu erhalten. Eine Studie aus Röntgenphotoelektronenspektroskopie (X-ray Photoelectron Spectroscopy, XPS) und Rasterelektronenmikroskopie (scanning electron microscopy, SEM) hat festgestellt, dass dieser Ansatz bei SnTe-Nanokristallen die besten Resultate hervorbringt [57]. Allerdings ist das Auflösungsvermögen der SEM nicht hoch genug um zu bestimmen, ob die Oberflächenqualität auch den hohen Anforderungen eines STM genügt. Auch stand während dieser Arbeit kein Wasserstoffcracker zur Verfügung, sodass sich im Folgenden auf die Kombination aus Tempern und Ionenzerstäubung konzentriert wird.

Bei der Ionenzerstäubung (englisch sputtering) wird ein Inertgas in die UHV-Kammer eingelassen, ionisiert und eine Potenzialdifferenz zwischen Ionenquelle und Probe erzeugt. Die Ionen werden so in Richtung der Probe beschleunigt und übertragen beim Auftreffen ihren Impuls an die Oberflächenatome. Diese werden dadurch aus dem Kristallgitter gelöst, sodass die Oberfläche abgetragen wird [58]. Um den größtmöglichen Impulsübertrag zu generieren wurde in dieser Arbeit daher Argon als Inertgas verwendet. Edelgase haben den Vorteil, dass sie nur in sehr seltenen Fällen chemische Verbindungen eingehen und daher nicht in die Bindungen der Probe aufgenommen werden. Allerdings besteht die Gefahr der Ionenimplantierung, wenn die kinetische Energie der Ar-Ionen zu groß ist. In diesem Fall dringen die Ionen tiefer in den Kristall ein, sodass sie diesen nicht mehr verlassen können. Daher ist die Beschleunigungsspannung ein wichtiger Prozessparameter. Darüber hinaus ist der Ionenfluss, also die Anzahl der pro Zeit und Flächenelement ankommenden Ar-Ionen, relevant. Er wird im wesentlichen durch den Ar-Partialdruck bestimmt. Aus diesen beiden Parametern ergibt sich eine Abtragsrate, welche typischerweise in Monolagen pro Zeiteinheit angegeben wird. Da der zugrundeliegende Mechanismus auf Impulsübertragung basiert, lässt sich eine Abtragsrate streng genommen nur für einzelne chemische Elemente (unter Vernachlässigung der Isotopenverteilung) definieren.

Aufgrund dieser elementspezifischen Sputterausbeute kann sich durch das Ionenzerstäuben von Mischkristallen wie $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ die Stöchiometrie verändern [58].

Da die Ionenzerstäubung nur dem Abtragen einzelner Atomlagen (MonoLayer, ML) bzw. einzelner Atome dient, wurden die Prozessparameter für diese Arbeit bei $E_{\text{kin}}^{\text{Ion}} = 350 \text{ eV}$ und $p_{\text{ges}} = 2 \cdot 10^{-6} \text{ mbar}$ fixiert, wobei anstelle des Partialdrucks aus praktischen Gründen der Gesamtdruck in der Kammer verwendet wurde. Die Abtragungsrate θ wurde für diese Parameterwerte auf Ag(111) kalibriert und auf $\theta \approx 0.2 \text{ ML} \cdot \text{min}^{-1}$ bestimmt. Auch die Dauer des Ionenzerstäubens wurde mit $t = 15 \text{ min}$ bei allen Versuchen konstant gehalten, sodass jedes mal etwa 2-3 Monolagen durch den Ionenbeschuss abgetragen wurden.

Da die Ionen zufällig auf die Probe treffen, wird die Oberfläche durch diesen Prozess aufgeraut. Ungetemperte Proben weisen größtenteils eine Korrugation von etwa 600 pm auf, was von der Höhe her zwei Atomlagen entspricht. Daraus kann geschlossen werden, dass die Abtragungsrate nicht signifikant kleiner als auf Ag(111) ist. Die exakte Rate ist jedoch unbekannt, weil die Anzahl der vollständig abgetragenen Lagen nicht bestimmbar ist. Um auf der Oberfläche wieder atomar glatte Terrassen zu bekommen, wird nach jeder Ionenzerstäubung die Probe anschließend geheizt. In dem bisher verwendeten Versuchsaufbau wurden die (Pb,Sn)Se-Einkristalle mit einem leitfähigen Silber-Epoxid-Kleber auf einem Probeträger befestigt und eine kurze Metallstange zum Spalten darauf geklebt. Da der Silber-Epoxid-Kleber jedoch nicht temperaturbeständig ist und beim Erhitzen der Druck in der UHV-Kammer um mehrere Größenordnungen ansteigt, wurde eine neue Probenaufnahme konstruiert. Bei dieser werden die (Pb,Sn)Se-Einkristalle seitlich zwischen zwei Metallblöcke eingeklemmt und die Blöcke durch zwei Schrauben gegeneinander gepresst. Zur Herstellung des thermischen Kontakts ist dann einer der Blöcke noch zusätzlich mit dem Probenträgerplättchen verschraubt.

Durch die erhöhte Temperatur während des Temperns wird ein signifikanter Anteil der Oberflächenatome thermisch aktiviert und überschreitet die Diffusionsbarriere, sodass die Gesamtenergie durch die Wiederherstellung eines geordneten Kristallgitters minimiert werden kann. Für die Endtemperatur spielt auch die Dauer des Ausheizens eine Rolle, weil die Thermalisierung von Probe, Probenträger und Manipulator relativ lange benötigt. Testmessungen mit einem Thermoelement haben gezeigt, dass das thermodynamische Gleichgewicht erst nach 20-25 Minuten erreicht wird. Konsistent damit lieferte eine Versuchsreihe mit 5, 10, 15, 20 und 30 minütigem Tempern bei gleicher Heizleistung sukzessive immer besser ausgeheilte Oberflächen, während noch längeres Tempern zu keiner weiteren signifikanten Verbesserung der Oberflächenqualität geführt hat. Daher wurden alle folgenden Versuche mit einer Ausheizdauer von $t = 30 \text{ min}$ durchgeführt.

Der Parameter mit dem größten Einfluss auf die Oberflächenqualität ist die Heizleistung bzw. die damit verbundene Endtemperatur. In Abb. 6.1(a) sind STM-Übersichtsaufnahmen mit ei-

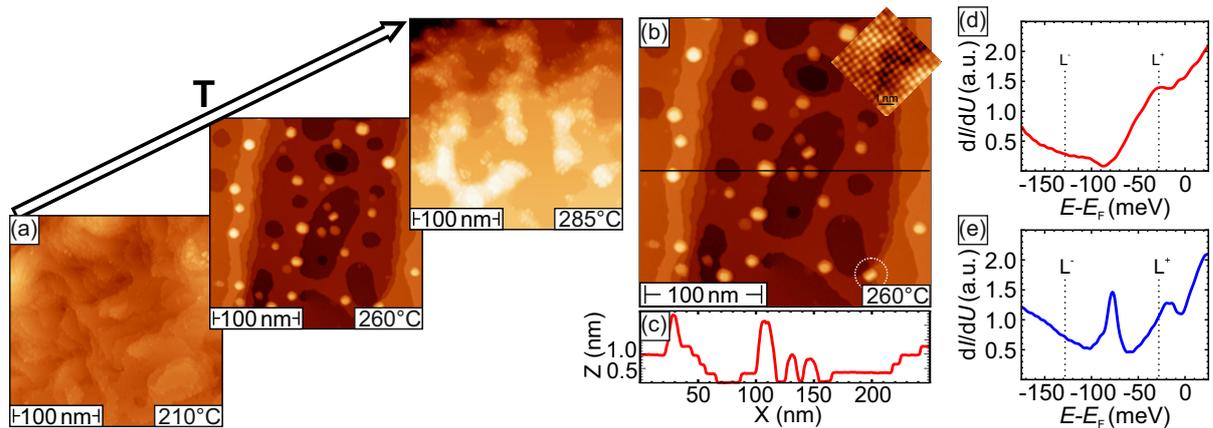


Abbildung 6.1 Temperaturabhängigkeit der Oberfläche: (a) STM-Topographien für verschiedene Ausheiztemperaturen nach vorheriger Ionenzerstäubung. Die Temperatur steigt von links nach rechts und ist zunächst zu niedrig und am Ende hoch. (b) Großaufnahme der optimalen Ausheiztemperatur. Neben atomar glatten Terrassen, siehe Teilbild oben rechts, sind Cluster und Löcher zu erkennen. Die Stufenkanten sind rau, besitzen keine Vorzugsrichtung und entsprechen alle einer Atomlage. Dies kann aus dem Linienprofil in (c) entnommen werden, das entlang der schwarzen Linie verläuft. (c,d) Spektren über einer Terrasse (rot) und einer Stufenkante (blau). Qualitativ ist kein Unterschied zu dem Zustand vor der Probenpräparation zu erkennen, quantitativ zeigt die Lage des Dirac-Punkts jedoch das die ursprünglich p-dotierte Probe jetzt den Dirac-Punkt unterhalb des Fermi-niveaus hat, also n-dotiert ist. Messparameter: $U = 1 \text{ V}$, $I = 10 \text{ pA}$ (a); $U = 500 \text{ mV}$, $I = 250 \text{ pA}$ (b). Stabilisierungsparameter: $U = 50 \text{ mV}$, $I = 200 \text{ pA}$ (c,d).

ner Größe von $250 \times 250 \text{ nm}^2$ für eine leicht zu niedrige (links), die optimierte (Mitte) und eine etwas zu hohe Temperatur (rechts) gezeigt. Bei einer zu niedrigen Temperatur kommt es nur teilweise zur Ausbildung glatter Terrassen, das Gesamtbild wird jedoch durch viele, unregelmäßig verlaufende Stufenkanten geprägt. Es ist anzumerken, dass die Rauigkeit in diesem Bild, genauso wie in dem rechts, definitiv erheblich größer als nur 2-3 Lagen ist. Die Ursache hierfür liegt wahrscheinlich in der Tatsache, dass diese Proben mehrere Zyklen aus Ionenzerstäubung und Tempern durchlaufen haben. Da die Temperatur dabei zunächst noch zu niedrig für eine vollständige Wiederherstellung der Oberfläche war, hat sich die Rauigkeit bis zu einem gewissen Grad von Zyklus zu Zyklus immer weiter erhöht. Die Topographie einer überhitzten Probe zeigt im Vergleich zur Topographie einer zu kalten getemperten Probe erheblich größere Terrassen, allerdings fehlen weiterhin geordnete, atomar glatte Stufenkanten. Eine zu hohe Probentemperatur führt außerdem zur Bildung von unregelmäßig geformten, wolkenartigen Clustern, die sich bevorzugt an Stufenkanten anlagern. Bemerkenswerterweise beträgt der Temperaturunterschied zwischen zu kalt und zu warm lediglich 65°C , sodass das Temperaturfenster für eine optimal ausgeheilte Oberfläche mit $250^\circ\text{C} \lesssim T \lesssim 280^\circ\text{C}$ ausgesprochen eng ist. Dementsprechend wichtig ist eine hinreichend lange Dauer des Temperns um sicherzustellen, dass die Temperatur in diesen Bereich liegt.

Die Topographie der optimierten Probentemperatur ist in Abb. 6.1(b) nochmals vergrößert dargestellt. Drei Besonderheiten fallen auf, die die Morphologie der ausgeheilten Oberfläche von der

von gespaltenen Einkristalle unterscheidet: Form und Orientierung der Stufenkanten, die Löcher und die Cluster. In dieser Arbeit konnte nach dem Tempern auf keiner Probe eine freistehende Mehrfachstufe gefunden werden. Wie dem Linienprofil in (c) zu entnehmen ist, wurden lediglich Einfachstufen beobachtet, die keine langreichweitige Präferenz bezüglich der Orientierung aufweisen. Lediglich kurze Teilstücke mit Längen $L \lesssim 25$ nm zeigen eine gewisse geradlinige Orientierung, wie z.B. bei dem großen Loch in der Bildmitte. Das Fehlen einer präferierten Stufenkantenorientierung wird auch durch die Abwesenheit atomar glatter Stufenkanten bestätigt. Stattdessen weisen alle Stufen eine selbst bei dieser Übersichtsaufnahme sichtbare Zerklüftung auf. Der überwiegend gekrümmte Verlauf der Stufen führt zu einer erheblichen Anzahl an Löchern, deren Häufigkeit und Größe mit der Gesamtzahl an Zyklen aus Ionenzerstäubung und Tempern variiert. In der Regel sind die meisten Fehlstelleninseln nach zwei Zyklen zu beobachten. Ebenfalls charakteristisch für die Topographie sind die vielen kleinen Cluster mit einer Höhe von wenigen Nanometern (siehe Linienprofil), von denen sich die meisten auf den Terrassen und einige an den Stufenkanten befinden. Letztere fungieren dabei oft als Defekte, an die die Stufen gepinnt sind. Ein solcher Cluster ist im unteren Bildteil mit einem weißen Kreis hervorgehoben, ein weiterer sitzt unter dem Teilbild oben rechts [siehe (a)].

Trotz dieser Materialansammlungen weist die Oberfläche eine hohe Ordnung auf, wie anhand des atomar aufgelösten Teilbilds in der oberen rechten Ecke zu sehen ist. Auch beeinflusst die Anwesenheit der Cluster die Spektren in der Umgebung nicht negativ, denn sowohl für die Terrasse als auch für die Stufenkante stimmen die Verläufe der beobachteten Spektren in (d) und (e) qualitativ sehr gut mit den bisherigen Messungen überein. Auf der Terrasse (rot) ist sehr gut die obere vHS bei L^+ und das V-förmige, globale Minimum beim Dirac-Punkt zu erkennen. Bei der gleichen Energie ist an der Stufe (blau) auch ein intensiver Peak in der LDOS zu sehen, der dem Kantenzustand zuzuordnen ist. Qualitativ sehen die Spektren also aus, wie es für eine topologisch nicht triviale Probe zu erwarten ist. Quantitativ deutet jedoch die Lage des Dirac-Punkts bei $E_{\text{DP}^*} = -88$ meV auf eine Änderung der Stöchiometrie hin, weil alle in dieser Arbeit verwendeten (Pb,Sn)Se-Einkristalle nach dem Spalten ihren Dirac-Punkt bei positiven Energien haben.

Ursache für diese Veränderung könnten unterschiedliche Aktivierungsenergien für die einzelnen Elemente Pb, Sn und Se sein, was zu verschiedenen Mobilitäten führt. Diese These wird auch von der Topographie gestützt, weil Löcher in der Regel auf eine etwas zu niedrige und Cluster auf eine bereits leicht zu hohe Proben temperatur während des Temperns hindeuten. Wenn es jedoch keine einheitliche Diffusionsbarriere gibt, so kann ein Szenario eintreten, bei dem Element A erst bei Temperaturen signifikant mobil wird, bei denen Element B bereits teilweise wieder aus dem Gitter gelöst wird und Materialansammlungen auf der Oberfläche bildet. Da diese Prozesse thermisch aktiviert werden, genügen sie einem Exponentialfaktor, sodass bereits

eine leichte Erhöhung der Temperatur zu deutlich mehr Clustern durch Element A führt. Andererseits kann die Temperatur mit dem selben Argument nicht nennenswert vermindert werden, da ansonsten Element B vollständig immobil und die Bildung von großen Terrassen unterdrückt wird. Als Konsequenz ergibt sich, dass ein gegebenenfalls existierendes Temperaturfenster nur sehr schmal sein kann, so wie es in den hier vorliegenden Experimenten der Fall ist.

Auf die naheliegende Frage, welches der drei Elemente für die Materialansammlungen verantwortlich sein könnte, wird später in diesem Kapitel eingegangen, sobald alle Ergebnisse vorgestellt wurden. Zunächst soll der Fokus auf der physikalisch interessanteren Frage liegen, ob dieser Effekt reproduzierbar und vor allem kontrollierbar ist. Falls dem so ist, ergibt sich die Möglichkeit das Fermi-niveau des Kristalls anzupassen ohne dabei auf Fremdatome zur Dotierung zurückgreifen zu müssen. Dadurch können ungewollte Wechselwirkungen zwischen $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ und den Dotieratomen, wie beispielsweise mit einem möglicherweise verbleibendem magnetischen Moment, von vornherein ausgeschlossen werden.

Die Resultate dreier ausgewählter Proben für verschiedene Probenpräparationen sind in Abb. 6.2 zusammengefasst. Um zu zeigen, dass die Verschiebung des Fermi-niveaus durch die Probenpräparation verursacht wird, wurden auf die eingeklemmten Proben zusätzlich eine Metallstange geklebt, die Probe zuerst gespalten und dann charakterisiert. Diese Ergebnisse sind in der Abbildung unter der Zykluszahl Null und in schwarz dargestellt. Als Zyklus wird von nun an die Kombination aus einmal Ionenzerstäubung und anschließendem Tempern bezeichnet. Wie zuvor bereits erläutert, sind die Prozessparameter für das Ionenzerstäuben dabei immer gleich, genauso wie die Dauer des Temperns. Als einziger freier Parameter verbleibt dann die Proben-temperatur, welche in Abb. 6.2 farbcodiert wurde. Zur Bestimmung der Energie des Dirac-Punkts wurde nach jedem Zyklus die Position des Kantenzustands bzw. seines Peaks in den jeweiligen Spektren herangezogen.

Für eine der drei Proben sind die dazugehörigen Spektren in Abb. 6.2 (a) gezeigt. Vor dem Beginn der Präparationsreihe liegt der Dirac-Punkt bei $E_{\text{DP}^*} = 80 \text{ meV}$. Der erste Zyklus wurde mit einer Ausheiztemperatur von 255°C durchgeführt und führt zu einer signifikanten n-Dotierung der Probe, was anhand der Lage des Dirac-Punkts von $E_{\text{DP}^*} = 32 \text{ meV}$ deutlich wird. Die Energieverschiebung des Kantenzustands ist etwa dreimal größer als die intrinsisch vorkommende Fluktuation und kann deshalb nicht auf zufällige Sn-Konzentrationschwankungen zurückgeführt werden.

Die Abhängigkeit der Dirac-Energie E_{DP^*} von der Zykluszahl ist nochmals unterhalb von (a) in Abb. 6.2 (b) graphisch aufgetragen. Für die Zyklen zwei und drei, ebenfalls bei einer Proben-temperatur von 255°C durchgeführt, setzt sich der Trend der n-Dotierung fort, weswegen auch der Energiebereich für die Spektren angepasst wurde. Ab dem vierten Zyklus wurde die Temperatur stufenweise erhöht, zunächst nur moderat um 5°C . Während diese Anpassung noch zu keiner

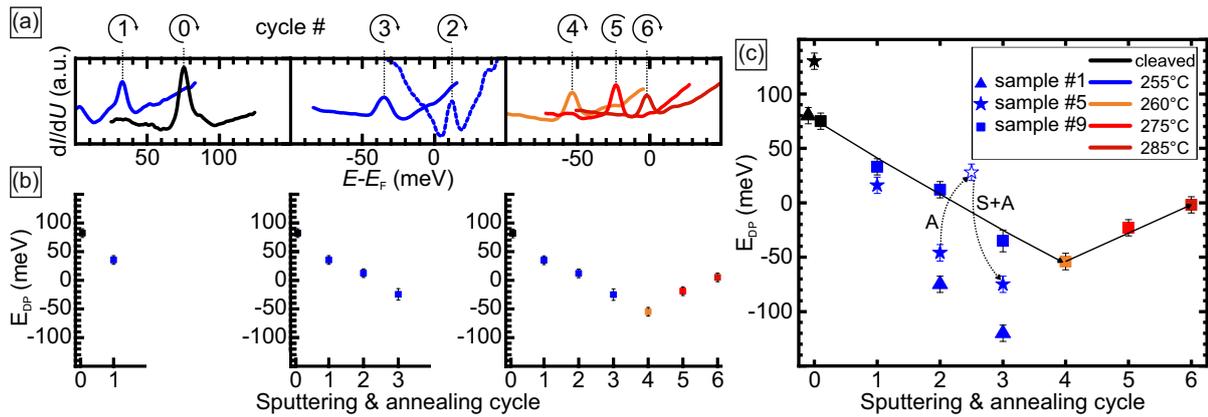


Abbildung 6.2 Ferminiveauverschiebung durch Ionenzerstäubung und Tempern: (a) Spektren des Kantenzustands nach dem Spalten (schwarz) und 6 Zyklen aus Ionenzerstäubung und Tempern. Alle Prozessparameter sind gleich, bis auf die farbkodierte Proben­temperatur während des Ausheizens. Die ersten vier Zyklen führen zu einer n-Dotierung, die letzten beiden aufgrund einer erhöhten Temperatur zu einer p-Dotierung. (b) Graphische Auftragung der Dirac-Energie gegen die Zyklenanzahl, basierend auf den Spektren aus (a). (c) Analoge Auftragung zu (b) für zwei weitere Proben. Bei Probe # 5 wurde nach zwei Zyklen einmal die Probe ohne vorherige Ionenzerstäubung für die doppelte Dauer getempert.

anderen Beobachtung geführt hat als bei den ersten drei Zyklen, so konnte bei einer Proben­temperatur von 275°C (Zyklus# 5) eine Umkehrung der Dotierung beobachtet werden. Lag der Dirac-Punkt nach vier Zyklen bereits etwa 55 meV unterhalb des Ferminiveaus, so hat sich seine Position um ungefähr 33 meV auf $E_{DP^*} = -22\text{ meV}$ erhöht. Auch bei noch höheren Proben­temperaturen setzt sich dieser Trend fort, wie anhand des Zyklus # 6 zu sehen ist. Dieser wurde bei $T = 285^\circ\text{C}$ durchgeführt und resultiert ebenfalls in einer positiven Energieverschiebung von $\delta E = +21\text{ meV}$. Auch hier sind die Verschiebungen zu groß, um sie auf wachstumsbedingte, zufällige Schwankungen der Stöchiometrie zurückzuführen. Die Versuchsreihe mit dieser Probe zeigt nicht nur, dass der Effekt einer Dotierung durch Zyklen aus Ionenzerstäubung und Tempern reproduzierbar ist, sondern dass durch die Wahl der Temperatur sogar die Art der Dotierung kontrollierbar ist.

Analog zu dieser Probe sind in Abb. 6.2 (c) die Dirac-Energien in Abhängigkeit der Zyklenanzahl auch für die anderen beiden Proben aufgetragen. Die Präparation von Proben #1 unterscheidet sich nur unwesentlich von der zuvor vorgestellten Probe # 9. Beide Proben stammen von dem selben Wirtskristall, haben daher zu Beginn die gleiche Dirac-Energie und haben anschließend drei Zyklen unter äquivalenten Bedingungen bei $T = 255^\circ\text{C}$ durchlaufen. Lediglich die Messung nach dem ersten Zyklus wurde bei Probe # 1 nicht durchgeführt, da es sich um den ersten Versuch mit einer neu angefertigten Probenhalterung gehandelt hat. Dementsprechend ist während des Temperns der Druck in der UHV-Kammer zu weit angestiegen. Davon abgesehen weist die Probe nach zwei bzw. drei Zyklen jeweils eine klare n-Dotierung auf, die allerdings jeweils signifikant stärker als bei Probe # 9 ist. Die Abweichung erreicht Werte von annähernd 100 meV

und zeigt damit deutlich, dass eine exakte Kontrolle durch stets gleiche Prozessparameter kaum zu realisieren ist. Dieser Umstand wird nochmals durch die dritte ausgewählte Probe (# 5) verdeutlicht. Diese stammt von einem anderen Wirtskristall und hat zu Anfang ihren Dirac-Punkt bei $E_{\text{DP}^*} = 130 \text{ meV}$. Obwohl der Dirac-Punkt also ursprünglich 50 meV weiter oberhalb sitzt, liegt ihre Dirac-Energie bereits nach dem ersten Zyklus unterhalb der von Probe # 9. Eine mögliche Ursache für diese Diskrepanz kann in der hohen Temperaturempfindlichkeit der Proben, in Kombination mit leichten Schwankungen des thermischen Kontakts liegen. Da jede Probe in Form und Größe ein Unikat ist, können zum einen die Kontaktflächen zwischen Probe und Metallblöcken variieren und zum anderen kann der Anpressdruck schwanken. Selbst der Drehmomentschlüssel mit der schwächsten Feder hat noch so viel Kraft auf die Proben ausgeübt, dass diese bereits *ex situ* gespalten wurden. Bei dem Anziehen per Hand ist aber davon auszugehen, dass das Drehmoment zwischen den einzelnen Proben unterschiedlich war.

Von besonderem Interesse an Probe # 5 ist jedoch der Zwischenschritt nach zwei Zyklen, bei dem die Probe ohne vorheriges Ionenzerstäuben direkt für die doppelte Dauer ($t = 60 \text{ min}$) bei $T = 255^\circ\text{C}$ getempert wurde. Auch in diesem Fall ist eine signifikante p-Dotierung beobachtbar. Der anschließende vollständige Zyklus unter den selben Bedingungen wie die ersten beiden führt dann wieder zu einer n-Dotierung. Damit ist eine reversible Manipulation des Fermi-niveaus möglich, ohne dabei die Proben höheren Temperaturen auszusetzen. Dadurch kann die Bildung von Clustern gehemmt und somit die Lebensdauer bzw. die Anzahl der durchführbaren Zyklen erhöht werden. Die Lebensdauer ist begrenzt, weil keine Oberflächen mit glatten Terrassen ohne Cluster präpariert werden konnten und die Materialablagerungen nicht vollständig beim Ionenzerstäuben entfernt wurden. Je größer der Anteil der Oberfläche ist, der von Clustern bedeckt wird, desto wahrscheinlicher ist es während des Temperns, dass sich weiteres Material anlagert. Deshalb verschlechtert sich nicht nur mit jedem Zyklus die Oberflächenqualität, sondern diese positive Rückkopplung führt dazu, dass das System irgendwann innerhalb von ein bis zwei weiteren Zyklen kippt und die Oberfläche vollständig zerstört wird.

In diesem Zusammenhang ist es nun naheliegend die Fragen zu diskutieren, aus welchen Elementen die Cluster bestehen könnten, wie das Ionenzerstäuben bzw. Tempern die Stöchiometrie beeinflussen und warum es bei einem reinen Ausheizen zu einer entgegengesetzten Verschiebung des Fermi-niveaus im Vergleich zu vollständigen Zyklen kommt. Eine definitive Aussage über die chemische Zusammensetzung einer Probe erlauben verschiedene Methoden wie XPS oder die Augerelektronenspektroskopie (Auger Electron Spectroscopy, AES), nicht jedoch die Tunnelmikroskopie. Die AES stand für diese Arbeit zur Verfügung und bietet im Vergleich zur XPS auch noch den Vorteil der Oberflächensensitivität, da die mittlere freie Weglänge für Elektronen in einem Festkörper wesentlich geringer ist als die von Photonen. Bedauerlicherweise wurde festgestellt, dass die in dieser Arbeit zur Verfügung stehenden Proben kleiner als der fokus-

sierte Elektronenstrahl sind, welcher die Augerelektronen aus der Probe löst. Das Augersignals der Probe wurde von dem der Probenhalterung überdeckt, sodass nur deren charakteristische Ta-Peaks im Augerspektrum zu sehen sind. Dementsprechend ist die Analyse der chemischen Zusammensetzung der Materialansammlungen auf die Kombination der vorhandenen STM-Daten und Literatur beschränkt.

Bei ARPES-Messungen von $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ -Filmen auf SrTiO_3 kann beobachtet werden, dass das Valenzband mit zunehmender Sn-Konzentration immer näher an die Bindungsenergie $E_{\text{Bin}} = 0\text{eV}$ heranrückt. Es stehen also weniger Elektronen zur Verfügung und das Fermi-niveau sinkt dementsprechend mit steigendem Sn-Anteil [59]. Dieses Verhalten ist auch aufgrund der höheren Elektronenaffinität von Sn gegenüber Pb zu erwarten. $(\text{Pb},\text{Sn})(\text{Se},\text{Te})$ -Molekularstrahlepitaxie-Proben (molecular beam epitaxy, MBE) werden üblicherweise aus den Vorprodukten SnSe, PbSe, SnTe und PbTe gewachsen und die Stöchiometrie durch deren Verhältnis zueinander kontrolliert. Wenn die sonstigen Wachstumsbedingungen dann gleich bleiben, so besitzen verschiedene Proben die gleiche Art und die gleiche Dichte von Defekten. Änderungen des Fermi-niveaus bzw. der Anzahl freier Ladungsträger sind dann auf die unterschiedlichen Zusammensetzungen zurückzuführen. Wie aus den Methoden von Ref. [60] hervorgeht, führt daher sowohl bei $(\text{Pb},\text{Sn})\text{Te}$ als auch bei $(\text{Pb},\text{Sn})\text{Se}$ eine Erhöhung der Sn-Konzentration zu mehr Löchern und dementsprechend zu einem absinkenden Fermi-niveau.

Ausgehend von diesen Ergebnissen werden die hier beobachteten Verschiebungen des Fermi-niveaus auf eine Änderung der Sn-Konzentration zurückgeführt. Selbst wenn durch das Tempern Kristalldefekte entstehen oder ausheilen, so dürften diese Defekte nicht isoelektronisch sein und müssten in einer erheblichen Anzahl auftreten. Die STM-Daten zeigen jedoch keine Anzeichen einer erhöhten Defektdichte z.B. in Form von Vakanzen. Lediglich die großen Löcher in Abb. 6.1 (b) deuten auf fehlendes Material hin. Da die Löcher jedoch unabhängig von der Tunnelspannung beobachtet werden können, also sowohl bei hohen Spannungen die das Se-Gitter zeigen, als auch bei niedrigen Spannungen, bei denen zusätzlich noch das (Pb,Sn) -Gitter sichtbar wird, ist davon auszugehen, dass sowohl Pb bzw. Sn als auch Se an diesen Stellen fehlt. Da nicht davon ausgegangen wird, dass während des Temperns eine signifikante Menge an Material ins Vakuum übertritt, muss geschlussfolgert werden, dass die Cluster zumindest teilweise auch aus Se bestehen.

Die in Abb. 6.2 (a) gezeigte Versuchsreihe zeigt zunächst ein Absinken der Dirac-Energie, hier als n-Dotierung bezeichnet ²⁴, was einem Ansteigen des Fermi-niveaus und dementsprechend einer verringerten Sn-Konzentration entspricht. Wie bereits diskutiert kann sowohl die Ionenzerstäu-

²⁴In dieser Arbeit bezieht sich der Begriff der Dotierung darauf, ob Dirac-Punkt ober- oder unterhalb des Fermi-niveaus sitzt bzw. sich von diesem weg (p) oder zu diesem hin (n) bewegt. In der konventionellen Sprache der Festkörperphysik wird jedoch geschaut, ob das System Elektronen abgibt oder aufnimmt. Da die Bandstruktur gleich bleibt und sich nur das Fermi-niveau verschiebt, ergibt sich daraus eine entgegengesetzte Nomenklatur.

bung als auch das Tempern zu einer Änderung der Zusammensetzung von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ führen. Allerdings konnte im Rahmen der Messgenauigkeit keine Änderung der Abtragsrate bei Proben mit einer unterschiedlichen Sn-Konzentrationen ($x_{\text{max}} = 0.07$), die in einem Argonplasma behandelt wurden, festgestellt werden [61]. Unter der Annahme, dass das Ionenzerstäuben keinen signifikanten Einfluss auf die Stöchiometrie nimmt, muss sich bei niedrigen Probertemperaturen zunächst ausschließlich oder mindestens überwiegend Sn in Form von Clustern an der Oberfläche ablagern. Es entsteht an der Oberfläche eine Sn-arme Zone, die bei erhöhten Temperaturen durch vermehrte Sn-Diffusion aus dem Volumen wieder abgebaut werden kann. Die bei Probe # 5 beobachtete p-Dotierung des Dirac-Punkts durch alleiniges Ausheizen könnte mit der verlängerten Dauer des Tempervorgangs zusammen hängen. Bei den anderen Zyklen erreichte die Probe lediglich für etwa 5 Minuten die Endtemperatur, was hinreichend lange zur Rekonstruktion der Oberfläche ist, aber möglicherweise nicht lange genug, um für einen nennenswerten Materialaustausch zwischen Oberfläche und Volumen zu sorgen. Basierend auf dieser Hypothese hat das lange Tempern dann dazu geführt, dass mehr Sn aus dem Volumen an die Oberfläche diffundiert ist als dort in Form von Clustern entnommen wurde. Infolgedessen steigt dadurch dann die Sn-Konzentration, das Fermi-niveau sinkt und die Energie des Dirac-Punkt verschiebt sich hin zu positiven Werten.

Weitere Hinweise darauf, dass die hier beobachteten Temperaturabhängigkeiten auf die Sensitivität der Sn-Konzentration zurückzuführen sind, liefern verschiedene Beobachtungen, die das Wachstum von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ betreffen [62–64]. So konnte bei hohen Temperaturen ($T \approx 550^\circ\text{C}$) eine Dekomposition von $\text{Pb}_{0.7}\text{Sn}_{0.3}\text{Se}$ -Mikrokristallen in PbSe und SnSe beobachtet werden [64]. Bei der Zucht von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ aus einer flüssigen Phase auf einem Si(100)-Substrat führte eine um 10°C gesteigerte Nukleationstemperatur zu einem fast viermal niedrigeren Sn-Anteil [62]. Ein ähnlich starken Einfluss der Temperatur auf die Zusammensetzung wurde bei dem Wachstum mit einem sogenannten hot wall Verdampfer beobachtet. Hier führt eine geringfügige Änderung der Substrattemperatur von 40°C ($T = 330 - 370^\circ\text{C}$) zu einer Drittelung der Sn-Konzentration ($x = 0.3 - 0.1$). Bemerkenswerterweise ist das Temperaturfenster, in dem hier eine glatte Oberfläche präpariert werden kann, mit etwa 30°C ($250^\circ\text{C} \lesssim T \lesssim 280^\circ\text{C}$) von ähnlicher Größe. Dieser Befund lässt sich durch die Annahme ähnlicher Wirkungsmechanismen, hier die starke Änderung des Sn-Anteils durch thermisch aktivierte Mobilität, deuten.

Das Ziel dieser Studie wurde damit nur teilweise erreicht, da keine Probenoberfläche ohne Materialansammlungen präpariert werden konnten. Außerdem hat sich gezeigt, dass in Folge dieser Clusterbildung die Proben nach einigen Zyklen erschöpft sind, da ihre Oberflächen zu rau für STM-Messungen werden. Trotzdem kann die hier genutzte Methode zur *in situ* Reinigung einen wertvollen Beitrag zur Untersuchung von Adatomen auf $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ spielen, da mit ihr eine Möglichkeit entdeckt wurde, das Fermi-niveau ohne externe Störatome zu variieren.

6.2 3d-Übergangsmetallatome auf $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$

Zum Abschluss dieser Arbeit sollen noch erste Zwischenergebnisse aus den Wachstumsstudien (magnetischer) Adatome vorgestellt werden. Als Adsorbate wurden die beiden 3d-Übergangsmetalle Cu und Fe verwendet. Während des Aufdampfens befand sich die Spaltprobe im STM, dessen Spitze zuvor etwa 200 nm von der Probe zurückgezogen wurde. Der geringe Abstand zwischen Probe und Spitze erhöht zwar die Gefahr eines Zusammenstoßes, jedoch auch die Wahrscheinlichkeit nach dem Aufdampfen die selbe Region wieder zu finden. Durch die Aufdampfposition im STM ist einerseits der Abstand zum Verdampfer sehr groß, die Menge der ankommenden Atome pro Fläche daher klein und die Bedeckung somit besser kontrollierbar. Andererseits bleibt aber vor allem der thermische Kontakt zum Kryostat bestehen, sodass die Proben temperatur die ganze Zeit über bei $T \lesssim 5.5 \text{ K}$ gehalten werden kann. Dadurch verfügen die auftreffenden Adatome über zu wenig thermische Energie um auf der Oberfläche mobil zu werden und Cluster zu bilden. Stattdessen bleiben sie an ihrer jeweiligen Ankunftsstelle oder deren direkter Umgebung haften.

Die Topographie und Spektroskopie der Kantenzustände für eine Wachstumsserie mit Cu (a-h) und Spektren für Fe (i-l) sind in Abb. 6.3 gezeigt. Zunächst wurde eine Probe gespalten und die Lage des Dirac-Punkts der reinen Oberflächen mittels der schwarzen Spektren in (b) und (i) bestimmt. Dieser liegt bei den gezeigten Proben bei $E_{\text{DP}^*} = 128 \text{ meV}$ für Cu und bei $E_{\text{DP}^*} = 85 \text{ meV}$ für Fe. Nach dem ersten Aufdampfschritt kann bei beiden Übergangsmetallen eine Verschiebung des Kantenzustands hin zu negativen Energien beobachtet werden. Der dafür verantwortliche Ladungstransfer von Elektronen der Übergangsmetalle hin zu $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ ist jedoch für Cu ausgeprägter. So konnte bei der gleichen Bedeckung ein signifikant größeres ΔE bei Cu beobachtet werden. Aus diesem Grund musste für Fe auf eine Probe mit niedrigerem intrinsischem Dirac-Punkt zurückgegriffen werden, um den Dirac-Punkt an das Fermi-niveau zu dotieren. Bei Proben mit einem Anfangswert der Dirac-Energie von etwa 125 meV war die nötige Adatomdichte so groß, dass für Fe gar kein Oberflächen- oder Kantenzustand mehr beobachtet werden konnte.

Ein typisches STM-Bild nach dem Aufdampfen einer kleinen Menge Adatome ist in (a) zu sehen. Neben einigen größeren Clustern, bei denen es sich wahrscheinlich um Spaltreste handelt, sind die Adsorbate als kleine Erhebungen zu erkennen. Jede einzelne repräsentiert ein Cu-Atom, wobei der Anteil der bedeckten Fläche jedoch erheblich größer erscheint, als die tatsächliche Bedeckung in Monolagen ist. Dieses Phänomen ist nicht nur für Adatome typisch, sondern tritt auch beim Wachstum von Inseln bzw. dünnen Filmen auf. Deshalb wurde die Bedeckung durch das Verhältnis aus der Anzahl der Adsorbate auf einer Fläche (z.B. 50 nm^2) mit der Anzahl der Oberflächenatome der selben Fläche ermittelt. Insgesamt wurden lediglich 0.2-1.8% einer

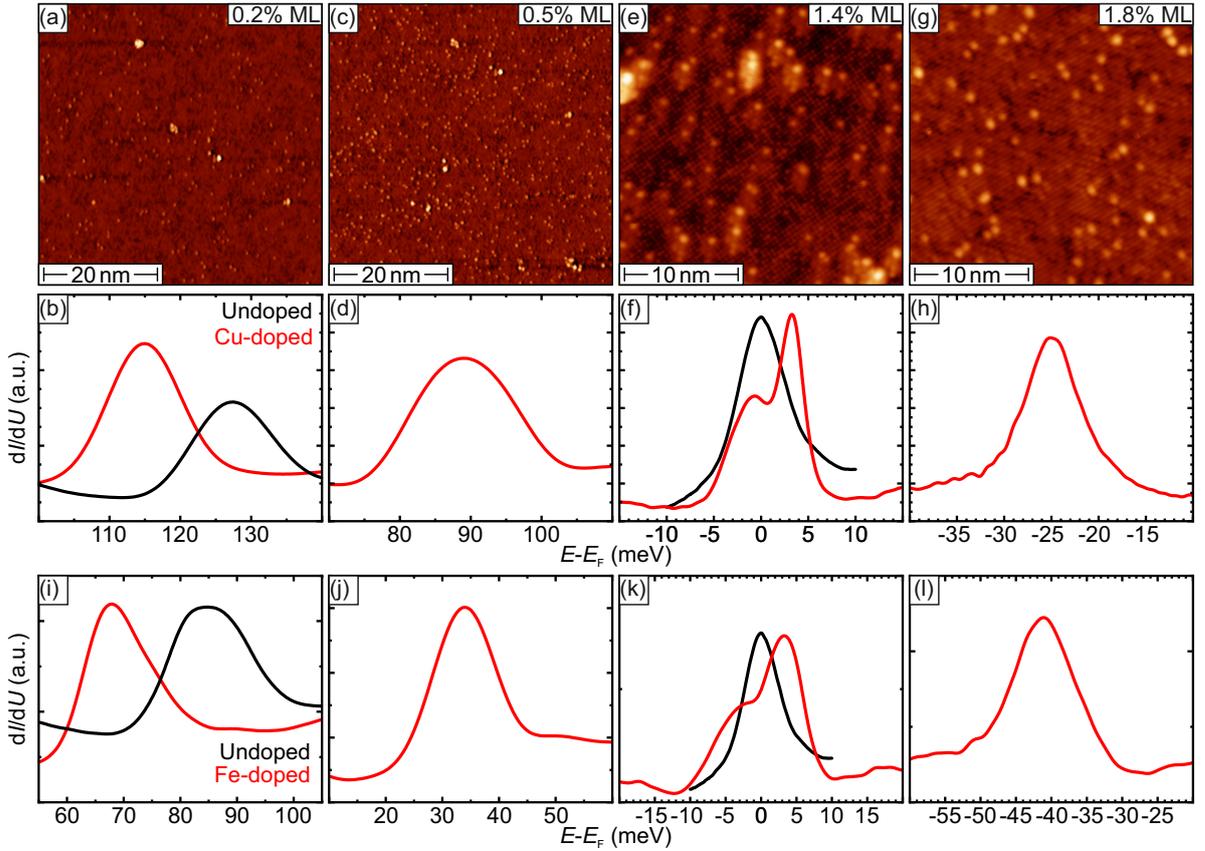


Abbildung 6.3 Ferminiveauverschiebung durch Cu- und Fe-Dotierung: (a-h) STM-Topographien und Spektren von Kantenzuständen einer mit Cu-dotierten Probe für 0.2, 0.5, 1.4 und 1.8% ML-Bedeckung. Die schwarzen Spektren sind auf Proben ohne Cu-Adsorbate unmittelbar nach dem Spalten und vor dem Aufdampfen (b) bzw. nach mehrere Zyklen aus Ionenzerstäubung und Tempern (f) gemessen. (i-l) Spektren für eine analoge Messreihe mit Fe anstatt Cu. In beiden Fällen ist die Bedeckung über die Probe homogen und nur für $E_{\text{DP}^*} \approx E_{\text{F}}$ tritt eine Aufspaltung des Kantenzustands auf. Messparameter: $U = 250$ mV (a,b), $U = 300$ mV (c), $U = 100$ mV (d), $I = 10$ pA (a-d). Stabilisierungsparameter: $U_{\text{Cu}}, U_{\text{Und}} = 250$ mV (b) $U = 200$ mV (c) $U_{\text{Cu}} = 15$ mV & $U_{\text{Und}} = 10$ mV (f), $U = -40$ mV (h), $U_{\text{Und}} = 125$ mV & $U_{\text{Fe}} = 150$ mV (i), $U = 120$ mV (j), $U_{\text{Fe}} = 20$ mV & $U_{\text{Und}} = 10$ mV (k), $U = -20$ mV (l), $I = 200$ pA (b,d,f,h,i-l).

Monolage auf (Pb,Sn)Se aufgedampft. Bereits eine Bedeckung von 0.2% ML Cu führt zu einer Verschiebung des Fermineaus von $\Delta E = 13$ meV und übersteigt damit die natürlichen Fluktuationen des Dirac-Punkts. Noch deutlicher wird dies nach einem weiteren Aufdampfschritt, bei dem die Bedeckung auf 0.5% ML wächst (d). Das Spektrum ist nun um fast 40 meV gegenüber dem undotierten Fall verschoben. Bei einer Bedeckung von etwa 1.4% ML Cu wurde der Punkt erreicht, an dem Dirac-Energie und Fermineau quasi entartet sind. An einigen Stellen der Probe liegt der Dirac-Punkt wenige meV ober und an anderen Stellen wenige meV unterhalb des Fermilevels. Der Kantenzustand zeigt unter diesen Bedingungen eine Aufspaltung in zwei Peaks, deren Größenordnung stets bei wenigen meV liegt. Eine klare Präferenz bezüglich der Asymmetrie konnte nicht festgestellt werden, dafür aber eine hohe Empfindlichkeit gegenüber dem energetischen Abstand zum Fermineau. Schon ab einem Abstand von etwa 5-10 meV

konnte keine Aufspaltung mehr beobachtet werden. Ebenfalls keine Aufspaltung wurde bei der Probe # 9 aus Abb. 6.2 nach dem letzten Zyklus festgestellt. Bei dieser Probe wurde ohne externe Störatome das Fermi-niveau an den Dirac-Punkt gebracht. Das dazugehörige Spektrum eines Kantenzustands ist in (f) und (k) jeweils in schwarz gezeigt. Zu beachten ist, dass die Daten aus unterschiedlichen Messungen stammen, sodass sie für eine übersichtliche Darstellung skaliert wurden. Dadurch bleibt der Verlauf und insbesondere die Anzahl der Peaks erhalten, jedoch ist kein aussagekräftiger Vergleich hinsichtlich der Intensitäten möglich. Die Korrelation zwischen Aufspaltung und der Bedingung $E_{\text{DP}^*} = E_{\text{F}}$ wird durch weiteres Aufdampfen erhärtet. In einem letzten Schritt wurde die Bedeckung auf 1.8% ML erhöht, was zu einer Überdotierung führt. Der Dirac-Punkt liegt nun unterhalb des Fermi-niveaus, wie aus dem Spektrum in (h) hervorgeht. Nun ist jedoch kein Anzeichen mehr für eine Aufspaltung erkennbar.

Analoge Versuche wurden auch mit Fe-Adatomen durchgeführt. Wie an den Spektren in (i-l) zu sehen ist, verschiebt sich zunächst ebenfalls der Dirac-Punkt hin zu negativen Energien (j). Sobald er in einem engen Intervall von $E_{\text{DP}^*} \approx \pm 5 \text{ meV}$ liegt (k), kann erneut eine Aufspaltung beobachtet werden. Auch in diesem Fall führt eine Überdotierung wieder zu einem Verschwinden der Aufspaltung (l). Ein völlig anderes Verhalten wurde jedoch beobachtet, als statt einer konstanten Bedeckung ein Fe-Gradient über einer Probe gewachsen wurde. Auch bei dieser Probe konnte eine Aufspaltung gezeigt werden, sobald der Kantenzustand nahe genug am Fermi-niveau lag. Zeitgleich gab es jedoch Probenbereiche mit höherer Fe-Bedeckung, sodass dort die Probe lokal überdotiert war. In diesen Bereichen wurde jedoch kein Einzelpeak beobachtet, sondern selbst bei $E = -36 \text{ meV}$ konnte noch eine klare Aufspaltung in zwei Peaks festgestellt werden. Weil bei diesen Messungen die Bedeckung auf der Probe nicht homogen war, sind die Daten nicht in Abb. 6.3 subsumiert.

Sowohl die geringere Größe der Aufspaltung als auch ihre hohe Sensitivität auf die exakte Lage der Dirac-Energie bei Proben mit konstanter Bedeckung, lassen auf einen völlig anderen Kopplungsmechanismus als im vorherigen Kapitel schließen. Da die Aufspaltung im Falle der intrinsisch dotierten Probe nicht beobachtet wurde, könnten die mit den $3d$ -Übergangsmetallen einhergehenden magnetischen Momente eine Schlüsselrolle spielen. Von besonderem Interesse ist dabei die anscheinend langreichweitigen Wechselwirkungen auf der Probe mit Fe-Gradient, die möglicherweise durch die robusten, spinpolarisierten Dirac-Elektronen vermittelt wird. Sofern eine magnetische Ordnung zugrunde liegt könnte es sich dann um einen völlig neuartigen Wechselwirkungsmechanismus handeln. Um diesen Fragen nachzugehen sind nicht nur weitere Messungen nötig, sondern auch umfassende Berechnungen, um zu verstehen, wie die $3d$ -Übergangsmetalle mit dem Substrat hybridisieren bzw. wechselwirken und ob die Ursache dieser Aufspaltung im induzierten Magnetismus liegt.

7 Zusammenfassung

In dieser Arbeit wurden intrinsisch auftretende Wechselwirkungsphänomene zwischen eindimensionalen, topologisch geschützten Kantenzuständen entdeckt und näher untersucht. Die Kantenzustände existieren nur an Stufen mit einer ungeraden Anzahl an fehlenden Atomlagen, da es dort zu einem Bruch der Translationssymmetrie kommt. Durch verschiedene geometrische Anordnungen von Stufenkanten konnten so verschiedene relevante Kopplungsparameter variiert werden. Während es bei Konfigurationen mit zusammenlaufenden Stufenkanten zu einer kontinuierlichen Verringerung des Stufe zu Stufe Abstands d und damit einer stetigen Zunahme der Kopplungsstärke kommt, so hängt bei sich kreuzenden Stufen die Projektion der Zustände aufeinander von dem eingeschlossenen Winkel ab. Durch die Rastertunnelspektroskopie wurden diese topographischen Informationen mit den spektroskopischen verknüpft. Der ungestörte Kantenzustand wird dabei in STS-Spektren durch ein ausgeprägtes Maximum an der Dirac-Punkt-Energie charakterisiert.

Abhängig von der Stärke der Wechselwirkung, respektive von d , konnten drei Kopplungsregime identifiziert werden. (I) Das Einsetzen der Kopplung tritt ab $d \lesssim 25$ nm auf und macht sich zunächst durch eine Verbreiterung des Peaks bei abnehmendem Maximalwert bemerkbar. Aus diesem breiten Peak bildet sich dann nach und nach eine Schulter aus, welche zu einem zweiten lokalen Maximum wird. In diesem Übergangsbereich koexistieren Spektren, denen tendenziell ein einzelner Peak zugeordnet wird, mit solchen die bereits einen Doppelpeak zeigen. (II) Ab einem Abstand von etwa 15 nm ist dann durchgehend ein Doppelpeak zu sehen, dessen Aufspaltung mit zunehmender Kopplung wächst. Dabei zeigt sich eine exponentielle Abhängigkeit vom d , was sich durch die Überlappung zweier exponentiell abklingender Wellenfunktionen erklären lässt. Der Maximalwert der Intensität der beiden Peaks bleibt zunächst annähernd konstant, allerdings schwankt das Intensitätsverhältnis zwischen den Doppelpeaks sehr stark. Sowohl symmetrische wie auch asymmetrische Verhältnisse wurden beobachtet, ohne dass eine Korrelation erkennbar war. (III) Die Intensität beider Peaks nimmt jedoch schnell ab, wenn der Abstand sich dem Wert $d \approx 5$ nm nähert oder ihn unterschreitet. Dann werden Aufspaltungen von mehr als 40 meV beobachtet, was dazu führt, dass die Doppelpeaks immer stärker mit den vHS überlappen. Dementsprechend konvergieren die Spektren der Kantenzustände gegen das Spektrum über einer Terrasse.

Auffällig an der Aufspaltung war deren asymmetrische Charakter bezüglich der Energie. Während in den Regimen I und II die Position von einem Peak nur wenige meV von der Dirac-Energie abweicht, verschiebt sich der andere Peak immer weiter. Erst im Bereich III beginnt sich auch der zweite Peak zu verschieben. Dieses Verhalten konnte erfolgreich mit einem Modell aus zwei Quantentöpfen mit variablem Abstand simuliert werden. Dabei musste von der Annahme identischer Quantentöpfe abgewichen werden, indem z.B die Tiefen nicht exakt identisch sind. Solche Unterschiede können z.B auf die stochastischen Schwankungen der Sn-Konzentration zurückgeführt werden.

Bei sich kreuzenden Stufenkanten kommt es am Schnittpunkt zu einer Doppelstufe, sodass kein Kantenzustand zu erwarten ist. Abhängig vom Winkel klingen die Kantenzustände ab einem Abstand von 10-15 nm ab. Dabei konnte nur für rechte Winkel gar keine Anzeichen für eine Aufspaltung in einen Doppelpack gefunden werden. Dies deutet auf eine Orthogonalität der beteiligten Kantenzustände bezüglich einer Quantenzahl hin, bei der es sich um den orbitalen Charakter (Se p_x und p_y) der beteiligten Dirac-Kegel handelt könnte.

Im zweiten Teil dieser Arbeit wurde die Oberfläche von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ durch Zyklen aus Ionenzerstäubung und Tempern gereinigt. Dabei konnte eine reversible Verschiebung des Fermi-niveaus erreicht werden, indem die Proben-temperatur oder die Dauer des Temperns angepasst wurde. Sowohl durch höhere Temperaturen, als auch durch längeres Tempern kann eine p-Dotierung des Kristalls erreicht werden, was zu einem Anstieg der Dirac-Energie relativ zum Fermi-niveau führt. Diese ist mit großer Wahrscheinlichkeit auf eine Änderung der Sn-Konzentration und der damit verbundenen Dichte an Defektelektronen zurückzuführen.

Als mögliche Senke für überschüssiges Sn kommen die beobachteten Materialansammlungen an der Oberfläche in Frage. Es konnte keine Probe ohne solche Cluster präpariert werden, wobei deren Anzahl mit der Temperatur des Temperns und der Anzahl der durchlaufenen Zyklen stark zunimmt. Dadurch ist nicht nur die Lebensdauer der Proben begrenzt, sondern auch das Temperaturfenster zur Probenpräparation relativ schmal. Lediglich in einem engen Intervall von $T \approx (265 \pm 15)^\circ\text{C}$ konnten einerseits großflächig atomar glatte Terrassen und andererseits eine hinreichend geringen Dichte an Clustern gefunden werden.

Durch Aufdampfen der $3d$ -Übergangsmetalle Fe und Cu konnten Proben n-dotiert und der Dirac-Punkt an das Fermi-niveau verschoben werden. Sobald deren energetischer Abstand betragsmäßig kleiner als 5 meV ist, konnte an freistehenden Stufenkanten eine feine Aufspaltung von wenigen meV in einen Doppelpack beobachtet werden. Dieser verschwindet bei Überdotierung und konnte auch nicht für Proben beobachtet werden, bei denen das Fermi-niveau durch Tempern und Ionenzerstäubung an die Dirac-Energie angeglichen wurde. Dieser Befund ist ein erster experimenteller Hinweis auf eine magnetische Natur der 1D-Kantenzustände.

Summary

In the present work, intrinsic interaction phenomena between one-dimensional, topologically protected edge states were discovered and investigated. Edge states only exist at steps with an odd number of missing atomic layers, because there the translational symmetry is broken. By investigating several different geometric step edge structures, the relevant coupling parameters could be varied. Since configurations with converging step edges reduce the step-to-step distance d continuously, the coupling strength increases monotonically. In the case of crossing steps the projection of the states onto one another depends on the included angle. Using scanning tunneling spectroscopy, this topographical information was linked to the spectroscopic information. The undisturbed edge state is characterized in STS spectra by a pronounced maximum at the Dirac point energy.

Depending on the strength of the interaction, or d , three coupling regimes could be identified. (I) The onset of coupling occurs from $d \lesssim 25$ nm and is noticeable as a broadening of the peak while the maximum value decreases. A shoulder then gradually forms from this broad peak, which becomes a second local maximum. In this transition region, spectra that tend to show a single peak coexist with those that already exhibit a double peak. (II) Below a distance of about 15 nm, a double peak can always be seen. The splitting grows with increasing coupling. There is an exponential dependence on d , which can be explained by the overlapping of two exponentially decaying wave functions. The maximum value of the intensity for both peaks remain approximately constant, but the intensity ratio between them fluctuates heavily. Both symmetrical and asymmetrical ratios were observed, but no correlation is recognized. (III) However, as the distance approaches or falls below $d \approx 5$ nm, the intensity of both peaks decreases rapidly. Splittings of more than 40 meV are observed, which causes the double peaks to overlap with the vHS. As a consequence the spectra form of the step edge converge to the spectra form over a terrace.

Is notable that the splitting has an asymmetric character in terms of energy. While in regimes I and II the position of one peak deviates from the Dirac energy only by a few meV, the other peak is shifting. The second peak starts to shift in regime III. This behavior could be successfully simulated with a model consisting of two quantum wells with variable spacing. However the assumption of identical quantum wells had to be rejected (e.g. the depths are not exactly identical). Such differences can be induced due to stochastic fluctuations in the Sn-concentration.

In the case of crossing step edges, a double step occurs at the point of intersection. Therefore no edge state is to be expected. Depending on the angle, the edge states decay below a distance of 10-15 nm. For right angles no indications of a splitting into a double peak could be found. This is a hint for orthogonality of the involved edge states with respect to a quantum number, which could be the orbital character (See p_x and p_y) of the involved Dirac cones.

In the second part of this work, the surface of $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ was cleaned by ion sputtering and annealing cycles. A tuneable shift of the Fermi level can be achieved by adjusting the sample temperature or the annealing time. Both a higher temperature and a longer annealing time lead to a p-doped crystal. In this case the Dirac energy rises, compared to the Fermi level. This is most likely due to a change in the Sn-concentration and the associated changes of the hole density.

The dissolved Sn could form the observed accumulations of material on the surface. It was not possible to prepare a sample without such clusters. The number of these clusters increases rapidly with the annealing temperature and the number of completed cycles. Due to this reason, the temperature window for sample preparation is relatively narrow and the sample lifetime is limited. Only in the small range between $T \approx (265 \pm 15)^\circ\text{C}$ could large areas of atomically smooth terraces be found, coexisting with a sufficiently low density of clusters.

The Dirac point of initially p-doped samples are shifted towards the Fermi level by evaporation of the 3d-transition metals Fe and Cu. As soon as the energetic distance between Fermi and Dirac energy is in the range of 5 meV, a fine splitting of a few meV into a double peak could be observed at free-standing step edges. This splitting disappears with overdoping and could not be observed for samples in which the Fermi level was adjusted to the Dirac energy by annealing and sputtering. This finding is a first experimental hint for a magnetic nature of the 1D edge states.

Anhang

Abbildungsverzeichnis

2.1	Su–Schrieffer–Heeger Modell	7
3.1	Übersicht STM–Theorie	19
3.2	Tunneleffekt	21
3.3	dI/dV –Beiträge	28
4.1	Gitter, Brillouin–Zone und Dirac–Kegel von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$	34
4.2	Topologischer Phasenübergang von $\text{Pb}_{1-x}\text{Sn}_x(\text{Se},\text{Te})$	38
4.3	Topologischer Phasenübergang durch Gleitspiegelsymmetrie	39
4.4	Topology der Fermifläche und STS–Spektrum von $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$	41
4.5	1D Kantenzustände der (001)–Oberfläche	43
4.6	1D Kantenzustände an ungeraden Stufen	44
4.7	Iaia–Modell für Kantenzustände in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$	46
4.8	Richtungsabhängigkeit von Kantenzustände in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$	48
5.1	Analyse gerader Stufenkanten I	56
5.2	Analyse gerader Stufenkanten II	57
5.3	Klassifizierung durch Anpassung mehrerer Gaussfunktionen	60
5.4	Kopplungseffekte an zusammenlaufenden, inselartigen ungeraden Stufenkanten	61
5.5	Kopplungseffekte an zusammenlaufenden Fehlstelleninseln mit ungeraden Stufenkanten	66
5.6	Doppelquantentopf–Modell zusammenlaufender Stufenkanten	69
5.7	Parallel verlaufende, gekoppelte Stufenkanten	75
5.8	Kreuzende Stufenkanten	79
5.9	Kreuzung zwischen ungerader Stufenkante und Versetzungslinie	83
5.10	Aufspaltung des Kantenzustands an Schraubversetzungen	86
6.1	Temperaturabhängigkeit der Oberfläche	95
6.2	Ferminiveauverschiebung durch Ionenzerstäubung und Tempern	98
6.3	Ferminiveauverschiebung durch Cu- und Fe–Dotierung	103

Tabellenverzeichnis

2.1	AZ–Klassifizierung	11
2.2	Symorphe Erweiterung der AZ–Klassifizierung	14
2.3	Nicht symorphe Erweiterung der AZ–Klassifizierung	16

Abkürzungsverzeichnis

AES	Auger Electron Spectroscopy	Augerelektronenspektroskopie
ARPES	Angle resolved photoemission spectroscopy	Winkelaufgelöste Photoelektronenspektroskopie
AZ	Altland Zirnbauer	Altland–Zirnbauer
BZ	Brillouin Zone	Brillouin Zone
DFT	Density Functional Theory	Dichtefunktionaltheorie
dI/dU	Differential conductivity	differenzielle Leitfähigkeit
DOS	Density Of States	Zustandsdichte
FFT	Fast Fourier Transform	schnelle Fouriertransformation
FWHM	Full Width At Half Maximum	Halbwertsbreite
HTF	High temperature flash	Hoch–Temperatur–Puls
LDOS	Local Density Of States	Lokale Zustandsdichte
MBE	Molecular beam epitaxy	Molekularstrahlepitaxie
ML	MonoLayer	Monolage
NSS	NonSymmorphic Symmetry	nicht symorphe Symmetrie
PHS	Particle Hole Symmetry	Teilchen–Loch–Symmetrie
QHE	Quantum Hall Effect	Quanten Hall Effekt
RKKY	Ruderman Kittel Kasuya Yosida	Ruderman–Kittel–Kasuya–Yosida
SEM	Scanning electron microscopy	Rasterelektronenmikroskopie
SGL	Schrödinger equation	Schrödingergleichung
SML	Spin–momentum locking	Spin–Impuls–Kopplung
SSH	Su Schrieffer Heeger	Su Schrieffer Heeger
STM	Scanning tunneling microscope	Rastertunnelmikroskop
STS	Scanning tunneling spectroscopy	Rastertunnelspektroskopie
TCI	Topological Crystalline Insulator	Kristalliner topologischer Isolator
TI	Topological Insulator	Topologischer Isolator
TKI	Topological Kondo Insulator	Topologischer Kondo-Isolator
TRS	Time Reversal Symmetry	Zeit-Umkehr-Symmetrie
UHV	Ultra-high vacuum	Ultrahochvakuum
vHS	Van Hove Singularity	Van Hove Singularität
WKB	Wentzel Kramers Brillouin	Wentzel Kramers Brillouin
WS	Topological Dirac/Weyl Semimetals	Dirac/Weyl–Halbmetallen
XPS	X-ray Photoelectron Spectroscopy	Röntgenphotoelektronenspektroskopie

Verzeichnis eigener Publikationen

- Jung, J., Odobesko, A. B., Boshuis, R., Szczerbakow, A., Story, T. und Bode, M.: „Systematic Investigation of the Coupling between One-Dimensional Edge States of a Topological Crystalline Insulator“. *Phys. Rev. Lett.* **126** 236402 (2021)
- Boshuis, R., Odobesko, A. B., Friedrich, F., Jung, J. und Bode, M.: „Comparative growth study of ultra-thin Bi films on clean and oxygen-reconstructed Nb(110)“. *Phys. Rev. Mat.* **5** 054801 (2021)
- Odobesko, A. B. , Haldar, S., Wilfert, S., Hagen, J., Jung, J., Schmidt, N., Sessi, P., Vogt, M., Heinze, S. und Bode, M.: „Preparation and electronic properties of clean superconducting Nb(110) surfaces“. *Phys. Rev. B* **99** 115437 (2019)
- Islam, M.F., Canali, C.M., Pertsova, A., Balatsky, A., Mahatha, S. K., Carbone, C., Barla, A., Kokh, K.A., Tereshchenko, O.E., Jimenez, E., Brookes, N., Gargiani, P., Valvidares, M., Schatz, S., Peixoto, T. R. F., Bentmann, H., Reinert F., Jung, J., Bathon, T., Fauth, K., Bode, M. und Sessi, P.: „Systematics of electronic and magnetic properties in the transition metal doped Sb₂Te₃ quantum anomalous Hall platform“. *Phys. Rev. B* **97** 155429 (2018)

Literatur

- [1] Klitzing, K. v., Dorda, G. und Pepper, M.: „New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance“. *Phys. Rev. Lett.* **45**, 494–497 (1980) (siehe S. 1).
- [2] Qi, X.-L. und Zhang, S.-C.: „The quantum spin Hall effect and topological insulators“. *Physics Today* **63**, 33–38 (2010) (siehe S. 1).
- [3] Maciejko, J., Hughes, T. L. und Zhang, S.-C.: „The Quantum Spin Hall Effect“. *Annual Review of Condensed Matter Physics* **2**, 31–53 (2011) (siehe S. 1).
- [4] Kane, C. L. und Mele, E. J.: „Quantum Spin Hall Effect in Graphene“. *Physical Review Letters* **95**, 226801 (2005) (siehe S. 1).
- [5] Bernevig, B. A. und Zhang, S.-C.: „Quantum Spin Hall Effect“. *Physical Review Letters* **96**, 106802 (2006) (siehe S. 1).
- [6] Bernevig, B. A., Hughes, T. L. und Zhang, S.-C.: „Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells“. *Science* **314**, 1757–1761 (2006) (siehe S. 1).
- [7] König, M., Wiedmann, S., Brüne, C. et al.: „Quantum Spin Hall Insulator State in HgTe Quantum Wells“. *Science* **318**, 766–770 (2007) (siehe S. 1, 2).
- [8] Fu, L., Kane, C. L. und Mele, E. J.: „Topological Insulators in Three Dimensions“. *Physical Review Letters* **98**, 106803 (2007) (siehe S. 1).
- [9] Moore, J. E. und Balents, L.: „Topological invariants of time-reversal-invariant band structures“. *Physical Review B* **75**, 121306(R) (2007) (siehe S. 1).
- [10] Roy, R.: „Topological phases and the quantum spin Hall effect in three dimensions“. *Physical Review B* **79**, 19 (2009) (siehe S. 1).
- [11] Hsieh, D., Qian, D., Wray, L. et al.: „A topological Dirac insulator in a quantum spin Hall phase“. *Nature* **452**, 970–974 (2008) (siehe S. 1).
- [12] Hsieh, D., Xia, Y., Wray, L. et al.: „Observation of Unconventional Quantum Spin Textures in Topological Insulators“. *Science* **323**, 919–922 (2009) (siehe S. 1).

- [13] Wan, X., Turner, A. M., Vishwanath, A. et al.: „Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates“. *Phys. Rev. B* **83**, 205101 (2011) (siehe S. 1).
- [14] Xu, G., Weng, H., Wang, Z. et al.: „Chern Semimetal and the Quantized Anomalous Hall Effect in HgCr_2Se_4 “. *Physical Review Letters* **107**, 186806 (2011) (siehe S. 1).
- [15] Yang, K.-Y., Lu, Y.-M. und Ran, Y.: „Quantum Hall effects in a Weyl semimetal: Possible application in pyrochlore iridates“. *Phys. Rev. B* **84**, 075129 (2011) (siehe S. 1).
- [16] Neupane, M., Xu, S.-Y., Sankar, R. et al.: „Observation of a three-dimensional topological Dirac semimetal phase in high-mobility Cd_3As_2 “. *Nature Communications* **5**, 3786 (2014) (siehe S. 1).
- [17] Liu, Z. K., Zhou, B., Zhang, Y. et al.: „Discovery of a Three-Dimensional Topological Dirac Semimetal, Na_3Bi “. *Science* **343**, 864–867 (2014) (siehe S. 1).
- [18] Fu, L.: „Topological Crystalline Insulators“. *Physical Review Letters* **106**.10 (März 2011) (siehe S. 1, 2, 35, 37).
- [19] Hsieh, T. H., Lin, H., Liu, J. et al.: „Topological crystalline insulators in the SnTe material class“. *Nature Communications* **3**, 982 (Jan. 2012) (siehe S. 1, 35, 37, 41, 45, 50).
- [20] Xu, S.-Y., Liu, C., Alidoust, N. et al.: „Observation of a topological crystalline insulator phase and topological phase transition in $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ “. *Nature Communications* **3**, 1192 (Jan. 2012) (siehe S. 1, 33, 37, 38).
- [21] Dziawa, P., Kowalski, B. J., Dybko, K. et al.: „Topological crystalline insulator states in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ “. *Nature Materials* **11**.12, 1023–1027 (Sep. 2012) (siehe S. 1, 33, 35, 37, 38).
- [22] Tanaka, Y., Ren, Z., Sato, T. et al.: „Experimental realization of a topological crystalline insulator in SnTe“. *Nature Physics* **8**, 800–803 (Sep. 2012) (siehe S. 1, 33, 37).
- [23] Hasan, M. Z. und Kane, C. L.: „Colloquium: Topological insulators“. *Reviews of Modern Physics* **82**, 3045–3067 (2010) (siehe S. 2, 10).
- [24] Kane, C. L. und Moore, J. E.: „Topological Insulators“. *Physics World* **24**, 32 (2011) (siehe S. 2).
- [25] Qi, X.-L. und Zhang, S.-C.: „The quantum spin Hall effect and topological insulators“. *Physics Today* **63**, 33–38 (2010) (siehe S. 2).
- [26] Moore, J. E.: „The birth of topological insulators“. *Nature* **464**, 194–198 (2010) (siehe S. 2).

- [27] Maciejko, J., Hughes, T. L. und Zhang, S.-C.: „The Quantum Spin Hall Effect“. *Annual Review of Condensed Matter Physics* **2**, 31–53 (2011) (siehe S. 2).
- [28] Ortmann, F., Roche, S. und Valenzuela, S.: „Topological Insulators: Fundamentals and Perspectives“. Wiley-VCH (2015) (siehe S. 2).
- [29] Altland, A. und Zirnbauer, M. R.: „Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures“. *Phys. Rev. B* **55**, 1142–1161 (Jan. 1997) (siehe S. 2, 10).
- [30] Ryu, S., Schnyder, A. P., Furusaki, A. et al.: „Topological insulators and superconductors: tenfold way and dimensional hierarchy“. *New Journal of Physics* **12.6**, 065010 (Juni 2010) (siehe S. 2, 10).
- [31] Asbóth, J. K., Oroszlány, L. und Pályi, A.: „The Su-Schrieffer-Heeger (SSH) Model“. Cham: Springer International Publishing (2016), S. 1–22 (siehe S. 5–8).
- [32] Trifunovic, L. und Brouwer, P.: „Bott periodicity for the topological classification of gapped states of matter with reflection symmetry“. *Phys. Rev. B* **96**, 195109 (Nov. 2017) (siehe S. 10).
- [33] Shiozaki, K. und Sato, M.: „Topology of crystalline insulators and superconductors“. *Physical Review B* **90**, 165114 (Okt. 2014) (siehe S. 13, 14, 40).
- [34] Shiozaki, K., Sato, M. und Gomi, K.: „Topology of nonsymmorphic crystalline insulators and superconductors“. *Physical Review B* **93**, 195 (Mai 2016) (siehe S. 13, 16, 40).
- [35] Araújo, A. L., Ferreira, G. J. und Schmidt, T. M.: „Suppressed topological phase transitions due to nonsymmorphism in SnTe stacking“. *Scientific Reports* **8.1** (Juni 2018) (siehe S. 17, 39).
- [36] Bai, C.: „Scanning Tunneling Microscopy and Its Applications“. Shanghai Scientific & Technical Publishers und Springer-Verlag Berlin Heidelberg (1992) (siehe S. 19).
- [37] Voigtlaender, B.: „Scanning Probe Microscopy. Atomic Force Microscopy and Scanning Tunneling Microscopy“. Springer-Verlag Berlin Heidelberg (2015) (siehe S. 19, 24, 28).
- [38] Wiesendanger, R.: „Scanning Probe Microscopy and Spectroscopy. Methodes and Applications“. Cambridge University Press (1994) (siehe S. 19, 25).
- [39] Sessi, P., Di Sante, D., Szczerbakow, A. et al.: „Robust spin-polarized midgap states at step edges of topological crystalline insulators“. *Science* **354**, 1269–1273 (2016) (siehe S. 34, 42–44, 53, 59, 70, 84, 91).

- [40] Zeljkovic, I., Okada, Y., Serbyn, M. et al.: „Dirac mass generation from crystal symmetry breaking on the surfaces of topological crystalline insulators“. *Nature Materials* **14.3**, 318–324 (März 2015) (siehe S. 35).
- [41] Okada, Y., Serbyn, M., Lin, H. et al.: „Observation of Dirac Node Formation and Mass Acquisition in a Topological Crystalline Insulator“. *Science* **341.6153**, 1496–1499 (2013) (siehe S. 35).
- [42] Liu, J., Duan, W. und Fu, L.: „Two types of surface states in topological crystalline insulators“. *Physical Review B* **88.24** (Dez. 2013) (siehe S. 35, 36, 41).
- [43] Polley, C. M., Buczko, R., Forsman, A. et al.: „Fragility of the Dirac Cone Splitting in Topological Crystalline Insulator Heterostructures“. *ACS Nano* **12.1**, 617–626 (Jan. 2018) (siehe S. 36).
- [44] Dimmock, J. O., Melngailis, I. und Strauss, A. J.: „Band Structure and Laser Action in $\text{Pb}_x\text{Sn}_{1-x}\text{Te}$ “. *Phys. Rev. Lett.* **16**, 1193–1196 (Juni 1966) (siehe S. 37).
- [45] Conklin, J. B., Johnson, L. E. und Pratt, G. W.: „Energy Bands in PbTe “. *Phys. Rev.* **137**, A1282–A1294 (Feb. 1965) (siehe S. 37).
- [46] Strauss, A. J.: „Inversion of Conduction and Valence Bands in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ Alloys“. *Phys. Rev.* **157**, 608–611 (Mai 1967) (siehe S. 37).
- [47] Brzezicki, W., Wysokiński, M. M. und Hyart, T.: „Topological properties of multilayers and surface steps in the SnTe material class“. *Physical Review B* **100.12** (Sep. 2019) (siehe S. 39, 40, 50).
- [48] Ozawa, H., Yamakage, A., Sato, M. et al.: „Topological phase transition in a topological crystalline insulator induced by finite-size effects“. *Physical Review B* **90.4** (Juli 2014) (siehe S. 40).
- [49] Wang, J., Wang, N., Huang, H. et al.: „Electronic properties of SnTe -class topological crystalline insulator materials“. *Chinese Physics B* **25.11**, 117313 (Nov. 2016) (siehe S. 41).
- [50] Sessi, P., Biswas, R. R., Bathon, T. et al.: „Dual nature of magnetic dopants and competing trends in topological insulators“. *Nature Communications* **7**, 12027 (Juni 2016) (siehe S. 43).
- [51] Iai, D., Wang, C.-Y., Maximenko, Y. et al.: „Topological nature of step-edge states on the surface of the topological crystalline insulator $\text{Pb}_{0.7}\text{Sn}_{0.3}\text{Se}$ “. *Physical Review B* **99.155116** (Apr. 2019) (siehe S. 44–47, 53, 56, 57, 91).
- [52] Wang, Y. J., Tsai, W.-F., Lin, H. et al.: „Nontrivial spin texture of the coaxial Dirac cones on the surface of topological crystalline insulator SnTe “. *Physical Review B* **87.23** (Juni 2013) (siehe S. 45).

- [53] Rechciński, R. und Buczko, R.: „Topological states on uneven (Pb,Sn)Se (001) surfaces“. *Physical Review B* **98.24** (Dez. 2018) (siehe S. 46, 48).
- [54] Ryu, S. und Hatsugai, Y.: „Topological Origin of Zero-Energy Edge States in Particle-Hole Symmetric Systems“. *Physical Review Letters* **89.7** (Juli 2002) (siehe S. 47).
- [55] Chiu, C.-K., Teo, J. C., Schnyder, A. P. et al.: „Classification of topological quantum matter with symmetries“. *Reviews of Modern Physics* **88.3** (Aug. 2016) (siehe S. 47).
- [56] Smoluchowski, R.: „Anisotropy of the Electronic Work Function of Metals“. *Physical Review* **60.661** (1941) (siehe S. 74).
- [57] Saghir, M., Walker, M., McConville, C. F. et al.: „SnTe microcrystals: Surface cleaning of a topological crystalline insulator“. *Appl. Phys. Lett.* **108.6**, 061602 (Apr. 2021) (siehe S. 93).
- [58] Kanno, I., Kotera, H. und Wasa, K.: „Handbook of sputter deposition technology. fundamentals and applications for functional thin films, nanomaterials, and MEMS“. Elsevier (2012) (siehe S. 93, 94).
- [59] Guo, H., Yan, C.-H., Liu, J.-W. et al.: „Topological crystalline insulator $\text{Pb}_x\text{Sn}_{1-x}\text{Te}$ thin films on SrTiO_3 (001) with tunable Fermi levels“. *APL Materials* **2.5**, 056106 (Apr. 2021) (siehe S. 100).
- [60] Assaf, B. A., Phuphachong, T., Volobuev, V. V. et al.: „Magneto-optical determination of a topological index“. *npj Quantum Materials* **2.1**, 26 (Mai 2017) (siehe S. 100).
- [61] Zimin, S. P., Gorlachev, E. S., Amirov, I. I. et al.: „Sputtering rates of lead chalcogenide-based ternary solid solutions during inductively coupled argon plasma treatment“. *Semiconductor Science and Technology* **26.10**, 105003 (Aug. 2011) (siehe S. 101).
- [62] Sachar, H. K., Chao, I., McCann, P. J. et al.: „Growth and characterization of PbSe and $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ on Si (100)“. *Journal of Applied Physics* **85.10**, 7398–7403 (2021) (siehe S. 101).
- [63] Martinez-Collazo, A., Santiago, F. und Chu, T.-K.: „Growth and material characterization of PbSnSe/PbSe heterostructures“. *Journal of Vacuum Science & Technology A* **7.3**, 1420–1423 (1989) (siehe S. 101).
- [64] Saghir, M., Sanchez, A. M., Hindmarsh, S. A. et al.: „Nanomaterials of the Topological Crystalline Insulators, $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ and $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ “. *Crystal Growth & Design* **15.11** (2015) (siehe S. 101).

Danksagung

Die Promotion umfasst nicht nur die wissenschaftliche Arbeit und die Dissertation, sondern man lernt in dieser Zeit auch viel über sich selbst. An dieser Stelle möchte ich all jenen danken, die mich dabei unterstützt und so einen wichtigen Beitrag zu dem Ergebnis dieser Arbeit beigetragen haben:

- *Prof. Dr. Matthias Bode* für die freundliche Aufnahme am Lehrstuhl und die Möglichkeit, meine Dissertation in der Experimentellen Physik II anzufertigen. Seine Expertise und sein Rat waren mir stets eine große Hilfe.
- *Dr. Artem Odobesko und Dr. Paolo Sessi* für die großartige fachliche Unterstützung und die mal mehr, mal weniger wissenschaftlichen Diskussionen.
- Den vielen anderen Beteiligten, Kooperationspartnern und Koautoren.
- *Dem LT1/1K-Team*, allen ehemaligen sowie aktuellen Mitgliedern, für die großartige Unterstützung im Labor. Eure bedingungslose Hilfsbereitschaft war insbesondere in den Phasen des Stillstands eine Inspiration.
- *Anita Gebhardt* für ihre hilfsbereite und angenehme Art.
- Allen aktuellen und ehemaligen Mitgliedern des Lehrstuhls für so ziemlich alles: Die kollegiale Atmosphäre, die Hilfsbereitschaft, die vielen Gelegenheiten Kuchen zu verspeisen, tolle Gespräche und und und...
- *Meinen Freunden* für so viele unvergessliche Jahre in Würzburg.
- *Meiner Familie* für die bedingungslose und überwältigende Unterstützung.