Title
Topological surface states in cadmium arsenide (001) thin films

Author
Kealhofer, David

Publication Date
2020

UNIVERSITY of CALIFORNIA
Santa Barbara

Topological surface states in cadmium arsenide (001) thin films

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

by

David A. Kealhofer

Committee in charge:

Professor Susanne Stemmer, Co-chair

Professor Fyl Pincus, Co-chair

Professor Leon Balents

Professor Andrea Young

December 2020

The dissertation of David A. Kealhofer is approved:

Leon Balents

Andrea Young

Fyl Pincus, Co-chair

Susanne Stemmer, Co-chair

December 2020

To you, the reader—may you find something to interest you!

Acknowledgements

The work reproduced or adapted here from refs. [38] and [37], to which APS retains the copyright, as well as the unpublished manuscript provisionally titled, "The topological insulator picture in (001)-oriented Cd_3As_2 films," has been performed in collaboration with the other authors of those works, Timo Schumann, Luca Galletti, Manik Goyal, Honggyu Kim, Tyler Pardue, Alexey Suslov, and Susanne Stemmer. A portion of this work (the parts of Chapter 4 discussing transport at fields higher than 14 T) was performed at the National High Magnetic Field Laboratory, which is supported by the National Science Foundation Cooperative Agreement No. DMR-1644779 and the State of Florida. Another portion of this work, the x-ray diffraction, atomic force microscopy, electron microscopy, and the low-temperature transport studies not performed at the NHMFL, made use of the shared facilities of the UCSB MRSEC (NSF DMR 1720256). Please see the Acknowledgements of refs. [38] and [37] for a full accounting of the funding agencies and programs that have supported this work.

While page v of my dissertation is clearly a suboptimal place to try to accurately sum up my feelings of gratitude and affection for the many people who have made my life better throughout my time in graduate school, I am truly grateful for the many people who can be said to have supported, aided, or rescued me over the past years in Santa Barbara, and here are a few of them:

Susanne Stemmer, my advisor, who gave me a real second chance at graduate school, and who has been a patient and encouraging mentor and an inspiring role model, besides providing for all of my scientific and professional needs

Everyone I've worked with in the Stemmer group, particularly my fellow cadmium arsenide enthusiasts, Timo (a great mentor and friend), Manik (to whom I am bound by astrology, a model of tirelessness), Luca (a wise and supportive colleague, cool in a crisis), Evan (a reminder of why we worship youth), and Alex (our hope for the future), with whom I've spent Many Good Evenings, from whom I have learned about as much as you may read in this dissertation, and whom I count as friends. I would also like to express particular gratitude to Santosh and Chris, who were generous with their time, patience, and guidance when I first joined the group. Thanks as well to Honggyu, who was always ready to teach me more crystallography, and Tyler and Salva, our crack team of TEM experts. It has been especially rewarding to work with Daniel, who managed to handle whatever gnarly problems were thrown his way, and, for less time, Paola. I'm also lucky to have learned a lot from the little I've worked with Omor and the less I worked with Biswajit, and to the rest of my groupmates I can say I look forward to working with you more at some point in the future.

My committee, for their support, assistance and responsiveness over the years

Kurt Olsson, for dealing with so many of my problems and being a great colleague (is that the right word for someone who teaches you loads and in his spare time fixes your broken dreams?), and John English, who, like an echidna, conceals his soft and playful nature behind a coat of spines

Amanda Strom—like several others here, she'd be just as much at home on a page entitled *Apologies*—who, in spite of having been responsible for a wider range of professional duties than fits here, has been a reliable collaborator and great help for however many years it's been since she helped Friedhelm and me with the MPMS2

Jennifer Farrar, without whom I would not have graduated

The numerous scientific and administrative staff and faculty of the MRL and MRSEC shared facilities, particularly Joanne McNie, Mark Cornish, Aidan Taylor, Youli Li, and Miguel Zepeda, without whom most of my work would have simply not come to pass

The pros at the Nanofab, which by their efforts must be the best-run institution in the world

My many officemates, particularly Doug, Emily, Emily and Erin, who, besides being great people to spend time with and excellent scientific role models, dealt with a truly unfair amount of watercraft, wet wetsuits, and weird occurrences

My old colleagues: father Friedhelm, who taught me so much that was sorely needed, and Ale, who taught me everything I know about zebrafish, Payam, Elijah, Adam, Hannah, Samhita, James (a friend in success and failure), Mishel, and others

My 2013 cohort—obviously Christina, who's been teaching me physics since about 2011, and also my apartmentmates Kevin and Blake, and all the rest of you—as well as my rough contemporaries Kevin, Andrew, and Matthew, for giving physics and -ists a good name and being great role models and colleagues (and friends, obviously), and, among housemates, Sam for resurrecting my hard drive and other *gesta*

Fellows in MBE, Nick, Kaveh, Yuewei, Akhil, Mario, Shouvik, Tony, Mihir, Kelsey, Richard, many more, thanks for all the help, inspiration, and collegiality

Jan Osimo and Andrea Stouffer, for getting me on my feet again

Jennifer Kloetzel, for teaching an old dog new tricks, Claire, Hannah, Ivan, Naomi, Chenoa, Britta, Joanna, Thomas, Katrina, Josh

Alex, Alex, Jess, Bretton, Jordan, Elliott, Vincent, Paul, Rachel, Mike, Christine, Julian, Mike, Richard, Jacquie, Rupert, Kira, Ian, Mary, Nakoa, Tibi, Gursharan, Jose, and all the others—thanks for all the passes and beers and more and sorry for letting all those headers go by

Kim, Aaron, and everyone else at Handlebar Coffee, for being the best reason to get up in the morning

And my family, Janet and Stephen, Marc, Robert, and Catherine, and Elizabeth, for their love and support, without which I couldn't do anything

Curriculum vitae

David A. Kealhofer

Education

2020	Ph.D., Physics, UC Santa Barbara
2017	M.A., Physics, UC Santa Barbara
2013	B.A., Physics and Classics, Williams College, Williamstown, MA
2009	San Francisco University High School, San Francisco, CA

Publications

15. **David A. Kealhofer,** Manik Goyal, Tyler N. Pardue, and Susanne Stemmer, "The three-dimensional topological insulator picture in (001)-oriented Cd_3As_2 films," in preparation

14. Manik Goyal, **David A. Kealhofer,** Timo Schumann, and Susanne Stemmer, "Magnetoresistance effects in cadmium arsenide thin films," *Applied Physics Letters* **117,** 170601 (2020) DOI

13. Omor F. Shoron, Manik Goyal, Binghao Guo, **David A. Kealhofer,** Timo Schumann, and Susanne Stemmer, "Prospects of terahertz transistors with the topological semimetal cadmium arsenide," *Advanced Electronic Materials* **2020**, 2000676 (2020) DOI

12. Manik Goyal, Salva Salmani-Rezaie, Tyler N. Pardue, Binghao Guo, **David A. Kealhofer**, and Susanne Stemmer, "Carrier mobilities of (001) cadmium arsenide films," *APL Materials* **8,** 051106 (2020) DOI

11. **David A. Kealhofer,** Luca Galletti, Timo Schumann, Alexey Suslov, and Susanne Stemmer, "Topological insulator state and collapse of the quantum Hall effect in a three-dimensional Dirac semimetal heterojunction," *Physical Review X* **10,** 011050 (2020) DOI

10. Omor F. Shoron, Timo Schumann, Manik Goyal, **David A. Kealhofer,** and Susanne Stemmer, "Field-effect transistors with the three-dimensional Dirac semimetal cadmium arsenide," *Applied Physics Letters* **115,** 062101 (2019) DOI

9. Luca Galletti, Timo Schumann, **David A. Kealhofer,** Manik Goyal, and Susanne Stemmer, "Absence of signatures of Weyl orbits in the thickness dependence of quantum transport in cadmium arsenide," *Physical Review B* **99,** 201401(R) (2019) [Editors' Suggestion] DOI

8. **David A. Kealhofer,** Honggyu Kim, Timo Schumann, Manik Goyal, Luca Galletti, and Susanne Stemmer, "Basal-plane growth of cadmium arsenide by molecular beam epitaxy," *Physical Review Materials* **3,** 031201(R) (2019) [Editors' Suggestion] DOI

7. Alessandro Mongera, Payam Rowghanian, Hannah J. Gustafson, Elijah Shelton, **David Kealhofer,** Emmet K. Carn, Friedhelm Serwane, Adam A. Lucio, James Giammona, and Otger Campàs, "A fluid-to-solid jamming transition underlies vertebrate body axis elongation," *Nature* **561,** 401 (2018) DOI

6. Luca Galletti, Timo Schumann, Omor F. Shoron, Manik Goyal, **David A. Kealhofer,** Honggyu Kim, Susanne Stemmer, "Two-Dimensional Dirac Fermions in Thin Films of Cd_3As_2," *Physical Review B* **97,** 115132 (2018) [Editors' Suggestion] DOI

5. Manik Goyal, Luca Galletti, Salva Salmani-Rezaie, Timo Schumann, **David A. Kealhofer,** and Susanne Stemmer, "Thickness dependence of the quantum Hall effect in films of the three-dimensional Dirac semimetal Cd_3As_2," *APL Materials* **6,** 026105 (2018) [Editor's Pick] DOI

4. Timo Schumann, Luca Galletti, **David A. Kealhofer,** Honggyu Kim, Manik Goyal, and Susanne Stemmer, "Observation of the quantum Hall effect in confined films of the three-dimensional Dirac semimetal Cd_3As_2," *Physical Review Letters* **120,** 016801 (2018) [Editors' Suggestion] DOI

3. Timo Schumann, Manik Goyal, **David A. Kealhofer,** and Susanne Stemmer, "Negative magnetoresistance due to conductivity fluctuations in films of the topological semimetal Cd_3As_2," *Physical Review B* **95,** 241113(R) (2017) DOI

2. Friedhelm Serwane, Alessandro Mongera, Payam Rowghanian, **David A. Kealhofer,** Adam A. Lucio, Zachary M. Hockenberry, and Otger Campàs, "*In vivo* quantification of spatially varying mechanical properties in developing tissues," *Nature Methods* **14,** 181 (2017) DOI

1. G Ranjit, **D. Kealhofer,** G. D. Vukasin, and P. K. Majumder, "Measurement of $7p_{1/2}$-state hyperfine structure and $7s_{1/2}$-$7p_{1/2}$ transition isotope shift in ^{203}Tl and ^{205}Tl," *Physical Review A* **89,** 012511 (2014) [Editors' Suggestion] DOI

Abstract

Topological surface states in cadmium arsenide (001) thin films

by

David A. Kealhofer

Thin films are fertile ground in the study of so-called topological materials. Electrostatic gating, quantum confinement, and strain engineering afford separate degrees of control over the bulk band structure and, as a result, the presence of topological surface states, and are candidate routes toward engineering more exotic, correlated topological states. Cadmium arsenide is a topological material that affords such possibilities. It is called a 3D Dirac semimetal because, in its electronic structure, the crossing of spin-degenerate bands generates low-energy physics resembling that of the Dirac equation, with a linear dispersion in three dimensions. These Dirac nodes, as the crossings are called, can be manipulated to realize different topological phases.

One part of this dissertation describes the growth of thin films of cadmium arsenide of a particular orientation, (001), previously unstudied, using molecular beam epitaxy. The films' morphology, crystal structure, and orientation are investigated with x-ray diffraction, x-ray reflectivity, and atomic force microscopy alongside results from scanning transmission electron microscopy. (001)-oriented films are of special interest because, in the (001) surface Brillouin zone, the 3D Dirac nodes project onto the same point, resulting in a surface state dif-

ferent from those on every other surface.

The rest of the dissertation focuses on studies of the quantum Hall effect in these films. Transport through the (001) surface states is observed and studied in the quantum Hall regime, where the sequences of filling factors, as a function of magnetic field and carrier density in the film, reveal the Dirac nature of these two-dimensional surface states. The picture of the cadmium arsenide (001) film, then, as a three-dimensional topological insulator, is developed with reference to its band structure, with help from an effective model of its surface states. This picture is tested experimentally via the evolution of the quantum Hall effect in films of diminishing thickness. This evolution is contrasted with the predictions of an alternative model for the film's electronic structure based on subband confinement. These growth and transport studies establish cadmium arsenide not just as a topological insulator of notably high carrier mobility, but as a versatile platform for studying and realizing topological phases in thin films.

Contents

Chapter 1

Cadmium arsenide as a topological material

1.1 Welcome

The study of cadmium arsenide (Cd_3As_2) dates to at least 1935 [82], its "unusually high electron mobility," as the letter title had it, remarked upon as early as 1959 [65]. The study of Cd_3As_2 continued more or less sporadically through the 60s and 80s, and the historian might identify long-running currents in the course of semiconductor physics and narrow-gap semiconductors particularly as the tides on which these works were borne. The contemporary study of Cd_3As_2 began in 2013, when it was identified by Wang *et al.* as a candidate material to realize an electronic state of recent interest, the three-dimensional Dirac semimetal (discussed in more detail in Section 1.2) [91], which itself had only entered the public imagination in 2012 [97]. This discovery followed that of the closely related Weyl semimetal, which, like the 3D Dirac semimetal, is in a sense a solid-state realization of relativistic physics. Though modern interest in this topic can probably be said to date to 2011 [88, 12], pioneering early work dates to the 80s [60, 85, 61] and by a coincidence its history might be traced to the 30s, like the study of Cd_3As_2 [31]. And since these two, the ranks of "topological semimetals," which includes such creatures as nodal line semimetals and triple point semimetals, seem to be expanding quickly.

The summer following the proposal of Cd_3As_2 as a 3D Dirac semimetal, four different research groups, all using angle-resolved photoemission spectroscopy (ARPES), a powerful technique for visualizing electronic band structures, published papers with essentially the same message: Cd_3As_2 is a 3D Dirac semimetal [9, 49, 59, 96]. Such intense research activity has to be the product of a certain kind of optimism, and, in retrospect, the optimistic view that Cd_3As_2 would prove to be of durable interest has been at least partially vindicated. One reason is that Cd_3As_2 remains unusual among 3D Dirac semimetals in that its Dirac nodes are not obscured by other bands. In a 3D Dirac semimetal, two bands cross—the crossing is called a Dirac node—and the low-energy physics resembles that of the Dirac equation, with a linear dispersion in three dimensions [97]. Signatures of these 3D Dirac nodes can be subtle, and a chorus of conventional electrons would tend to drown them out. Another reason for its continued relevance, no doubt, is its "unusually high electron mobility." While hope has waned for achieving the mobility reported in an early single-crystal sample, $9,000,000\ \mathrm{cm^2/V\,s}$ [46], it seems certain that numbers like $20,000\ \mathrm{cm^2/V\,s}$ [69] and $30,000\ \mathrm{cm^2/V\,s}$ [57], which have been reported in thin films, are not the highest possible. Besides enabling the observation of the quantum Hall effect [67, 81], this has supported research in the direction of specialty electronic devices [14, 72].

These two properties—the prominence of the Dirac nodes and the high electronic mobility—are at the core of our interest in Cd_3As_2, and indeed there is a thread in Cd_3As_2 research that follows high-mobility semiconductor-like physics. The other thread, which is latterly dominant, might be called topological in its goals, and this also motivates our efforts in thin films. From a certain perspective, "topological" has the main force of meaning "edge states." The unifying framework is the bulk-boundary correspondence, which relates the number of protected edge modes to features of the bulk band structure that are non-local (e.g., they may be localized in reciprocal space) [30]. (In that sense—the presence of the edge modes is robust to local changes to the bulk—are the edge modes protected.) Now, 3D Dirac semimetals are not (that) topological, if the expression can be forgiven, in the sense that the (Fermi arc) surface states are not topologically protected [36], but it might be seen as a platform for achieving other topological states by modification of its bulk band structure via accessible experimental controls, particularly in thin films [97, 89].

Thin films are an exciting frontier in the study of so-called topological materials. Electrostatic gating, quantum confinement, and strain engineering afford separate degrees of control over the bulk band structure and, as a result, the presence of topological surface states, and suggest routes toward engineering more exotic, correlated topological states. In addition, interest in 3D Dirac semimetals and 3D topological insulators (TIs), as components in a hybrid platform for realizing topological quantum computing, continues to grow [19, 94, 90]. Finally, unlike bulk single crystals, thin films can be integrated into conventional semiconductor device micro- and nano-fabrication processes, harnessing the capabilities of decades of engineering and optimization.

This dissertation focuses mainly on our early growth and electronic transport studies of Cd_3As_2 thin films grown at a particular crystal orientation, identified as (001), which, among all other surfaces of the crystal, has a surface state of a unique kind. The dissertation is organized as follows. The rest of this Chapter consists of a brief introduction of the 3D Dirac semimetal. Chapter 2 introduces the crystal and electronic structure of Cd_3As_2, the focus of our studies, along with a couple of different approaches to modeling the electronic states of (001) films. We were the first to publish a method for growing these (001)-oriented films, which we accomplish using a technique called molecular beam epitaxy (MBE). These first studies were published in ref. [38], and substantial parts of that work are reproduced in Chapter 3, which describes in broader terms the MBE growth of Cd_3As_2. The subsequent Chapter describes our studies of electronic transport, particularly the integer quantum Hall effect, in (001) films, and reproduces parts of two works, one published [37], and the other in preparation. Chapter 5 contains some brief suggestions for future research directions. In addition, there are two Appendices, one summarizing in-progress work on capacitance measurements in Cd_3As_2, with some comments on apparatus, the other, directed mainly towards prospective Cd_3As_2 MBE growers, containing some suggestions for best growth practices.

1.2 The three-dimensional Dirac semimetal and its cousins

The 3D Dirac semimetal is so called because its low-energy dispersion realizes the Dirac equation in three dimensions in the way that, say, graphene's realizes it two. This is generally caused by the crossing of two spin-degenerate bands; the 3D Dirac semimetal requires inversion and time-reversal symmetry (or their combination) [43]. Band crossings due to accidental degeneracy are well-known to be ordinarily avoided—that is, a gap opens instead—but they can be stabilized by symmetries of the crystal lattice [97]. There are a couple of different routes to achieving this [4]. The most relevant to our study is sometimes called the 3D Dirac semimetal from an accidental band crossing [23, 95] or, sounding somewhat more encouraging, from band inversion [4]—these refer to the same thing. In Cd_3As_2, it is thought that the band crossings are driven by an inversion of two bands that are prevented from hybridizing because they belong to different irreducible representations of the fourfold rotational symmetry of the k_z axis [91]. (A technical point, revisited in Section 2.1, is that the band inversion of the Cd 5s-like bands with respect to the As 4p-like bands [91, 56] is of separate consequence.) *These next two paragraphs mainly paraphrase ref. [88].*

A close cousin of the 3D Dirac semimetal is the Weyl semimetal [88, 12], whose low-energy excitations are described by the Weyl equation, which we mention for the sake of introducing its surface states and some contrasting topological properties. Weyl semimetals, just as 3D Dirac semimetals host Dirac nodes, host Weyl nodes, which are caused by the crossing of bands with no spin degeneracy. To eliminate the spin degeneracy, reducing the degeneracy of the node to twofold from fourfold, the Weyl semimetal must lack either time reversal or inversion symmetry [88]. In the vicinity of the Weyl node, the Hamiltonian can be written in terms of the momentum measured from the position of the node at $\vec{k}_0$ as $\vec{q} = \vec{k} - \vec{k}_0$ as (reproducing eq. 1 of ref. [88])

$$H(\vec{k}) = (E_0 + \vec{v}_0 \cdot \vec{q})\mathbb{1} + \vec{v}_1 \cdot \vec{q}\sigma_1 + \vec{v}_2 \cdot \vec{q}\sigma_2 + \vec{v}_3 \cdot \vec{q}\sigma_3,$$

where we have tediously written out the sum. (The Pauli matrices σ_i act on the pseudospin band degree of freedom.) A couple of observations are relevant. First, the wavefunctions will clearly have two components; in the 3D Dirac semimetal, an extra true spin degree of freedom means that the wavefunctions are four-component spinors. Second, the Hamiltonian uses up all three Pauli matrices (and the identity). A generic perturbation, which we can write in terms of the same matrices, thus shifts the position of the Weyl node but does not create a gap. In that sense the Weyl semimetal is robust to perturbation; this is not true of the 2D Dirac semimetal (e.g., graphene), whose low-energy Dirac-like Hamiltonian does not include σ_z, or the 3D Dirac semimetal, because there are more 4×4 matrices than appear in the Hamiltonian [97].

There is an invariant associated with each node, calculated as $\text{sign}(\vec{v}_1 \cdot \vec{v}_2 \times \vec{v}_3)$ [88]. This quantity is proportional to the Berry flux that pierces any surface surrounding a single node, and, as a result, the Chern number of the 2D band structure on the surface of such a surface is ± 1 [4, 88]. The nodes come in pairs of opposite “chirality” [$= \text{sign}(\vec{v}_1 \cdot \vec{v}_2 \times \vec{v}_3)$], which is necessary in order that a 2D surface enclosing the entire Brillouin zone has a Chern number of zero [88]. More interestingly, considering now 2D surfaces that enclose only one Weyl node, each of these surfaces, which can be construed as a Chern insulator with Chern number ±1, itself has a gapless edge state [88]. Since this is true of any surface enclosing a single Weyl node, but not of a surface enclosing both, there is a surface state, composed of the edge states of the 2D surfaces, that begins at one Weyl node and terminates at its opposite-chirality partner [88]. One feature of these "Fermi arc" surface states is that they only exist on surfaces for which the Weyl nodes project onto different points in the surface Brillouin zone [88, 12]. (Fermi arcs were observed, using ARPES, in the Weyl semimetal TaAs [92].)

In 3D Dirac semimetals, some have proposed that the surface states resemble these Fermi arcs [64, 4]. They may well do, but it is worth noting that 3D Dirac semimetals are "less topological" than their cousins the Weyl semimetals. The band structure of a 2D surface enclosing a Dirac node, unlike one enclosing a Weyl node, has Chern number zero, and there is no analogous topologically protected surface state [4]. It has been shown, for example, that the Fermi arcs can be disconnected from the projected Dirac nodes even by small changes to the chemical potential [36]. (In this work, we study (001) Cd_3As_2 films. In Cd_3As_2, as discussed in Section 2.1, the (001) surface is precisely the surface where the Dirac nodes project onto the same point in the surface Brillouin zone, and there are no Fermi arcs.) But, by a number of different perspectives, 3D Dirac semimetals are close to being topological, with suitable modifications of the bulk band structure, whether the breaking of inversion symmetry, which can create a Weyl semimetal [97], the breaking of rotational symmetries to create other semimetal or insulator phases [70, 5], or thin film confinement, which has been proposed as a way to realize a 2D topological insulator (quantum spin Hall insulator) [91]. This ability to tune the 3D Dirac semimetal into other phases is at the heart of our motivation to grow and study Cd_3As_2.

Chapter 2

Two models for 2D transport in (001) thin films

Some of the work in this chapter, notably Figure 2.2 and its caption, reproduces or adapts portions of our work published previously [37], to which APS retains the copyright.

In this chapter, we introduce the crystal and electronic structure of Cd_3As_2 in more detail. We then discuss a basic approach to understanding how this is modified in a (001) thin film, before introducing a different perspective, based on the surface states of a topological insulator. We contrast the results of these two pictures in more detail in Chapter 4, when we discuss our transport experiments.

2.1 Crystal and electronic structure

The crystal structure of Cd_3As_2 is complicated. The prevailing contemporary view is that the unit cell, which contains 160 atoms, consists of a nearly cubic subunit that is arranged in rotating fashion $2 \times 2 \times 4$ to form the large tetragonal unit cell shown in Figure 2.1a [3]. The subunit can be described in terms of the antifluorite structure, in which the anions (here, arsenic) sit on fcc sites, and the cations (cadmium atoms) form the corners of a cube on the tetrahedrally coordinated interstitial sites. For Cd_3As_2, however, the stoichiometric compound is formed only when two of the interstitial sites, which in the antifluorite structure form a cubic sublattice, are vacant. These vacant sites lie diagonally opposite on a single face of the subunit, which distorts the cubic subunit, as shown in Figure 2.1b. The full unit cell is assembled by stacking these subunits, each containing two formula units, on each other in clockwise and counterclockwise blocks, four subunits high. In the *ab*-plane, each $1 \times 1 \times 4$ block has neighbors of opposite handedness; the unit cell is thus $2 \times 2 \times 4$. The crystal structure has lattice parameters $a \approx 12.6$ Å and $c \approx 25.4$ Å.

The crystal structure is shown projected along [001] in Figure 2.1c. The essentially cubic fcc coordination of the arsenic atoms is visible, and the relatively distorted positions of the cadmium atoms can also be seen—they do not all project onto the same point. In single crystal growth, however, the easy cleavage plane is (112), shown in Figure 2.1d, not (001). Since the tetragonal structure is made up of the quasi-cubic subunit described above, $a = b \approx c/2$, and the tetragonal (112) plane can be construed as the (111) plane of the quasi-cubic subunit. (Note

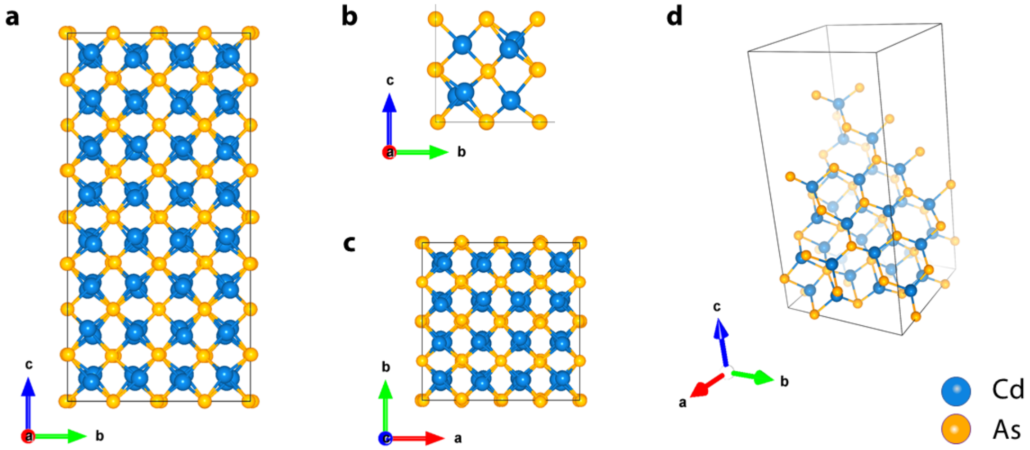

Figure 2.1: Crystal structure of Cd_3As_2 according to ref. [3], rendered in VESTA [54]. (a) Single unit cell projected along [100]. The c-axis, [001], is oriented upwards. Cadmium atoms are rendered in blue, and arsenic atoms are rendered in orange. (b) Distorted subunit ($(0, 0, 0)$ to $(a/2, b/2, c/4)$ is shown) in the same orientation as (a). The cadmium atoms (blue) sit in simple cubic coordination about the center of the subcell, except, since two sites are vacant, the atomic positions lie closer to the center of the cube face. (c) Single unit cell projected along [001], showing the square sublattice of arsenic atoms. The b-axis is oriented upwards. (d) The easy cleavage plane is (112). A slice of the unit cell along this plane is shown, with perspective added to more easily identify the hexagonal structures along this surface.

that the cleavage planes of fluorite are {111}.) As a result, all ARPES studies but one have been on (112) planes (and the claim of the odd study out, to have studied the (001) surface, is not supported beyond examination of the emission spectrum at the center of the Brillouin zone), and so has been the majority of the thin film growth.

Band structure calculations have found two Dirac nodes close to the Brillouin zone center, located on the k_z-axis at $k_z = \pm k_D$, where $k_D \approx 0.1c^*$ [91, 3, 56]. There is some minor quantitative disagreement about the relevant energy range for Dirac physics, but it seems to be on the order of 10–50 meV [91, 3, 56]. The Dirac nodes lie at the Fermi energy, and there are not any other bands nearby. Zooming out from the nodes, still near the center of the Brillouin zone, one sees that the two bands that make up the nodes originate as two As $4p$ bands; these lie a couple hundred meV above another such As $4p$ band, which lies another couple hundred meV above a Cd $5s$ band (see Figure 2.2) [91, 3, 56].

The complexity of the crystal structure is evident from the historical disagreement about it. While it has always been known to be a tetragonal unit cell containing many formula units, it has been variously claimed to belong to space group $P4_2/nmc$, with lattice parameters a = 8.95 Å, c = 12.65 Å [82]; $I4_1cd$, with lattice parameters a = 12.67 Å, c = 25.48 Å [74] or, elsewhere, a = 11.78 Å, c = 23.64 Å [62]; or, most recently, $I4_1/acd$, with lattice parameters a = 12.633 Å, c = 25.427 Å [3]. Of these, the strongest evidence is for the structure reported by Steigmann and Goodyear [74] and the one reported by Ali *et al.* [3]. The main stakes are the inversion center present in the $I4_1/acd$ space group, which is absent in the $I4_1cd$ structure. An admirable treatment of the disagreement is afforded by ref. [3]. There, they imply that, since the authors of ref. [74] did not consider the centrosymmetric structure, and besides their apparatus was less sensitive to quantitative intensity than that now commonplace, the weight of ref. [74] does not go against the $I4_1/acd$ stucture. For their part, Ali *et al.* put it that certain goodness-of-fit parameters are slightly improved for the centrosymmetric structure, which, combined with the general admonition that centrosymmetric structures be favored versus noncentrosymmetric ones in the refinement of crystal structures from diffraction data [52], favors the centrosymmetric $I4_1/acd$ structure sufficiently to support their finding. If one finds this line of argument perhaps too much an argument, one can be reassured that the actual differences in atomic positions being considered are very small (approximately 0.03 Å) [3]. The resulting differences in the calculated band structure (compare Fig. 5 of ref. [3] to Fig. 1b of ref. [91]) are perhaps merely quantitative, but, in the vicinity of the Dirac nodes, the differences are extremely subtle. (The most important difference, which has no bearing on this work, may be that the band lying 250 meV below the Fermi level is spin-split by about 10 meV between Γ and X in the noncentrosymmetric structure and, of course, is not spin-split in the centrosymmetric structure.) The crystal structure described at the beginning of this chapter is that of Ali *et al.* [3].

One highly reasonable perspective on the last paragraph is that Cd_3As_2 is *too* complicated. The complexity of the crystal structure—really, the size of the unit cell—is problematic for band structure calculations, and it stands in contrast to the simplicity of the 3D Dirac semimetal's definition, which from one perspective only demands that spin-degenerate bands cross. (For comparison's sake, consider graphene, where the crystal structure has a two-atom basis, and a tight-binding model suffices to uncover its Dirac nature.) So why is Cd_3As_2 a Dirac semimetal? The short answer is that the two As $4p$-like bands that cross at $k_z = \pm k_D$ cannot hybridize because they belong to different representations of the C_4 group, which is the symmetry along the $\Gamma - Z$ line [91]. Since the bands are spin-degenerate, as the centrosymmetric structure guarantees [3], the crossings are Dirac rather than Weyl nodes. Why one band lies higher than the other has a more detailed microscopic explanation, but it is thought to be related to the ordering of the vacant sites (relative to the antifluorite structure) on the Cd sublattice [56]. A structure without any ordering of the vacant sites lacks Dirac nodes [91].

The inversion of the Cd $5s$ and As $4p$ bands has separate causes. It is driven by a large spin-orbit interaction, like in HgTe. In fact, the band inversion appears to be a feature of the 10-atom pseudocubic subunit that makes up the more complicated structure [56]. On that basis one might speculate that the band inversion is a very robust aspect of the band structure—perhaps more so than the 3D Dirac nodes, at least with respect to disordering of the vacancies. If one assumes that the variance in the early crystal structures is at least partly a reflection of the range of crystal growth practices and concomitant diversity of sample quality than the inaccuracy of the XRD, the robustness of the band inversion to variations in sample quality might be seen in the comparatively early determination of the inverted band structure of Cd_3As_2 versus its discovery as a 3D Dirac semimetal. That is, in 1969 it was first suggested that Cd_3As_2 had an inverted band gap of nearly 400 meV [86], in decent agreement with recent band structure calculations [91, 3, 56], but it took the combined powers of band structure calculations, modern synthesis techniques, and ARPES to reveal the Dirac nodes in 2014.

2.1.1 Effective model

The complexity of the crystal structure and the comparative simplicity of the 3D Dirac semimetal encourage us to treat the band structure of Cd_3As_2 with an effective model. Let's define k_D so the Dirac nodes lie at $\vec{k} = (0, 0, \pm k_D)$. Most first-principles calculations have found $k_D < 0.05\ \text{Å}^{-1}$, [91, 3, 56] consistent with several experimental studies [33, 2, 28], though there are some discrepancies—for a recent review, see [15]. Since the length of the first Brillouin zone is 5 or $10 \times k_D$, we consider a $k \cdot p$ approach to modeling the bulk band structure near the Dirac nodes to be accurate, as has been done elsewhere [91, 13]. The original $k \cdot p$ model along these lines adapted one for a cubic semiconductor by introducing a parameter δ to split the p_z state from the p_x and p_y ones: thus was achieved the model for the tetragonal crystal structure of Cd_3As_2 [8].

This model (sometimes referred to, for obvious reasons, as the Bodnar model) is applied in a number of recent works [91, 13]. The full eight-band Hamiltonian, per ref. [91], is written

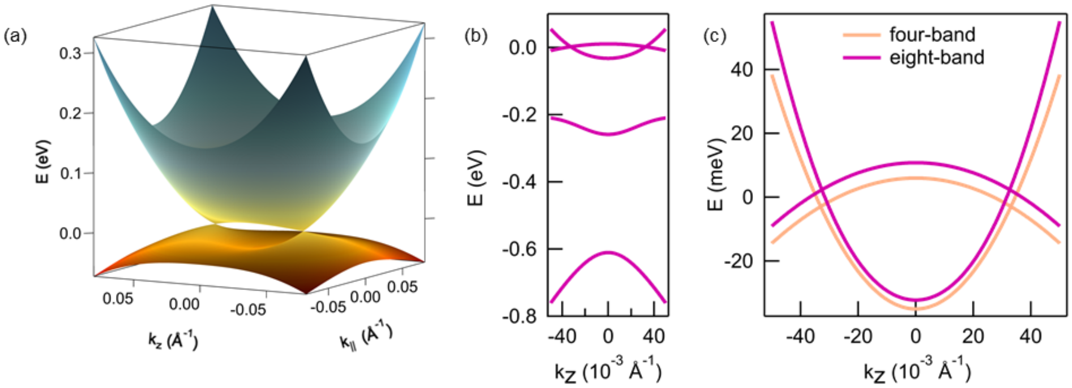

Figure 2.2: Band structure of Cd_3As_2 from $k \cdot p$ models. (a) Bulk band structure near the Dirac nodes. Two Dirac nodes lie along the $\Gamma - Z$ direction. In their vicinity the dispersion is linear in all three directions, here plotted as E vs. $k_{||} = \pm\sqrt{k_x^2 + k_y^2}$ and k_z. The parameters are those of ref. [13]. (b) Eight-band $k \cdot p$ model from ref. [91]. The energies are plotted as a function of k_z. Each of the four visible bands is doubly degenerate. (c) Four-band model with parameters from ref. [13] compared to the eight-band model from ref. [91], demonstrating that, while the difference in the model parameters produces a small energy offset, we can refer to either one equivalently.

in terms of the following four-band Hamiltonian,

$$H_4(\vec{k}) = \begin{pmatrix} E_s + A'\vec{k}^2 & ik_xP & ik_yP & ik_zP \\ -ik_xP & E_p + M\vec{k}^2 - L'k_x^2 & Nk_xk_y & Nk_xk_z \\ -ik_yP & Nk_xk_y & E_p + M\vec{k}^2 - L'k_y^2 & Nk_yk_z \\ -ik_zP & Nk_xk_z & Nk_yk_z & E_p - \delta + M\vec{k}^2 - L'k_z^2 \end{pmatrix}, \tag{2.1}$$

as

$$H_8(\vec{k}) = \begin{pmatrix} H_4(\vec{k}) & 0 \\ 0 & H_4(\vec{k}) \end{pmatrix} + H_{\mathrm{SO}}, \tag{2.2}$$

with

$$H_{\mathrm{SO}} = \frac{\Delta}{2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & 1 \\ 0 & i & 0 & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i & 0 & -i & 0 & 0 \\ 0 & 1 & i & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.3}$$

where we have only written terms to $O(k^2)$, and the parameters $E_s, E_p, L, M, N, P, \delta$ and Δ are to be fit to match band structure calculations or experimental data, as was done in ref. [91]. In discussing the bands that form the nodes, it is sometimes expedient to retain only these, and we use the parameters of ref. [13]. The resulting dispersion relations are plotted in Fig. 2.2.

2.2 (001) thin films

The main motivation to study (001) thin films is the surface state. Wang *et al.* have proposed that very thin films (a few nm thick) are a way to realize a 2D TI (aka a quantum spin Hall insulator) [91]. The premise is that, at certain thicknesses, the subband quantization of the bulk bands can produce an inverted gap in the 2D spectrum, as has been predicted in the topological insulators Bi_2Te_3 and Bi_2Se_3 [48]. In these next sections, we use a simplified model of the (001) quantum well to set some expectations for transport results in thin films (ranging from 12 nm to 50 nm) discussed in Chapter 4. This model differs slightly from that of ref. [91], but nevertheless makes similar predictions. We also expose an old phenomenological model for the surface of a 3D TI, which, at the risk of spoiling the surprise, matches the transport results better.

2.2.1 Thin film confinement

The remaining sections of this Chapter reproduce and adapt portions of the unpublished manuscript provisionally titled, "The three-dimensional topological insulator picture in (001)-oriented Cd_3As_2 films," which contains work performed in collaboration with my coauthors, Manik Goyal, Tyler Pardue, and Susanne Stemmer.

It is instructive to examine the picture that emerges from considering only the bands that form the 3D Dirac nodes. A naïve but effective way to model the thin film confinement is to treat an infinitely deep well, that is, quantize $k_z = n\pi/L$, with $n = 1, 2, 3\ldots$ and L the thickness of the film. In doing this, we have explicitly discarded surface states from our analysis. Perhaps more troubling is that the confining potential $V(z)$ does not break the fourfold symmetry of the k_z axis, but the bulk Dirac nodes are nevertheless destroyed. (One reason for this is that the $k \cdot p$ Hamiltonian is really only an effective treatment of a small portion of the Brillouin zone; it does not have the same microscopic origins as the true band structure.) We will see, however, that these objections do not invalidate our conclusions.

In-plane spectra, $E(k_x = k_y)$, are plotted in Figure 2.3 across a range of thickness that includes the films studied here. (The numerical values of the $k \cdot p$ coefficients are taken from ref. [13].) Panels (a) through (f) show thicknesses from 6 nm (a) to 30 nm (f). It can be seen that the gap shrinks non-monotonically as thickness is increased: panel (g) shows the

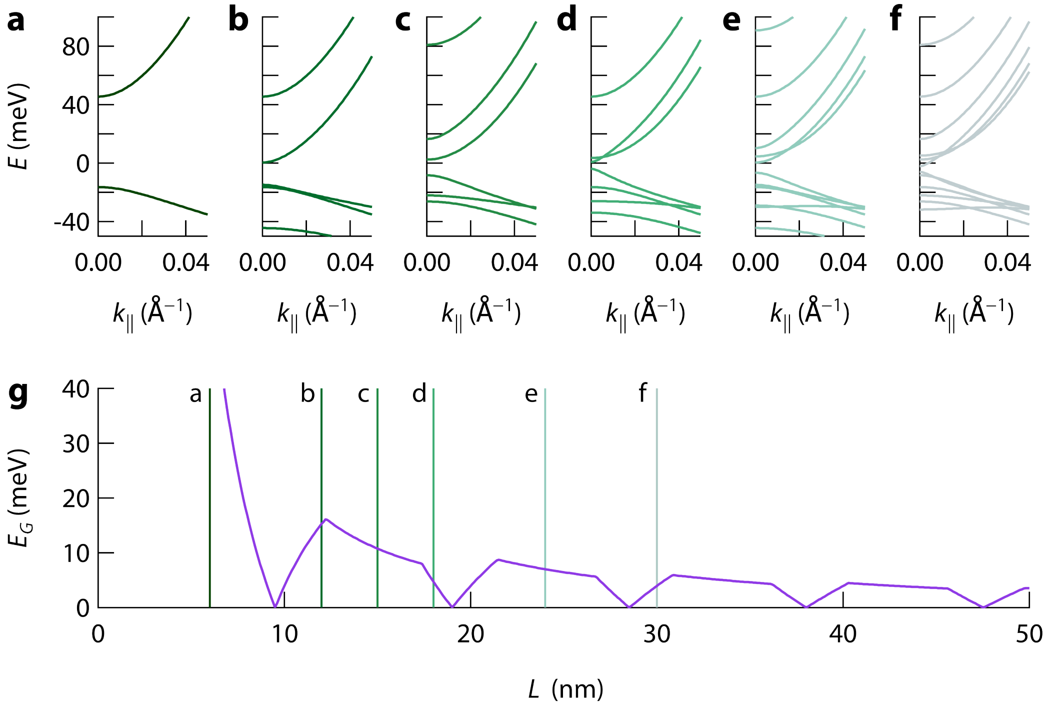

Figure 2.3: Thickness dependence in the subband picture. (The $k \cdot p$ model and coefficients are those from ref. [13].) In-plane spectra ($k_{\|} = k_x = k_y$) are plotted for a film thickness L equal to (a) 6 nm, (b) 12 nm, (c) 15 nm, (d) 18 nm, (e) 24 nm, and (f) 30 nm. (g) Evolution of the gap E_G as a function of L. The thicknesses corresponding to panels (a)-(f) are marked with labeled, colored vertical lines.

evolution of the gap at $k = 0$ as a function of thickness. Across the range of thickness studied in our experiments, the gap should decrease from a maximum of about 20 meV to as low as 5 or 10 meV; if some uncertainty is allowed in the correspondence between the model and reality, we should expect that the gap can take on arbitrarily small values near certain critical thicknesses. In any event, the prediction of the subband picture is that the thickness is a key parameter in determining the size of the gap. (As remarked elsewhere [37], the agreement between this heuristic approach and more sophisticated ones [91] is nearly quantitative and verifies that the analysis is not ill-founded.) More qualitatively, as L increases, so does the number of subbands in any particular low-energy window. Comparing panels (b) through (e), the number of conduction bands relevant to the transport increases from one to two or three (depending on E_F). As a result, the model predicts a commensurate increase in the complexity and/or apparent degeneracy of the Landau level spectrum.

A more enlightened model that reflects the robustness of the Dirac nodes to the confining potential and does not exclude out of hand the surface states in the quantization of k_z has the same difficulties. Even if the Dirac nodes are preserved, the increasing number of relevant subbands would still cause an increase in the degeneracy factor or the complexity of the sequence of filling factors. And by the same logic as in the previous paragraph, the possibility of observing a surface state lying within the subband gap (see Fig. 2.3g), is a strong function of the thickness. In other words, the simplified model makes fewer constraints on the data than the more realistic one.

2.2.2 Topological insulator surface states in (001) films

The motivation for the TI picture in the film is the inversion at Γ between the Cd 5s band that lies a couple hundred meV below the three As 4p bands, two of which form the bulk Dirac node [91]. If the energy scale for the resulting crossing is simply the size of the energy difference between the inverted bands, the resulting Dirac state might be observable—modulo interference from the bulk states—over more than 100 meV. That scenario, which is what we refer to as the 3D TI picture, is attractive for its consistency with the transport, which is discussed in Chapter 4. It is also similar to a model that has been proposed recently in compressively strained HgTe films, which have a similar band structure to Cd_3As_2 [51]. That work, unlike ours [37], also identifies a number of massive Dirac states that can be bound to such an interface. The presence of such massive states, named after Volkov and Pankratov [85, 61], depends on the microscopics of the interface, e.g. how abrupt it is. The massless Volkov-Pankratov state, however, which is identified with the 2D Dirac surface state in a 3D TI, is always present at the interface between materials of inverted and normal band orderings [77]. In the following sections, we quickly introduce an old phenomenological model for the surface states of a TI and briefly discuss how some of the model parameters might be seen in an experiment. (In Chapter 4, we discuss our experiments in the context of this model and the subband picture described in the previous section.)

2.2.3 Basic model and spectrum

We apply a simple continuum model for the surface states of a 3D TI to parametrize the relevant underlying physics [11]. Near the center of the Brillouin zone, we expect that the surface states can be described by the following Hamiltonian,

$$H_0 = \left[\hbar v_F (k_x \sigma_y - k_y \sigma_x) + \frac{\Delta_i}{2}\mathbb{1}\right] \otimes \tau_z + \frac{\Delta_h}{2}\mathbb{1} \otimes \tau_x, \tag{2.4}$$

where the σ_i and τ_i are Pauli matrices, referring to a spin degree of freedom and a surface pseudospin degree of freedom, respectively. The k_i are the 2D crystal momenta, and v_F parametrizes the steepness of the (identical) Dirac cones when $\Delta_h = 0$. The term Δ_i is formally an inversion-breaking term in the sense that τ_z is an inversion symmetry operator. In a more concrete sense, it describes the effect of the inter-surface energy difference that we ascribe to the asymmetry of the Cd_3As_2 heterostructure, that is, the difference in band offsets on either side of the film, as depicted schematically in Fig. 2.4. Finally, Δ_h allows a gap to open via the hybridization of the two surfaces. This is expected to be relevant only in films thin enough for there to be significant wavefunction overlap between the surface states.

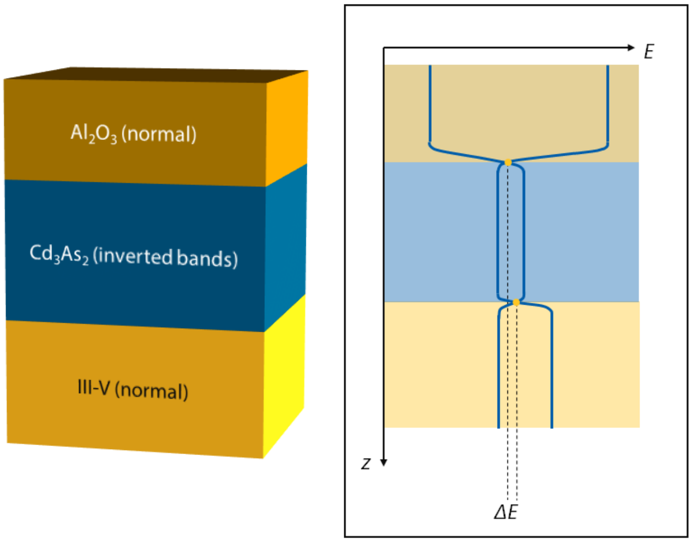

Figure 2.4: Schematic of the band crossings, driven by the inverted band structure of Cd_3As_2, at the interfaces between the bottom and top layers in our normal heterostructure (see Chapter 3). The difference in energy of the surface Dirac nodes (labeled as ΔE) is driven by a difference in band offsets between the top and bottom layers.

In the absence of a magnetic field, the spectrum of this Hamiltonian is

$$E_{0,\alpha,\beta}(k) = \alpha\sqrt{(\hbar v_F k)^2 + (\Delta/2)^2 + \beta\, \hbar v_F |k| \Delta_i}, \tag{2.5}$$

where $\alpha, \beta = \pm 1$ independently and we define $\Delta = \sqrt{\Delta_i^2 + \Delta_h^2}$. Evidently, in the circumstance where $\Delta_i = \Delta_h = 0$, the result is a doubly degenerate Dirac cone.

In a magnetic field $\vec{B} = B\,\hat{z}$, the Zeeman effect acts on the real spin degree of freedom and is captured by the addition of the following term to the Hamiltonian,

$$H_Z = g^* \mu_B B\, \sigma_z \otimes \mathbb{1}, \tag{2.6}$$

where g^* is an effective g factor, and μ_B is the Bohr magneton. The much more important effect, though, is the quantization of the spectrum into Landau levels. The calculation of this is accomplished by the Peierls substitution, $\hbar\vec{k} \longrightarrow \vec{\Pi} = \hbar\vec{k} + e\vec{A}$, where we use the gauge $\vec{A} = xB\,\hat{y}$, and e is the magnitude of the electron charge. Because $\vec{k}$ and $\vec{A}$ do not commute, we introduce the ladder operators $a = (2eB\hbar)^{-1/2}(\Pi_y + i\,\Pi_x)$ and $a^\dagger = (2eB\hbar)^{-1/2}(\Pi_y - i\,\Pi_x)$. The resulting Hamiltonian $H(B)$ reads,

$$H = \begin{pmatrix} \frac{\Delta_i}{2} + g^*\mu_B B & -\sqrt{b(B)}\,a & \frac{\Delta_h}{2} & 0 \\ -\sqrt{b(B)}\,a^\dagger & \frac{\Delta_i}{2} - g^*\mu_B B & 0 & \frac{\Delta_h}{2} \\ \frac{\Delta_h}{2} & 0 & -\frac{\Delta_i}{2} + g^*\mu_B B & +\sqrt{b(B)}\,a \\ 0 & \frac{\Delta_h}{2} & +\sqrt{b(B)}\,a^\dagger & -\frac{\Delta_i}{2} - g^*\mu_B B \end{pmatrix}, \tag{2.7}$$

where $b(B) := 2eB\hbar v_F^2$. The ladder operators are associated to states $|n\rangle$, where n is an integer ≥ 0, for which $a^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle$, $a\,|n > 0\rangle = \sqrt{n}\,|n-1\rangle$, and $a\,|0\rangle = 0$.

As long as $n > 0$, the eigenvectors $\Psi_n(B)$ of $H(B)$ have the form

$$\Psi_{n>0}(B) = \begin{pmatrix} \Phi_{1,n} \left|n-1\right\rangle \\ \Phi_{2,n} \left|n\right\rangle \\ \Phi_{3,n} \left|n-1\right\rangle \\ \Phi_{4,n} \left|n\right\rangle \end{pmatrix}, \tag{2.8}$$

where the $\Phi_{i,n}$ are numbers. When $n = 0$, however,

$$\Psi_0(B) = \begin{pmatrix} 0 \\ \Phi_{2,0} \left|0\right\rangle \\ 0 \\ \Phi_{4,0} \left|0\right\rangle \end{pmatrix}. \tag{2.9}$$

Combining all these, we find the spectrum in a perpendicular field to be

$$E_{\alpha,\beta,\,n>0}(B) = \alpha\sqrt{(g^*\mu_B B)^2 + (\Delta/2)^2 + n\,b(B) + \beta\sqrt{n\,b(B)\Delta_i^2 + (g^*\mu_B B)^2\Delta^2}}, \tag{2.10}$$

where, as above, $\Delta := \sqrt{\Delta_i^2 + \Delta_h^2}$, and $\alpha, \beta = \pm 1$ independently. For $n = 0$,

$$E_{\beta,\,n=0} = -g^*\mu_B B + \beta\Delta. \tag{2.11}$$

The effect of tuning the model parameters Δ_h and Δ_i is illustrated in Figure 2.5. If both

$\Delta_h = 0$ and $\Delta_i = 0$, then the two Dirac cones are degenerate everywhere, as shown in Fig. 2.5a. A finite Δ_i has the effect of shifting each cone relative to the other in energy, as seen in Fig. 2.5b, so that the energy difference between the Dirac points is equal to Δ_i. By contrast, the effect of Δ_h, shown in Fig. 2.5c, is to open a gap at the Dirac point; no degeneracy is split. Far from $k = 0$, the dispersion looks like that of Fig. 2.5a. If both Δ_i and Δ_h are nonzero, the case of Fig. 2.5d, then the dispersion relation far from Γ looks like that of Fig. 2.5b, while a gap opens at $k = 0$.

Since Δ_i essentially tunes a splitting while Δ_h opens a gap, small changes in Δ_i substantially affect the Landau level spectrum, in contrast to even fairly large changes in Δ_h. This difference is illustrated in Fig. 2.5e-f. In Fig. 2.5e, Δ_i is fixed at 75 meV, and the spectra of Equation 2.7 are plotted for different values of Δ_h, ranging from 0 to 60 meV. The effect of this is subtle and most noticeable for the lowest Landau levels at low field, or, in the quantum Hall regime, the smallest filling factors. In a spectroscopic experiment, sensitive to quantitative shifts in the Landau level energies, perhaps ARPES (provided good enough linewidth), could in principle be sensitive to the shift of the lowest couple of Landau levels. Such energy shifts are, however, invisible to a transport experiment. Equivalently small changes in Δ_i, by contrast, drastically affect the Landau level spectrum, and so in the event that an experimentally relevant parameter should control Δ_i, one would expect to see large, qualitative changes to the Landau level spectrum. This point is illustrated in Fig. 2.5f. Here Δ_h, now, is set at a fixed value and spectra are plotted for various Δ_i, ranging from 0 to 60 meV. These small changes result in both splitting and shifting of the Landau levels, causing crossings and changes in the sequence of filling factors that would be measured in a Hall measurement.

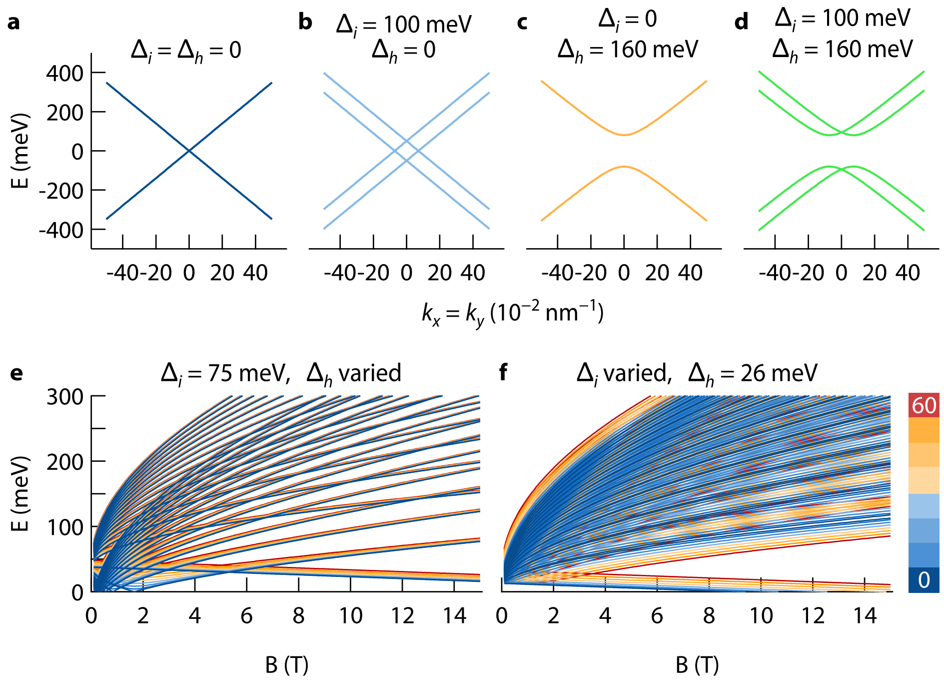

Figure 2.5: Model Hamiltonian spectra (see equations 2.4 through 2.11). (a)-(d) Zero-field spectra for small k and $v_F = 8 \times 10^5$ m/s. The parameter values are chosen for clarity. In panel (a), the cones are doubly degenerate everywhere because $\Delta_i = \Delta_h = 0$. (b) The effect of finite Δ_i is to displace the two cones in energy. The upper and lower cones are associated one to the eigenvalues of the τ_z operator. (c) By contrast, a finite Δ_h opens a gap in the spectrum, minimal at $k = 0$, which affects both surfaces in equal measure; the two bands are doubly degenerate everywhere. (d) When both Δ_h and Δ_i are appreciable, a gap opens and the two bands are in general nondegenerate. Whether away from $k = 0$ or $E = 0$, however, the spectrum is qualitatively and quantitatively similar to that in panel (b). (e) and (f) Landau level spectra as a function of Δ_h and Δ_i. Throughout g^* is set at +25. (The positive sign means that the zeroth Landau level disperses lower in energy with increasing field.) The changing parameter is colored according to the scale at the far right. (e) Effect of changing Δ_h with finite Δ_i. The Landau levels are shifted in energy, which is more noticeable for Landau levels with low indices. (f) Effect of changing Δ_i with finite Δ_h. Small changes to Δ_i cause large changes to the spectrum because the two fans [more visible in panel (e)] are pushed to higher and lower energies, respectively.

Chapter 3

Molecular beam epitaxy of (001) films

This Chapter contains a brief introduction to molecular beam epitaxy (MBE) of Cd_3As_2 generally, before switching focus to (001) films in particular. Ultimately, the lack of a suitable substrate has motivated the development of III–V semiconductor (III = Ga, Al, In; V = As, Sb) "virtual substrates" or buffer layers, which present additional engineering opportunities. Structural characterization, mainly via x-ray diffraction (XRD), is discussed, alongside a preview of electronic transport that is the topic of Chapter 4.

3.1 MBE checklist

An MBE growth system mainly consists of an ultra-high vacuum system, effusion cells to supply source material, and a heater for a substrate. As a concept, it is simple. One need only provide the right atoms to the right substrate to grow a material, provide enough thermal energy for them to defeat various activation hurdles, and do it all in vacuum to prevent impurities from stealing the show. Difficulty lurks around every corner, however. The right substrate may not be commercially available, or even exist. The right atoms must be delivered in the right quantity, which can be hard to measure, and is typically a function of the substrate temperature, which is itself hard to measure: accurate, fast, and cheap thermometry operable over large temperature ranges and on diverse materials is, empirically, nonexistent. Nature must ordinarily supply a growth window—a range of fluxes and temperatures at which stoichiometric growth is favored. And the demands of the ultrahigh vacuum environment of MBE add a layer of complexity to every task, from substrate preparation to maintenance.

3.2 Substrate selection for (001) plane growth

For Cd_3As_2, the right substrate does not exist: there are no commercially available substrates that have the right crystal structure or lattice parameter. (Substrates could, however, be fashioned from large single crystal growth of Cd_3As_2, which is not that outlandish a possibility.) As discussed in Section 2.1, the a lattice parameter of Cd_3As_2 is approximately 12.6 Å. This is very large (the lattice parameter of silicon is a = 5.4307 Å), and a reflection of the fact that the unit cell consists of $2 \times 2 \times 4$ rotated copies of a nearly cubic building block, which has the rather more modest in-plane lattice constant of 6.3 Å. This is still large—14% larger than silicon's—but falls in between that of GaSb (6.096 Å) or AlSb (6.1355 Å) and InSb (6.479 Å). High quality GaSb substrates are commercially available.

Our group's work on (112) Cd_3As_2 had demonstrated the possibility of growing Cd_3As_2 on (111)B GaSb [69]. This was accomplished by growing a buffer layer of GaSb on (111)B GaAs substrates, which, compared to (111)B GaSb substrates, have relative advantages in cost, mechanical properties, and electrical isolation: there are no semi-insulating GaSb substrates. The disadvantage of growing on GaAs, however, is the very large lattice mismatch (more than 10%) with Cd_3As_2, which must be mitigated by an intermediate buffer layer of GaSb, AlSb, InSb, or their alloys. While the buffer layer thereby provides an opportunity for tuning electronic properties of the full heterostructure, its growth is a research question (or, more optimistically, an optimization problem) in its own right: the growth of 6.3 Å III–Sb alloys is, among the well-traveled paths of III–V MBE growth, forbidding terrain. (Further difficulties lie in micro- or nanofabrication.)

One of the essential difficulties is that the buffer layer must accommodate the lattice mismatch with the substrate. In general this occurs via the presence of extended defects, such as misfit dislocations, in the buffer layer. Under some circumstances, which can be complicated to predict, such defects do not propagate through the buffer layer, or can be engineered to diminish in number as a function of the buffer layer's thickness. In spite of the approximately 7% mismatch of lattice parameter, the growth of GaSb on a GaAs (111)B substrate proceeds in atomically smooth fashion, though the growth window is smaller than for (001) growth [29]. Following those results, most of our group's work on (112) Cd_3As_2 films has been on relaxed GaSb layers grown on GaAs (111)B [69]. Since the (001) project began after the (112) project, however, a growing awareness of the extra difficulties posed by growing two latticemismatched layers combined with a burgeoning interest in more complicated heterostructures and devices led us to pursue the more straightforward approach of growing Cd_3As_2 on (001) GaSb substrates.

It is worth mentioning that another difficulty that the buffer layer adds, rather than subtracts, is related to its greater symmetry than the Cd_3As_2 epilayer. If certain (translational or rotational) symmetry elements are present in the buffer layer, but not in the epilayer, defects such as twins and antiphase boundaries can be expected to result. (Perhaps the best-known realization of this is in the growth of III–V growth on Si, where the epilayer has a polar structure absent in the layer beneath it [44].) This can be mitigated in different ways, but, for example, rotational twins in the (112) epitaxy of Cd_3As_2on (111) zincblende surfaces is suppressed on vicinally cut substrates [26, 40].

3.3 A few details and challenges for prospective growers

The growth of Cd_3As_2 by MBE in general has some pecularities in comparison to more well-known processes. We include an exposition of some of the problems of its growth along with some of the mitigation measures or solutions we have effected.

3.3.1 Sources

Growth was performed on a Veeco Gen II MBE system with conventional III–V sources: dual zone Ga, Al, and In effusion cells, a valved cracking As source, and a three-zone Sb cracker.

Cadmium arsenide flux was supplied from two low-temperature sources, each loaded with 6N Cd_3As_2 chunks (supplied by American Elements).

Early work in our group on (112) growth taught us that the best growth was achieved by supplying about 2×10^{-6} torr "beam equivalent pressure," which is what we call the reading on an ion gauge in the beam path, aka the beam flux monitor or BFM [69]. Running a single Cd_3As_2 source at the temperature required to achieve this flux rapidly depletes the source material, but two sources, each at half the desired flux, last slightly more than twice as long, due to the exponential dependence of the flux on temperature and the unavoidable rampup and idle times during normal growth and flux calibration. Even with this measure, reloading Cd_3As_2 source material remained the most frequent reason for system maintenance. To reduce the downtime involved, we replaced the conventional low-temperature sources with extra-long sources mounted inside a bellows and on carriages that ride a rail, driven by a lead screw, with a gate valve and pump line attached, in order to be able to separately vent and pump the Cd_3As_2 sources without venting the chamber. These retractable sources can be loaded and, in our system, a chamber pressure of mid-10^{-11} torr achieved after a night of pumping.

3.3.2 Thermometry and sample mounting

The growth of Cd_3As_2 proceeds at very low temperature, less than 200 °C, and the window for good growth is small, probably around 30 °C. The design of most MBE hardware, by contrast, is optimized for much higher temperature growth. For example, the oxide removal temperature of GaAs is about 590 °C, which may require heating the substrate manipulator to temperatures in excess of 700 °C; it is not unusual for most of the coupling between substrate heater and substrate to be radiative. Of all our evidence for the sensitivity of the cadmium arsenide film quality to the growth temperature, the strongest and most quantitative comes from growing (112) films, where the nucleation time provides an extra measurement of the true substrate temperature. More qualitatively, we had earlier noticed that films grown using solder-free holders typically showed defects in increased density near where the faceplate pins contact the substrate, indicating that small temperature differences, which we might estimate to be 10 °C, are relevant to the growth. Combined with flawed but nevertheless systematic studies of the effect of substrate temperature on growth, these observations lead us to claim that inconsistency in the growth is mainly due to the difficulty of measuring and controlling the substrate temperature at these low temperatures.

We measure the temperature of the substrate by two means, a thermocouple mounted in the substrate heater and an infrared (2.3 μm) pyrometer. In the conventional terminology we say that the thermocouple is part of an inner control loop that the pyrometer controls in an outer loop, which means that the control of the substrate heater power supply is determined by the thermocouple reading relative to the heater's setpoint, but the heater's setpoint is controlled by the pyrometer reading. For temperature homogeneity, the sample is mounted to a solid metal (tungsten) faceplate with liquid gallium (thanks to Tony McFadden and Mihir Pendharkar for the suggestion).

There are three main weaknesses of this control scheme. First, the two-loop configuration is necessary because the pyrometer window cannot have continuous access to the substrate without becoming coated. Ordinarily this challenge is overcome by using a so-called hot window, where an integrated or external heater mitigates the deposition on the window. In our case, the window temperature would need to exceed the cadmium arsenide growth temperature, resulting in a background signal (even if out of focus) that would overwhelm the substrate signal. So, lacking a hot window, we provide only intermittent optical access (with a shutter), and bake out the window occasionally to remove deposited material. The intermittent access necessitates the two-loop control, realized in software, which is very slow and may be fundamentally flawed by the weak coupling between the pyrometer control signal and the heater output power.

Pyrometry is inherently beholden to the emissivity of the blackbody under measurement, and this is the second major weakness of our control scheme. The emissivity is a phenomenological parameter whose value depends on many physical parameters of the blackbody, such as its dielectric properties and the surface finish, and so is fundamentally difficult to control. For example, consider the situation where the substrate is transparent to the pyrometer wavelength, but over the course of the growth becomes opaque or roughens, changing the emissivity. If the substrate is transparent throughout the whole growth and the surface roughness is identical throughout, run-to-run reproducibility will still be affected by the surface finish of the tungsten faceplate (which is scraped and chemically etched in between growths) and the shape of the gallium droplet adhering the substrate to it. Nevertheless, to the extent that even these essentially uncontrolled aspects of the film growth are roughly consistent as a result of standardized operating procedures, the growth can still be reproducible, even if the temperature readings are inaccurate. For example, highly reproducible oxide desorption and buffer layer growth indicate that the surface finish of the faceplate and the gallium droplet's disposition are minor or negligible concerns at high temperature. (In this case, though, the measured temperatures are inaccurate because of an emissivity correction that likely varies according to each layer.) Developing new heterostructures and working on new substrates, however, must rely mainly on a combination of educated guesswork and trial and error.

Finally, and perhaps most importantly, low-temperature pyrometry suffers from ubiquitous IR background. In our MBE system, the main sources of IR background are the sources and their filaments (the idle temperature for the Sb cracking zone is 1000 °C; the entire Al crucible idles at 800 °C; the As cracking zone is held at 850 °C; all sources send power to their filaments to maintain idle temperature) and the filaments of the ion gauges and RHEED gun. In practice, a combination of the unknown emissivity of the substrate and buffer layers and the background IR lead to a highly variable minimum temperature reading in the chamber, ranging from about 100 °C to 140 °C, which is unacceptable if we recall that the growth temperature needs to be controlled to better than 10 °C. In the long run, as a result, spectroscopic methods of thermometry are the most promising because, in the noisy environment of the growth chamber, the ability to modulate the input beam (say, a lamp) will be crucial to a measurable signal.

3.3.3 RHEED

Arguably MBE's greatest strength is that the ultrahigh vacuum environment is conducive to powerful studies of the growing surface. In contemporary MBE, reflection high-energy electron diffraction (RHEED) is often the only real-time feedback outside of thermometry that the grower has—certainly this is the case on our Veeco Gen II—and so the RHEED system conventionally holds the honor of being considered one of the few mission-critical pieces of the whole operation.

The information we get from RHEED is useful for several tasks:

1. Oxide desorption. As discussed below in Section 3.3.4, it is common practice to remove the wafer oxide thermally, at least for (001) GaSb. It is, however, possible to ruin a newly loaded substrate by overheating; the window between good desorption and irreversible damage appears to be around 30 °C. It is therefore helpful to witness the moment when the oxide desorbs. It also provides a calibration point for the substrate thermometer, from which later growth can be indexed.

2. Cadmium arsenide nucleation time. When grown on (111)B GaSb, the Cd_3As_2 growth rate is not constant in time until after a nucleation period, which we can identify in RHEED. Upon opening the shutters, no change in the diffraction intensity is observed for a certain amount of time, which is strongly temperature-dependent. After that time, the intensity drops sharply, corresponding to nanoscopic roughening of the surface, before it increases as the average island size increases. Without noting the nucleation

time, it is impossible to grow films of reproducible thickness.

3. Group III and V flux calibration. This procedure is described in Section B.3.

4. Assuaging thermometry paranoia. One benefit of our reliance on III–V materials for our substrates and buffer layers comes from how well known the system is. While III–Sb growth is much less studied than III–As growth, a number of useful points remain.

    (a) The oxide desorption provides a reliable temperature calibration point. Ordinarily, good GaSb homoepitaxy can proceed at about 10 – 15 °C lower than the oxide desorption temperature.

    (b) We distinguish "static reconstructions," when there is no growth, from "dynamic" reconstructions, when there is. In both cases, given a GaSb surface under Sb-rich conditions, there is a temperature- and flux-dependent transition between a high-temperature 1×3 reconstruction and a low-temperature 2×5 reconstruction [10]. The transition point can be identified to within about 10 °C. If one has a good flux calibration, this can be used as a calibration point for the substrate temperature, or vice versa. Similar statements hold for AlSb, InAs, GaAs, etc.

3.3.4 Oxide removal

An early bottleneck for the (111)B growth was the substrate preparation. Ordinarily, in (001) growth, the "epi-ready" oxide delivered by the wafer manufacturer can be removed thermally (under group V flux), which can be confirmed in RHEED. This roughens the surface:

nanoscopic pits are formed when pieces of oxide desorb. After a small amount of homoepitaxy, however, a smooth surface can be recovered. By contrast, on the (111)B surface, the pits formed during the oxide desorption nucleate large defects that are not diminished, but rather enlarged, by further growth. We developed a two-step oxide etch to solve this problem. The substrate is etched in diluted HCl (37% aqueous HCl 1:1 vol/vol with water), for 45 seconds. The etch is quenched in methanol, and the substrate is rinsed in isopropanol. Once loaded into the buffer chamber, it is subjected to an etch by atomic (thermally cracked) hydrogen. The principle of the hydrogen etch is that the supplied atomic hydrogen reduces the Ga and As oxides, which occurs preferentially at different temperatures, though there are incidental benefits related to its removal of carbon-containing compounds [76]. We have found for (111)B GaAs substrates that the process is not very sensitive to the dose (chamber pressure × time), but it is sensitive to the temperature: there is a window of about ±25 °C for good etching. After loading the substrate into the growth chamber, RHEED appears bright and streaky, with faint reconstruction.

The normal recipe is unfortunately the result of iteration on what is known to be a bad procedure—it destroys the ion pump element—and as a result it seems unnecessarily arcane. In the end, the dose was originally optimized according to the buffer chamber pressure with the ion pump running—i.e. it was known relative to the pumping speed of the ion pump. Since the ion pump lifetime at the hydrogen pressure used during the etch is months rather than a decade, the procedure was modified after the untimely replacement of the pump elements. As a result, the pressure is set while the pump is running, the pump is switched off, and the hydrogen is pumped through the load lock turbo pump.

In detail, the temperature is set to about 525 °C on the buffer heater station; the hydrogen source is opened slightly to produce a buffer chamber pressure of mid-10^{-8} torr before heating up the cracking zone per the instructions in the manual. Once the cracking zone is at process temperature, the pressure is increased in the buffer chamber to 2-3 $\times 10^{-6}$ torr on the buffer ion gauge. After stability, the buffer ion gauge is turned off, the load lock–buffer gate valve is opened, and the ion pump is switched off; the hydrogen should be pumped using the load lock turbo pump. After an hour at 525 °C under hydrogen flux, the pressure is reduced while the substrate cools down, and the chamber is returned to its original state.

This procedure is in fact superior in that it produces a smoother surface, as is clear from the (111)B growth as well as studies on GaSb [6, 83], relative to thermally desorbing the GaSb oxide, which has nevertheless remained our common practice. The (001) film results reported in this thesis are entirely from films grown on substrates whose oxide was removed thermally. (The main advantage to thermal desorption, besides convenience, is that it provides a calibration point for the pyrometer, to ±10 °C.)

3.4 Buffer layer

The growth of Cd_3As_2 from compound (Cd_3As_2, rather than separate Cd and As) sources does not afford many engineering parameters. Essentially, the Cd_3As_2 flux and substrate temperature are the only ones, and it turns out that the growth is fairly insensitive to the first and very sensitive to the second. By contrast, opportunities for engineering of the buffer layer are numerous and promising, harnessing as they do the possibility of tuning the electronic properties of the interface.

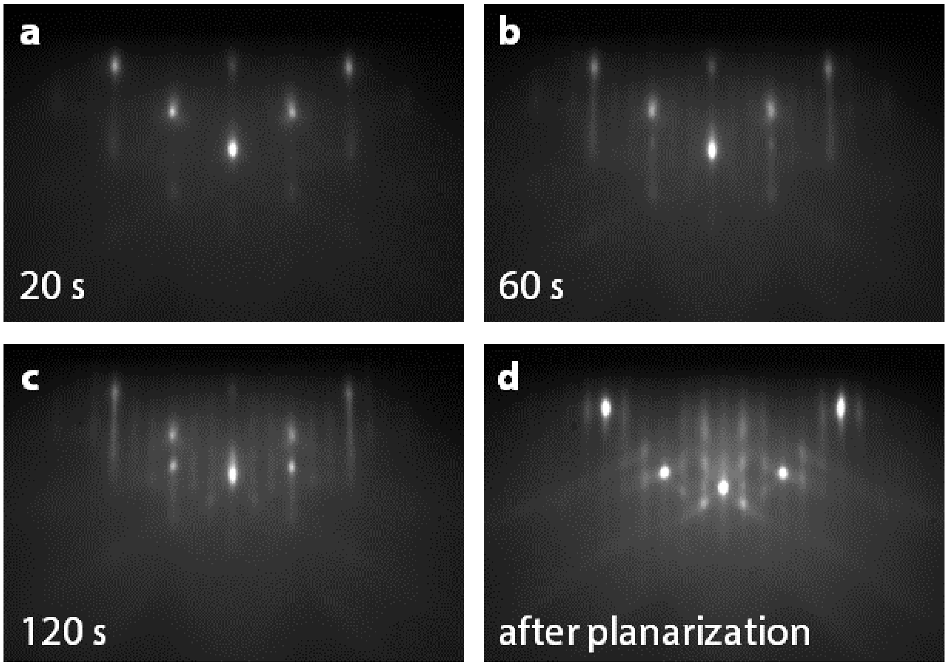

Figure 3.1: Evolution of RHEED during GaSb growth, starting after oxide desorption, viewed along the [110] direction, (a) after 20 seconds of growth at approximately 0.3 monolayers/s, (b) one minutes, (c) two minutes, and (d) about twenty minutes. Note that the faint 3× reconstruction is already visible after twenty seconds.

These buffer layers can be matched to the GaSb substrate, and thin strained layers of InAs, AlSb, and AlGaSb can be grown on them:

1. GaSb. GaSb homoepitaxy is an essential part of growing on GaSb substrates, since a hundred nm or so is enough to smooth out pits formed during oxide desorption. A good starting point is to grow 10 °C cooler than the oxide desorption temperature. Desorb the oxide under 3×10^{-7} torr BEP Sb flux; set the Ga source at about 900 °C or 1×10^{-7}

torr BEP Ga flux. RHEED should evolve as in Fig. 3.1. Under normal circumstances, the RHEED should stabilize in about a minute. If after two minutes it doesn't look right, close the Ga shutter and lower the Ga flux by 20%, and try again. If that does nothing, then start reducing the Sb flux in the same increment. In AFM, GaSb/GaSb(001) should look like Fig. 3.2—that is, essentially perfect: rows of atomic-height steps.

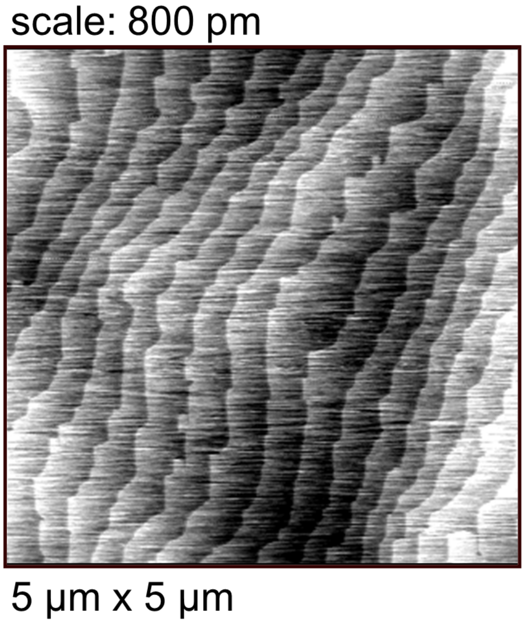

Figure 3.2: Atomic force micrograph showing surface after GaSb/GaSb(001) homoepitaxy. White is high and black is low.

2. AlGaSbAs. The keys to good AlGaSbAs growth are good GaSb growth and decent or better flux calibration. (The flux calibration is discussed in the Appendix.) AlGaSbAs is a highly attractive candidate buffer layer. The Al alloy fraction and the As alloy fraction accommodate appreciable tuning of the band offset [80]. AlSb and GaSb have fairly similar growth temperatures, and there is a comparatively large temperature window

for the alloy to be grown. The result is an insulating buffer layer that can be grown at very high quality. The main drawbacks are (1) that by its nature the alloy is disordered (see Section 5.1 for more discussion) and (2) that the As composition can be difficult to control. Ordinarily the issue is that the flux calibration effectively depends on the substrate temperature. On our system, there is an additional difficulty that the valve on the arsenic cracker source does not produce repeatable fluxes, and so the best practice is to check the flux directly before growing. (This is accomplished quickly with the valved source.)

As discussed for a particular film below (Section 3.5), perhaps the best way to measure the alloy composition relies upon calibration of the group III flux, which can be done to high precision using RHEED intensity oscillations (see Section B.3). At the relevant growth temperature, the incorporation of the incident Al and Ga atoms is unity, so the Al and Ga composition can be set in the growth recipe (provided the fluxes are known absolutely, i.e. in the units atoms/cm^2/s, rather than torr BEP—see Section B.3).

This procedure can be checked separately by measuring the lattice parameter of an AlGaSb/GaSb(001) layer. A hiccup one might encounter, however, is that one typically will not know *a priori* at what thickness the AlGaSb relaxes, what is called the “critical thickness” [53]. For AlSb/GaSb(001), the critical thickness has been measured to be about 14 nm [25], but elsewhere “above 150 nm” [24] and 154 nm [84]. In any event, the actual thickness at which a film relaxes is driven by kinetics, since it depends on the nucleation of the dislocations that allow the relaxation, so one usually needs to measure the amount of relaxation.

It is easy enough to measure if the film is relaxed or not. The best way is using a reciprocal space map (RSM) of a so-called off-axis peak (see Fig. 3.3 and Fig. 3.4). (As an aside, it may turn out that a desirable off-axis peak has a small scattering factor. This is one reason why it is usually preferable, at least for checking calibration layers whose thickness is not engineered, to grow thick films that scatter more x-rays, even if the tradeoff is partial relaxation of the layer. As mentioned below, even near-total ignorance of the elastic constants does not affect the result appreciably.) The RSM's main drawback is cost: one typically requires approximately 12 hours of measurement time on an x-ray diffractometer (say, at current rates, \$500). The measurement consists of a series of coupled $2\theta - \omega$ scans at different ω offsets. The ω and $2\theta - \omega$ axes can be transformed into the relevant components of the scattering vector, and it can be tested via peak finding (or, better, fitting) whether the buffer layer peak and substrate peak have the same in-plane scattering vector, as is the case in 3.3. At this point one can tell if the film is relaxed or not. For cubic crystals (like the zincblende III–Vs), the film is relaxed if in-plane lattice parameter is the same as the out-of-plane one, which can be seen from the associated scattering vector components. This condition can be drawn as a line on an RSM (see Fig. 3.4). If it is completely strained, the in-plane lattice parameter matches the substrate one, and if the film is partially relaxed, the peak will fall between the two conditions. Partial relaxation of the film or else its being coherently strained adds some uncertainty into the determination of the lattice parameters because they require knowledge of the elastic constants and Poisson ratio. These data have been tabulated [1], and in any event uncertainties in the elastic constants—say, in an exotic alloy—are reduced by orders of magnitude in the determination of the final lattice constant because they do not vary over a particularly large range and because they multiply a small number that is added to a larger one.

Once the Ga and Al fluxes are determined precisely, the Sb and As fractions are found by an RSM again.

3. Alloys containing indium should be avoided if possible due to the difference in growth temperature between InAs, or, worse, InSb, and GaSb or AlSb.

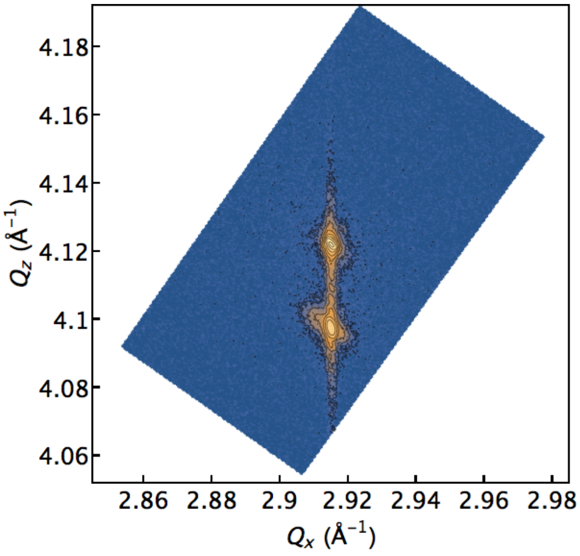

Figure 3.3: Reciprocal space map of an AlGaSb layer (bottom peak), lattice matched to GaSb (top peak).

These buffer layers can be matched to Cd_3As_2:

1. InGaSb

2. InAlSb
3. InGaAlSb, and
4. InSbAs.

The difficulties here are prohibitive. These are all alloys of InSb. The challenge of growing usable InSb is well-known. Perhaps the shortest proof of its troublesome nature is the fact that HgCdTe—a canonically difficult material to grow, with a growth window small enough to require advanced thermometry and growth rate monitoring, and toxic to boot—remains the industry standard for long-wavelength detectors. One difficulty is that, unlike for the members of the III–As family, the In:Sb flux ratio is an important growth parameter, which, worse, depends on temperature. Add to that its growth temperature is in the mid-300s °C, and its growth window is probably less than 20 °C, and it is enough to conclude that InSb is difficult to grow by MBE. Given that GaSb and AlSb grow, by contrast, at more like 500 °C, it is clear that some compromises are necessary to produce a good alloy. Considering besides the ancient controversy over whether the indium-containing alloys are actually stable mixtures [75], the dearth of reliable experimental results in the 6.3 Å neighborhood, and—almost as an afterthought—the lack of a suitable substrate, the challenge may seem hopeless. Kudos, then, to Manik Goyal for his work on the subject, which demonstrated material improvements in Cd_3As_2 films grown on such buffer layers [27]. Iteration on his recipes for buffer layers like nos. 1 and 2 above may be a fruitful avenue of future research.

But, at time of writing, the main motivation for pursuing exotic and challenging buffer layer growth is for additional control (strain, say), rather than improvement to the Cd_3As_2 transport. This is because the mobility-limiting mechanisms in Cd_3As_2 films are unknown in general, and the main evidence suggests that the exposure of the top surface masks whatever intrinsic differences can be controlled by the buffer layer (see Section 5.1 below). In other words, the 2020 state of the art in thin films can be achieved with whatever smooth buffer layer you like, whether lattice-matched (e.g. InGaSb) or not (e.g. AlGaSbAs, with a suitable wetting layer, per Section 3.5).

3.5 Structure for basic transport measurements

Much of the text and figures of this and the subsequent sections of this chapter reproduces or adapts portions of our work published previously [38], to which APS retains the copyright.

We present here a study based around the AlGaSbAs buffer layer mentioned above. A schematic of the heterostructure is shown in Figure 3.4a. After a solvent rinse, a cleaved piece of an undoped (001) GaSb wafer was loaded into a buffer chamber for outgassing. The chip was adhered to a tungsten plate with liquid gallium, as mentioned above, to improve temperature homogeneity and consistency in thermometry. The native oxide was thermally desorbed in the growth chamber under Sb flux, and a 150-nm-thick layer of GaSb was grown to smooth out any resulting pits before growth of a $Ga_xAl_{1-x}Sb_{1-y}As_y$ buffer layer for electrical isolation from the GaSb. While in principle an AlSb buffer layer electrically isolates the Cd_3As_2, the +0.649% lattice mismatch with GaSb can result in defects. The incorporation of small amount of Ga is thought to suppress the oxidation besides slightly reducing the lattice mismatch [45]; addition of As further reduces the lattice mismatch, allowing for the growth of thick buffer layers with low densities of extended defects. For example, high quality InAs quantum wells have been grown using $Ga_xAl_{1-x}Sb_{1-y}As_y$ buffer layers [78].

In quaternary alloys of Al, Ga, Sb, and As, the incorporation of the incident Al and Ga atoms is unity across a wide range of viable substrate temperatures, and as a result the desired group III composition can be achieved by setting the absolute flux of the sources. The relative Sb versus As incorporation, however, depends strongly on the substrate temperature. Here, the desired composition was achieved by maintaining a constant substrate temperature, measured by a pyrometer, and iterating on the beam equivalent pressure ratio of the group V sources. After the buffer layer, four monolayers (1.2 nm) of a thin wetting layer, either InAs, GaSb, or AlSb, were grown. The lattice constant of InAs differs from that of GaSb by -0.614%. We call the four monolayers a wetting layer because, given that the layer grows pseudomorphically on the underlying buffer layer, the differences in subsequent Cd_3As_2 growth on each must reflect differences in surface and/or interfacial energies. Finally, the growth parameters for (001) Cd_3As_2 were in most respects identical to those reported previously [69]. As mentioned above, high flux from two molecular Cd_3As_2 sources, totaling 2.6×10^{-6} torr on a flux monitoring ion gauge, was supplied to a low-temperature substrate (thermocouple temperature around 200 °C, uncorrected pyrometer temperature—see Section 3.3.2—around 115 °C).

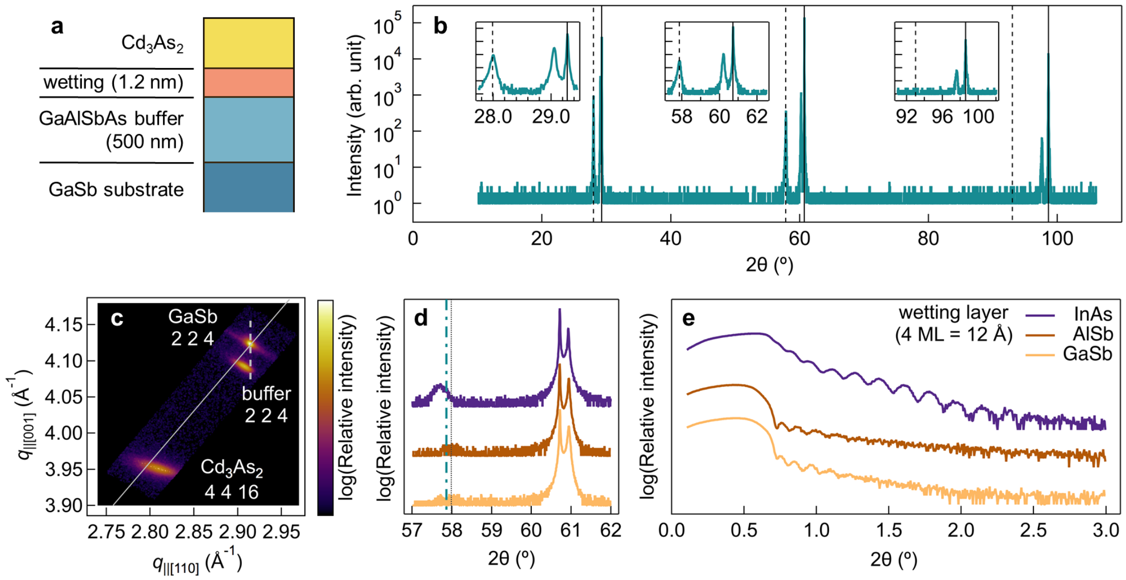

Figure 3.4: (a) Schematic of the heterostructure (not to scale). (b) Out-of-plane $2\theta - \omega$ x-ray diffraction scans from a 450-nm-thick (001) Cd_3As_2 film grown on an AlGaSb buffer layer with a 1.2-nm-thick InAs wetting layer. The vertical lines show literature values for the GaSb substrate (solid) and Cd_3As_2 layer (dashed). The reflections around $00l$ GaSb ($l = 2, 4$, and 6 from left to right) are shown in more detail in the insets. The $00l$ Cd_3As_2 reflections have $l = 8, 16$, and 24, but the $00\underline{24}$ Cd_3As_2 reflection is too faint to be seen. Unindexed peaks are from the AlGaSb buffer layer. (c) Reciprocal space map in the vicinity of the 224 GaSb reflection performed on the same film as shown in (b). The axes show the out-of-plane and in-plane components of the scattering vector, $q_{||[001]}$ and $q_{||[110]}$, respectively. The diagonal line (gray, solid) shows the cubic condition, i.e. $a = c$, or $q_{||[110]} = \sqrt{2}q_{||[001]}$. The short vertical line (gray, dashed) shows the condition for the buffer layer being fully strained to the substrate. The Cd_3As_2 peak is not expected to fall along the cubic condition line because its lattice is tetragonal. (d) Out-of-plane $2\theta - \omega$ x-ray diffraction scan in the vicinity of 004 GaSb for approximately 45 nm films grown on different wetting layers, InAs, AlSb, and GaSb [see legend in (e)]. The buffer layer (second most intense peak) is a quaternary alloy $Ga_{0.2}Al_{0.8}Sb_{0.9}As_{0.1}$. The scans are offset for clarity. Two literature values for the $00\underline{16}$ Cd_3As_2scattering angle are indicated with vertical lines, one in agreement with the reflections in (b) [74] (dot-dash, teal), and the other corresponding to a slightly smaller lattice constant (dot, black) [3]. The fact that the $00\underline{16}$ Cd_3As_2peak in the top trace (InAs wetting layer) does not fall along the dot-dash line is likely due to residual epitaxial strain. (e) X-ray reflectivity of the three films shown in (d). The film grown on InAs is the smoothest and also apparently thinnest, judged by the tail-off of the oscillations and their period, respectively, even though all growth parameters were held constant.

3.6 Basic structural and transport properties of (001) films

Figure 3.4b shows a high-resolution, out-of-plane $2\theta - \omega$ XRD scan of a sample with a 450-nm-thick Cd_3As_2 film on an InAs wetting layer and AlGaSb buffer layer. In the vicinity of the 002 GaSb substrate reflection (magnified in the left inset), the 008 Cd_3As_2 and 002 AlGaSb reflections are seen. The 1.2-nm-thick InAs wetting layer does not have enough volume to scatter x-rays above the detection limit. Another sequence of peaks is visible near the 004 GaSb reflection (center inset), starting with $00\underline{16}$ Cd_3As_2. The kinematical scattering intensity of the $00\underline{24}$ Cd_3As_2 reflection, expected near 006 GaSb (right inset), is less than that of the 008 Cd_3As_2 peak by a factor of more than 10^3 and is not observed. The lack of other film peaks confirms the (001) orientation of the Cd_3As_2 film.

To verify the in-plane epitaxial alignment, the reciprocal space map in Fig. 3.4c shows a region in the vicinity of the 224 GaSb reflection. The diagonal line indicates the 224 peak position for a relaxed cubic layer. The vertical dashed line shows the position for a coherently strained film (that is, the in-plane lattice parameter is identical to that of the substrate). The buffer layer peak falls between the two lines, which indicates partial strain relaxation, and its lattice parameter is about 0.5% larger than that of the substrate. From the position of the $44\underline{16}$ Cd_3As_2 peak, its lattice parameters are determined to be a = 12.65 Å and c = 25.44 Å, which agree well with the published values for bulk Cd_3As_2[3, 74]. The c/a ratio is 2.01, consistent with a (001)-oriented Cd_3As_2 film.

The data shown in Figs. 3.4d and 3.4e are from samples grown under nominally identical conditions but on $Ga_xAl_{1-x}Sb_{1-y}As_y$ buffer layers with different wetting layers: InAs, AlSb, and GaSb, respectively. Figure 3.4d shows out-of-plane $2\theta - \omega$ XRD scans in the vicinity of the 004 GaSb reflection. The $Ga_xAl_{1-x}Sb_{1-y}As_y$ buffer layer composition can be calculated using $x = 0.8$ (from flux calibration) and regarding the alloy as a mixture of $Ga_{0.2}Al_{0.8}Sb$ and $Ga_{0.2}Al_{0.8}As$. Assuming a fully relaxed buffer layer, we find $y = 0.09$; in the case of a fully strained buffer layer, $y = 0.11$. This lattice mismatch relative to the Cd_3As_2 layer is -4%. Despite the large mismatch, thin Cd_3As_2 films grown on heterostructures with InAs wetting layers show an elongation of the out-of-plane lattice parameter (cf. lines in Fig. 3.4d) that is, at least in part, due to residual epitaxial coherency strain. Furthermore, the differences in the scattered intensity of the $00\underline{16}$ Cd_3As_2 peak reflect differences in their defect densities on the three wetting layers. The film on InAs is also smoother, because oscillations in XRR survive to higher scattering angles (see Fig. 1e), and this is also evident in the AFM images shown in Fig. 3.5. Nevertheless, the morphology of the (001) films differs from the smooth, step-like surfaces of (112) Cd_3As_2 MBE films [68]. Furthermore, as judged by the period of the fringes in x-ray reflectivity (XRR), the Cd_3As_2 films grown on AlSb and GaSb are thicker than on InAs, which should increase the intensity of the XRD peaks relative to those of the film grown on InAs.

Cross-section transmission electron microscopy samples were prepared by Honggyu Kim using Ga ions in a focused ion beam system with a final milling energy of 2 keV. High-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) imaging was performed on an FEI Titan S/TEM ($C_s = 1.2$ mm) operated at 300 kV with a convergence semi-

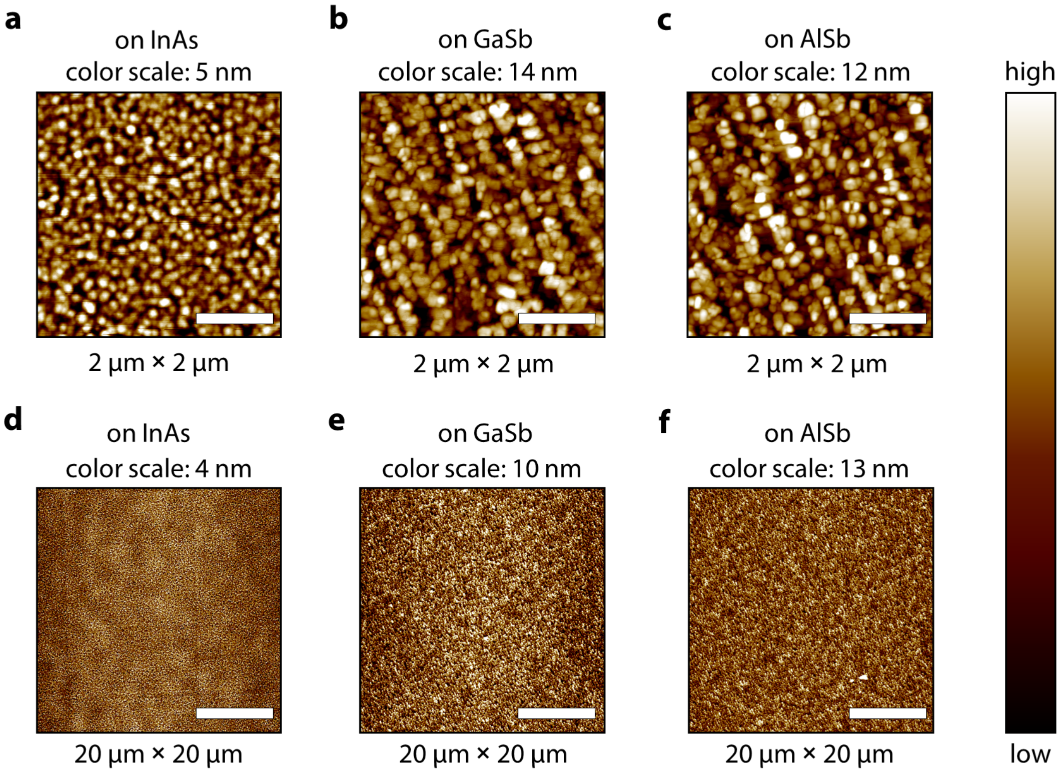

Figure 3.5: AFM images of the surfaces of the three Cd_3As_2 films shown in Figs. 3.4d and 3.4e, grown on different wetting layers (see labels). The lateral scan dimensions are 2 μm × 2 μm in a-c. Note the different height scales, indicated as the color scale for each image. Lower magnification images are shown in panels d-f, with lateral dimensions 20 μm × 20 μm. All images have undergone a first-order plane leveling procedure.

angle of 9.6 mrad. The experimental HAADF-STEM images were denoised with the Wiener filter [47]. The simulated image of Cd_3As_2 (space group $I4_1/acd$ [3]) was calculated using the Kirkland multislice algorithm [41].

Figure 3.6a shows a cross-section HAADF-STEM image, thanks to Honggyu Kim, viewed along [100]. Shown are the Cd_3As_2/InAs interface and the top few unit cells of the buffer layer. Comparison with simulations (Fig. 3.6b), noting especially the orientation of rows of subtly elongated butterfly-shaped features, which are caused by the displacements of cadmium atoms, further confirms the [001] growth direction as found from the lattice constant measured using XRD.

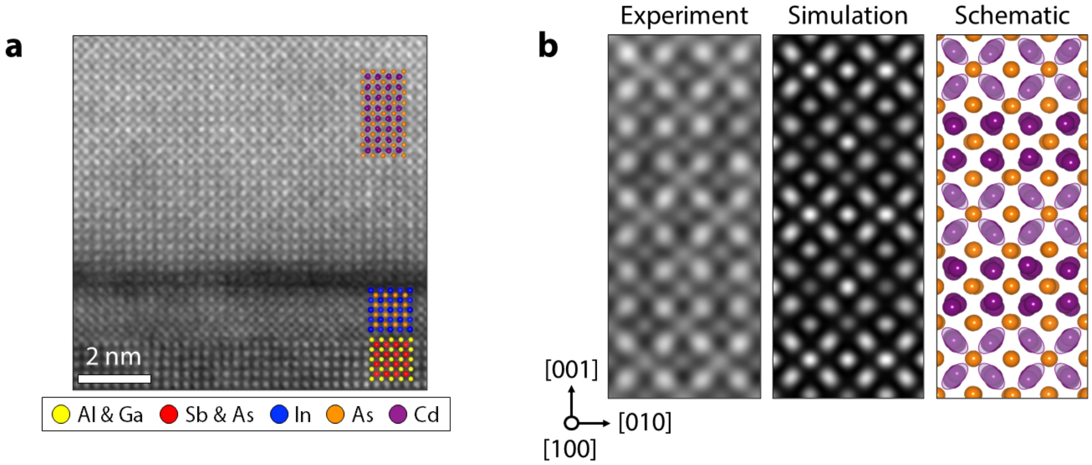

Figure 3.6: (a) HAADF-STEM image of a Cd_3As_2 film on an InAs wetting layer with a $Ga_{0.2}Al_{0.8}Sb_{0.9}As_{0.1}$. buffer layer, recorded along [100]. (b) Magnified image of the Cd_3As_2 film, simulated image, and schematic. Due to the displacement of the Cd atoms, the projected Cd atomic columns are seen as a characteristic elongated, butterfly-like shape (purple shaded ellipses), which form rows parallel to [010].

After structural characterization by atomic force microscopy (AFM) and x-ray diffraction and reflectivity, Hall bar structures were fabricated using standard photolithographic techniques. Mesas were isolated by argon ion milling, and ohmic contacts were deposited as Au/Pt/Ti stacks. Unless stated otherwise, samples were exposed to a nitrogen plasma treatment prior to electrical measurement to reduce the density of low-mobility p-type carriers on the top surface [21]. Electrical measurements were performed in a Quantum Design Physical Properties Measurement System using a low-frequency ac excitation, detected with a lock-in amplifier. Unless stated otherwise, the longitudinal magnetoresistance data shown was symmetrized and the Hall magnetoresistance antisymmetrized under a change in magnetic field direction.

Electrical transport measurements on three Hall bar devices fabricated out of approximately 45-nm-thick Cd_3As_2 films on different wetting layers, measured at 2 K, are shown in Fig. 3.7a. The similarity in the longitudinal magnetoresistance (R_{xx}) for films on AlSb and GaSb may be attributed to the granularity of the Cd_3As_2 films on both wetting layers (see Fig. 3.5), which causes the current traveling through a small number of paths that are only weakly connected. By contrast, the device on the film with the InAs wetting layer (solid lines in Fig. 3.7a) shows a factor of 50 lower zero-field resistance. The difference is especially evident in Fig. 3.7b, which shows only the results from the film on InAs. The incipient quantum Hall effect seen in Figs. 3.7a and 3.7b is a further sign of the high quality of the Cd_3As_2 film grown on InAs.

The differences in the longitudinal resistance between the three devices cannot be attributed solely to the difference in carrier concentration. Specifically, the 2D carrier concentrations, as determined by the Hall effect (Fig. 3.7a) are 1.02×10^{12} cm^{-2} (on InAs), 7.76×10^{11} cm^{-2} (on AlSb), and 6.47×10^{11} cm^{-2} (on GaSb), i.e. they vary by less than a factor of two. The differences in carrier concentrations correlate to the relative positions of the conduction bands in the three wetting layers [80], which suggests that band bending at the bottom interface may play a role. It is important to note, however, that transport is not through the wetting layer itself. The main evidence is the fifty-fold reduction in longitudinal resistance of samples with InAs wetting layers, relative to the other two, which is accompanied by only a two-fold increase in charge carriers. Furthermore, consider the raw (that is, not antisymmetrized or symmetrized) resistance data shown in Fig. 3.7 before and after a N* plasma treatment. In this sample, the Cd_3As_2 layer is slightly thicker (53 nm), and the carrier density is 1.25×10^{12} cm^{-2}. As shown in an earlier study on (112) films [21], the low-energy nitrogen plasma changes the band bending at the top surface to result in high-mobility, single-carrier (n-type) 2D transport. As seen in Fig. 3.7, the condition of the top surface appreciably changes the measured resistance also for these (001) films. This is further evidence that the measured resistance is not mainly due a buried conductive layer. The sensitivity of the transport properties, including the emergent quantum Hall effect, to the surface treatment, hints at the importance of surface states also in (001) films. Finally, the Hall mobility of $9,300$ $cm^2/V\,s$ in the 45-nm-thick Cd_3As_2 film on InAs (Fig. 3.7b) is nearly identical to that of the thicker 53-nm-thick film, $9,170$ $cm^2/V\,s$ (Fig. 3.7c) and compares well to that reported for MBE-grown (112) films of similar thickness [67].

3.6.1 Postscript

Recent work in our group has shown that good growth can be achieved on indium-containing alloy buffer layers such as InGaSb without growing an InAs wetting or adhesion layer. The demands of the desired device rule here: there's a lot of tuning possibilities with InAlSb and InGaSb alloys, such as the strain, band offsets, and even whether the buffer layer has a direct bandgap; the tradeoff is that the growth is more complicated to achieve, which may be prohibitive for the structures described in Section 5.1. Both the InAlSb and InGaSb alloys, at least in the lattice constant neighborhood of about 6.3 Å, tend to produce highly defective films,

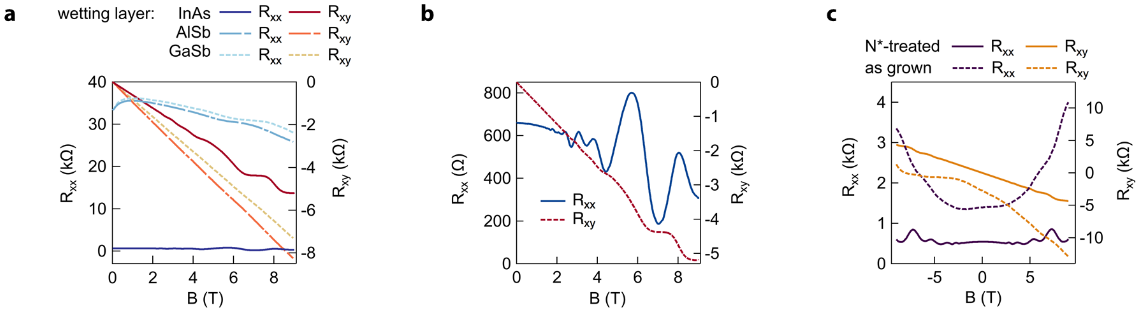

Figure 3.7: (a) Comparison of 2 K magnetotransport behavior of Cd_3As_2 films differing only by wetting layer. These are the same films as in Figs. 3.4d and 3.4e and Fig. 3.5. The longitudinal resistance (R_{xx}) is shown in shades of blue and the Hall resistance (R_{xy}) is shown in shades of red. The device sizes are: 10 μm × 10 μm on InAs (solid), 50 μm × 50 μm on AlSb (dot-dash), and 100 μm × 100 μm on GaSb (dash). (b) Higher magnification of magnetoresistance behavior for the 45-nm-thick Cd_3As_2 film (on an InAs wetting layer) shown in panel a. Strong quantum oscillations in the longitudinal magnetoresistance (solid blue) are accompanied by incipient plateaus in the Hall magnetoresistance (red dashes). (c) Comparison of magnetoresistance behavior for a 53-nm-thick Cd_3As_2 film on an InAs wetting layer before (dashes) and after (solid) surface treatment by a low-energy nitrogen plasma. The more symmetric R_{xx} behavior after cleaning (dark blue) and more antisymmetric R_{xy} (orange) after cleaning reflect the fact that the charge transport is less scattered, and thus there is less mixing between the two channels. Unlike data in panels a and b, these data were not symmetrized or antisymmetrized with respect to reversal of the magnetic field.

though mitigation is possible. Perhaps more interestingly, it appears that the relationship between how defective the buffer layer is and versus how defective is the Cd_3As_2 layer is more complicated than a simple correspondence [27].

Chapter 4

The quantum Hall effect in (001) films

The main topic of this chapter is the basic transport properties of (001) thin films. After a summary of some of our earliest results, we discuss the quantum Hall effect in very high magnetic fields (up to 45 T) in the context of the models introduced in Chapter 2. Transport in thinner films and at lower fields provides an additional check on our conclusions, as subsequently discussed. Inconclusive experiments using rotation of the sample inside the magnetic field are also presented.

4.1 Summary of early (001) transport results

The InAs wetting layer is a route to highly consistent (001) Cd_3As_2 films. The immediate hurdle, following that result, to observing the quantum spin Hall insulator—per ref. [91], films of thickness between about 10 and 20 nm are necessary—was the continuing difficulty in reproducibly fabricating devices on thin films. Given the constraints of immediate possibility, and the high mobilities achieved in the roughly 50-nm-thick (001) films, about $10,000\ \mathrm{cm^2/V\,s}$ [38], we turned our attention to these slightly thicker films. Early results, shown in Fig. 4.1, differed from those on (112) films.

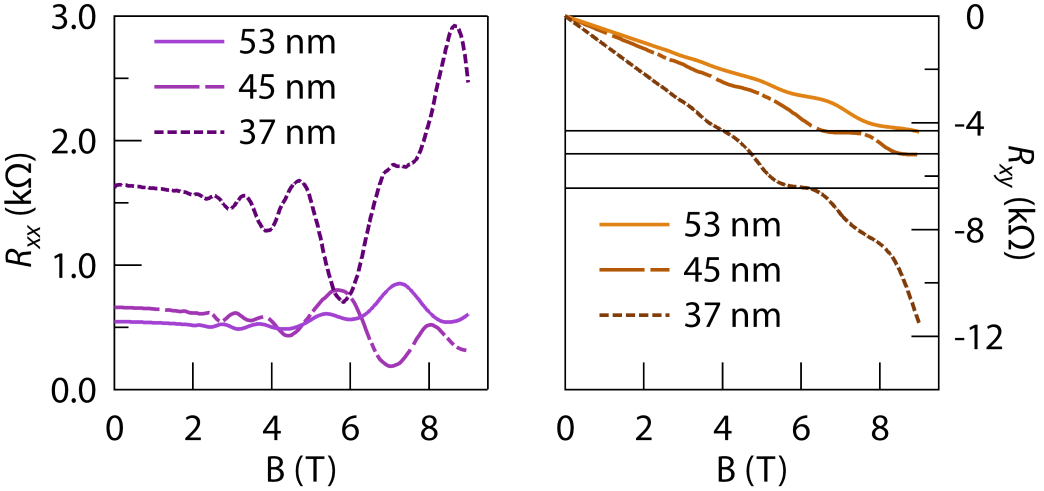

Figure 4.1: The quantum Hall effect in three (001) films, no cap, no gate. In the righthand plot, of the Hall resistance, three horizontal lines are shown at $\nu = 4, 5, 6$.

In the quantum Hall effect in (112) films [67, 22, 26], an apparently robust but not particularly well-understood feature was the sequence of filling factors. For example, in ref. [67], the first observation of the quantum Hall effect in Cd_3As_2 films, there appeared to be some features in R_{xx} corresponding to $\nu = 6$ and $\nu = 4$, essentially no feature at $\nu = 3$ or 5, and prominent plateaus at $\nu = 2$ and $\nu = 1$, which, to the sample-agnostic, might be said to look "conventional," in the sense that, on its own, such a filling factor sequence might be observed in an AlGaAs/GaAs quantum well, where the steps of two in the filling factor are due to spin degeneracy but, at high field, every single filling factor appears due to "spin splitting." In

Cd_3As_2, though, the spin degeneracy is thought to be lifted at very low field; measurements of the g factor vary, but have been reported as 16 [58] or 15-30, depending on the carrier density and orientation of the 2D orbit [7]; early modeling literature suggests slightly higher, 30-40 [8, 87]. As a result, the origin of the near-degeneracy of two in the filling factors was a topic of some interest and speculation.

In that context arrived these (001) Cd_3As_2 films. Even though they were too thick to demonstrate the quantum spin Hall insulator phase, the origin of the quantum Hall effect, as shown in Fig. 4.1, seemed to be different from that in the (112) films. In films of different thickness, plateaus at filling factors both odd and even appeared. Moreover, which ones appeared seemed to depend on the thickness or carrier density: since the carrier density changed with the thickness, it was unclear which was explanatory. Finally, the thickness dependence seemed slightly different from that in the (112) films [26]. There, as the thickness increased to 60 or 70 nm, the quantum Hall effect, visible in thinner films, was suppressed by 3D transport. Here, in films as thick as 55 nm or so, the transport was clearly 2D, as supported by the observation of the quantum Hall effect. The change of the carrier density with thickness apparently matched the change in carriers measured by the Shubnikov–de Haas effect in the (112) films, which in that work were attributed to surface states. Taken all together, with the evidence from the nitrogen plasma treatment (see Fig. 3.7), we hypothesized that the difference in filling factor sequences between the (001) and (112) films was due to the difference in surface states in either case.

4.2 Gated structure in high fields

The text and figures of this section reproduce or adapt portions of our work published previously [37], to which APS retains the copyright.

Disentangling the carrier density from the thickness was addressed by the addition of a top gate, the spin splitting ambiguity by a large magnetic field. Two (001)-oriented Cd_3As_2 thin films, of thickness 45 nm and 50 nm, as confirmed by x-ray reflectivity, were grown by molecular beam epitaxy, as described in Chapter 3, on a compound semiconductor structure, which is shown schematically in Fig. 4.2a. The structure consisted of a doped GaSb substrate, covered by an insulating GaAlSbAs buffer layer, a thin InAs wetting layer, and Cd_3As_2. For the gated Hall bars (Fig. 4.2b), mesas were isolated by Ar ion milling, and the Al_2O_3 gate dielectric was deposited using a low-temperature (120 °C) atomic layer deposition process. Ohmic and gate contacts (Au/Pt/Ti) were deposited by electron-beam deposition. Magnetoresistance measurements were performed at the 45 T hybrid magnet at the National High Magnetic Field Laboratory, using the 300 mK ^{3}He insert. The resistance data were acquired using standard low-frequency lock-in techniques and subjected to a moving average smoothing routine before plotting.

Figures 4.2c and 4.2d show the longitudinal resistance R_{xx} and Hall resistance R_{xy} for the 45-nm-thick film as a function of magnetic field under varying top gate bias, corresponding to a variation in the total 2D carrier density ranging from 8.2×10^{12} cm^{-2} to 1.7×10^{12} cm^{-2}. The corresponding plots for the 50-nm-thick film are shown in Figure 4.3.

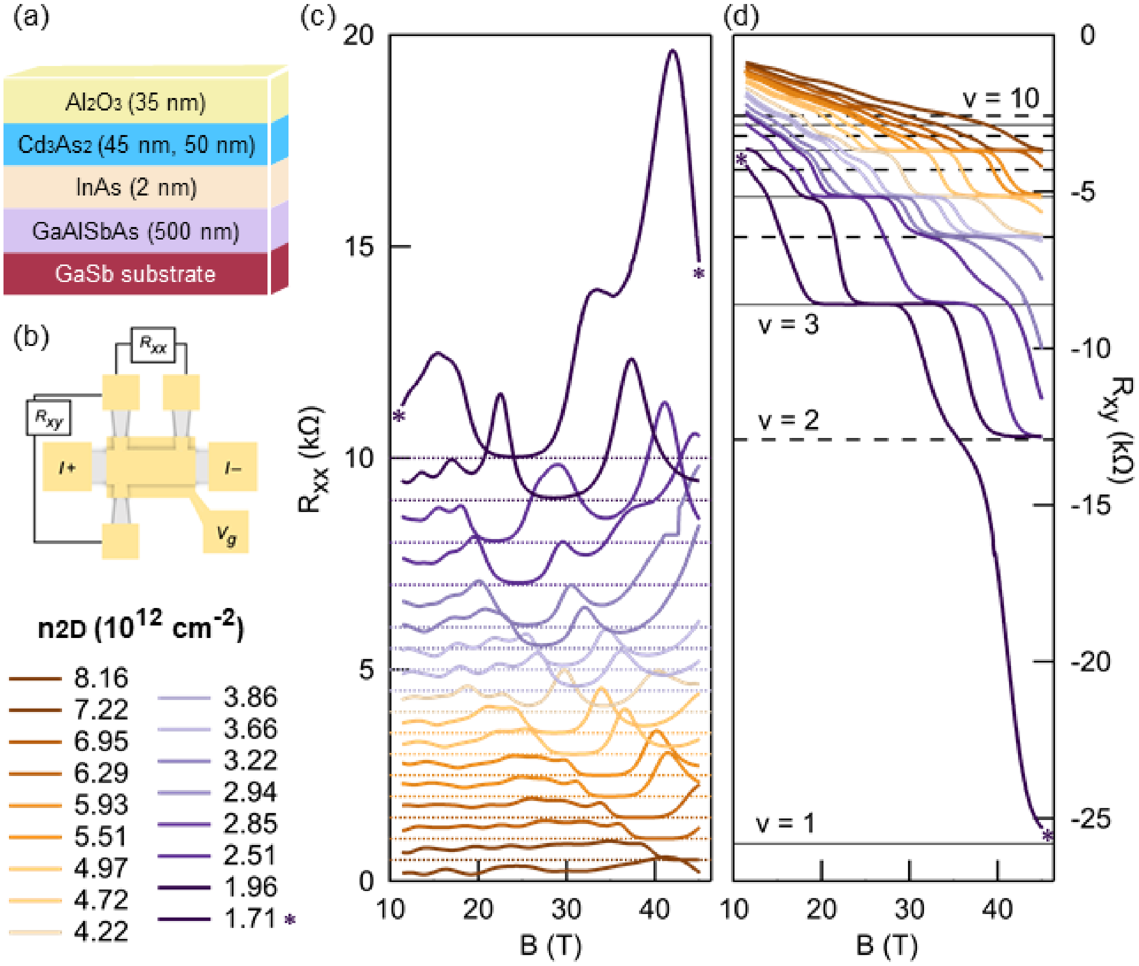

Figure 4.2: (a) Schematic of heterostructure (not to scale). (b) Schematic of the gated Hall bar device (not to scale). The gate metal is separated from the mesa by the Al_2O_3 layer and the other contacts are in contact with the Cd_3As_2 layer. (c) Longitudinal magnetoresistance (R_{xx}) and (d) Hall resistance (R_{xy}) for the device fabricated on a 45-nm-thick Cd_3As_2 layer. Different traces correspond to different gate bias or, equivalently, carrier density. The 2D carrier densities are calculated from the 11.5 T – 11.75 T Hall resistance and correspond to gate bias from about +3 V to −4 V. In panel (c), traces are offset for legibility. For each trace, $R_{xx} = 0$ is indicated by a dotted horizontal line. In panel (d), odd (even) filling factors are indicated by solid (dashed) lines.

As in the low-field results (Fig. 4.1), an unusual feature of the quantum Hall effect observed here is the prominence of plateaus with odd filling factors (indicated by the solid lines in

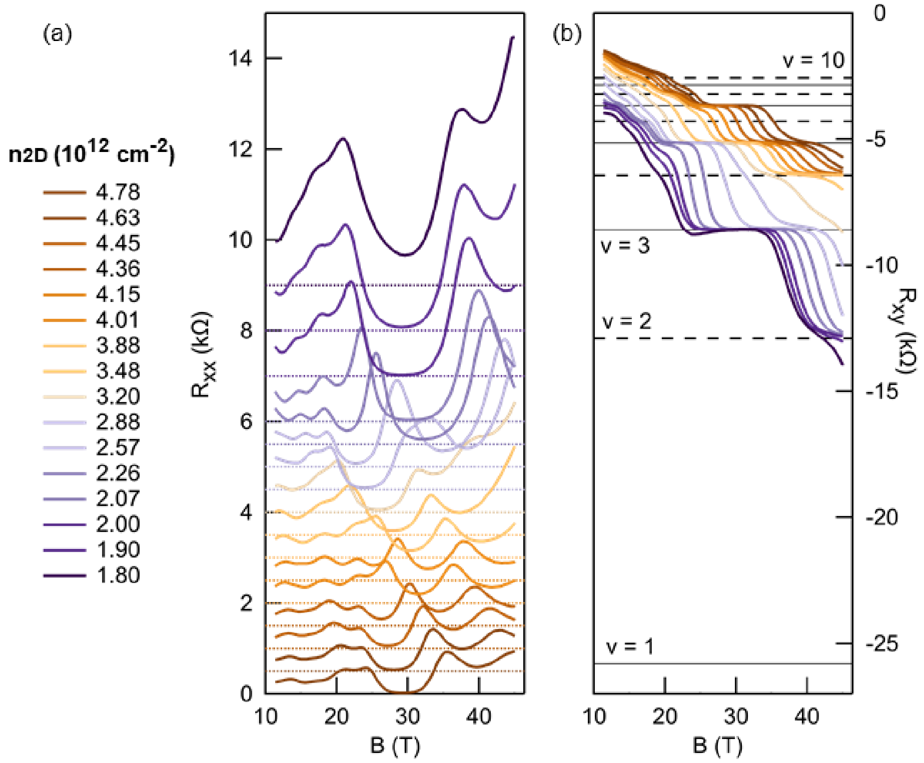

Figure 4.3: Magnetoresistance data for device fabricated on 50-nm-thick Cd_3As_2 layer. (a) Longitudinal resistance R_{xx} at different carrier densities, calculated from the 11.5 T to 11.75 T Hall resistance R_{xy} shown in (b). The gate bias ranges from about +1 V to −2.25 V.

Figs. 4.2d and 4.3b). Independent of the carrier density, the quantization of $R_{xy} = (1/\nu)(h/e^2)$ is more exact, and the plateaus more fully developed, for $\nu = 3, 5, 7$. By contrast, plateaus occurring at even filling factors (dashed lines) in many cases do not develop fully: there are plateau-like features in the Hall resistance, but the slope $\mathrm{d}R_{xy}/\mathrm{d}B$ does not reach zero. This difference is also reflected in R_{xx}, where the minima approach zero more closely as the Hall resistance reaches plateaus at odd filling factors than at even ones.

The Landau level structure is more apparent in the two-dimensional maps of the longitudinal resistance, R_{xx}, plotted as a function of carrier density and magnetic field, shown for both samples in Fig. 4.4a. The behavior of the two samples is nearly identical. Regions with the same filling factor, determined by their R_{xy} values, are oblong in shape and obey a complicated ordering scheme. For example, the filling factor increases by three between the regions $\nu = 5$ and $\nu = 8$; across the entire diagram, adjacent minima of R_{xx} coincide with steps in the filling factor of 1, 2, and 3. That the steps of 1 and 2 are not restricted to the high- or low-field regime indicates that these patterns are not due to the resolution of a degeneracy such as spin-splitting, but rather to the overlap of more than one fan of Landau levels, each originating from different zero-field energy states. As shown in Figs. 4.4b and 4.4c, the lowest values of R_{xx} coincide with the highest activation energies, deduced from an Arrhenius-type fit of the $T > 1.5$ K temperature dependence of the R_{xx} minima, while higher R_{xx} minima correspond to lower activation energies. In particular, the activation energy is greater at higher odd values of ν than lower even ones, in contrast to most conventional two-dimensional systems for which the activation energy increases with decreasing ν. Apparent deviations in the temperature plots are due to where the plateau regions were traversed; see white lines in Fig. 4.4a. In other words, when the path crosses the edge of an oblong minimum, the fitted activation energy is lower.

Figure 4.5 shows the behavior of the devices (on the 45 nm film) under larger negative gate bias than shown in Fig. 4.2, with the darkest purple trace common to both figures (as indicated by the asterisks). This trace still has a well-developed $\nu = 3$ plateau and approaches one at $\nu = 1$. Under increasingly negative bias, with $n_{2D} < 1.7 \times 10^{12}$ cm^{-2}, however, a striking and abrupt change appears in the Hall resistance: it becomes nonlinear in magnetic

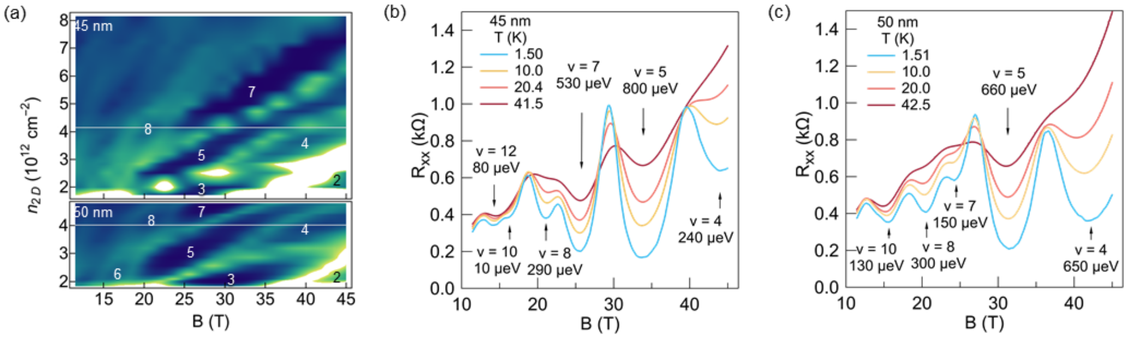

Figure 4.4: (a) Maps of the longitudinal magnetoresistance (scaled from dark blue to light green, indicating low to high) for devices made from the 45 nm and 50 nm films, respectively, as a function of magnetic field and carrier density. The two panels have the same vertical and horizontal scales. Thin horizontal gray lines correspond to the carrier density for the traces in (b) and (c). The spotty features are artifacts of the 2D plotting procedure. These are the same data as shown in Figs. 4.2c,d and 4.3a,b. (b) Temperature-dependent Shubnikov–de Haas oscillations from 1.5 K to 42 K for the 45-nm-thick sample. The activation energies shown are calculated from an Arrhenius-type fit to the indicated R_{xx} minima, labeled by the filling factor deduced from the corresponding value of R_{xy}. These traces would fall along the thin gray horizontal line on the 2D map in panel a. (c) Same as (b) from 1.5 K to 43 K for the 50-nm-thick sample.

field and ceases to form clear plateaus. Most tellingly, this behavior is not restricted to a particular magnetic field or carrier density (gate bias) alone, but instead coincides with the combined parameters for which the plateau at $\nu = 1$ is expected to develop fully. For some gate voltages, the quantum Hall plateaus vanish altogether. At the same time, the longitudinal resistance increases by more than a factor of 10. Nevertheless, features in the longitudinal magnetoresistance depend on both the magnetic field and the carrier density. The behavior was similar for both films (data for the 50-nm-thick film are shown in Fig. 4.6 and reproducible in successive bias sweeps.

The exact quantization of the Hall resistance and the concomitant near-vanishing of the longitudinal magnetoresistance (which fall to 10 Ω in the widest plateaus) is strong evidence

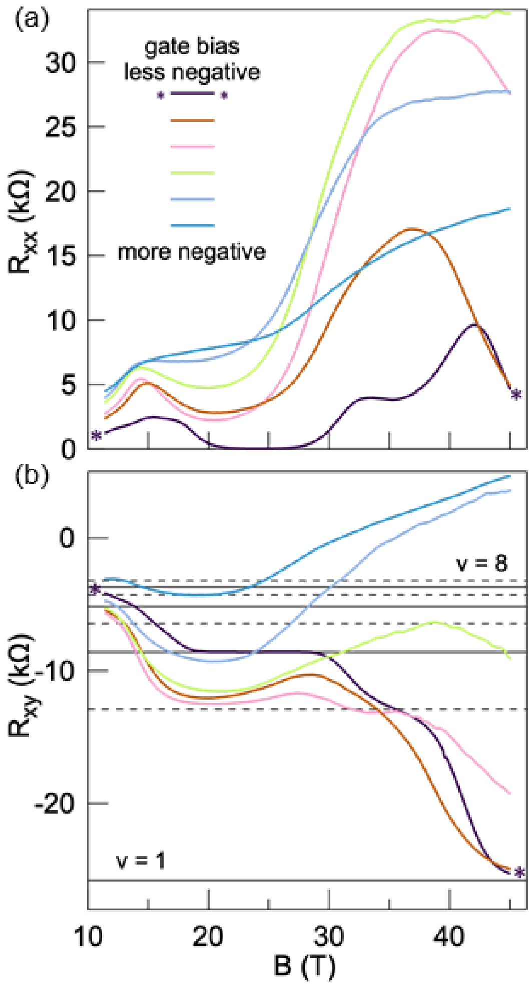

Figure 4.5: (a) R_{xx} and (b) R_{xy} data under large negative bias ranging from -4 V to -11 V for the 45-nm-thick sample. The dark purple traces marked with asterisks are carried over from Fig. 4.2c and 4.2d for comparison. In panel (b), odd (even) filling factors are indicated by solid (dashed) lines. The analogous data are shown for the 50-nm-thick sample in Fig. 4.6.

for the two-dimensional nature of the transport. Given the relatively large thickness of the films, purely two-dimensional transport is strong evidence that surface states are responsible for the observed spectrum. Such surface states can be regarded as a consequence of the inverted band structure of the Cd_3As_2 bulk that is in contact with, on either side of the film, (1)

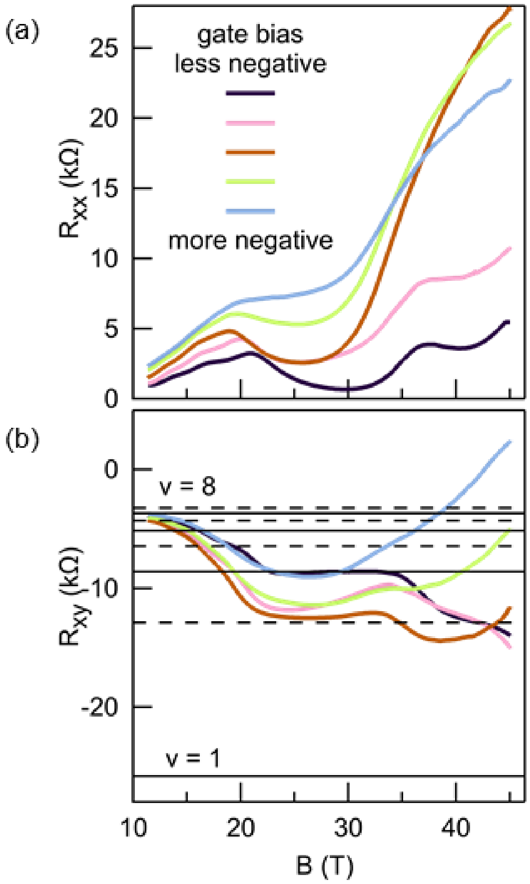

Figure 4.6: Higher bias R_{xx} and (d) R_{xy} data from the 50-nm-thick sample. The largest gate bias applied is −6 V (cf. Fig. 4.5).

the conventionally ordered bands of the III–V semiconductor layer and (2) the gate dielectric. As in the surface states of a 3D TI, this should give rise to a half-quantized Hall effect on each surface, and only odd integer filling factors should be observed in the absence of electron-electron interactions and/or a potential difference across the bulk [32, 66, 20, 18, 93, 55, 11]. Here, we show that the sequence of Landau levels observed in Figs. 4.2, 4.3, and 4.4 is consistent with these helical 3D TI like surface states. Figure 4.8a shows the combined Landau level spectrum of the top (purple) and bottom (orange) surface states of a generic 3D TI [77]. In

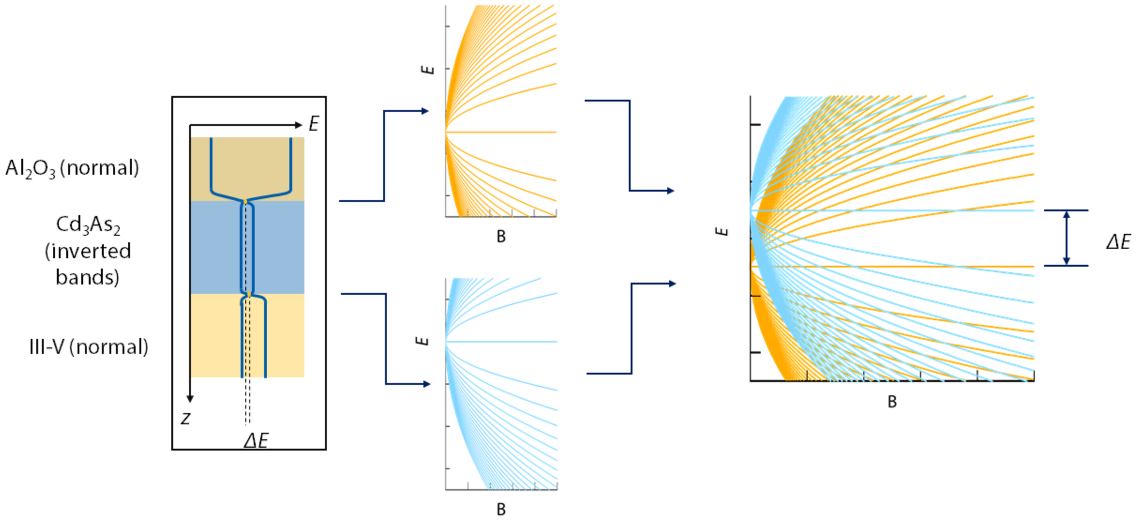

Figure 4.7: Cartoon of 3D TI surface state model: Dirac cones on either surface differ by an energy offset. Left: Zero-field band structure. There is a 2D Dirac node on either surface of the Cd_3As_2 layer (orange dot) where the band inversion is resolved. Middle: In a quantizing field, the Dirac cone either surface is split into Landau levels. Right: The combined spectrum is two overlapping sets of Landau levels separated by an energy offset ΔE. No effect of the gate is shown. Compare Fig. 4.8 for the actual model.

Fig. 4.8a we take into account that the two surfaces, each hosting a massless surface state, differ by a small offset in potential energy due to differences between bottom and top surface inherent to the heterostructure, which we refer to elsewhere as Δ_i (see particularly Sections 2.2.3 and 4.3.3; also Fig. 2.4, and Fig. 4.7). Furthermore, a multiplicative factor heuristically accounts for the screening by the top surface in the action of the gate bias on the bottom one, which effectively modifies the potential difference as a function of gate voltage. This scenario (in which the 2D carrier density is split between the two surfaces) satisfactorily explains the crossings of the Landau levels originating from each surface. More to the point, it reproduces the sequence of filling factors seen in both samples. There are a couple of different ways of doing this, which might seem more or less artificial—in the end, the real physical quantity is the total filling factor as quantized in the Hall voltage. It may seem slightly artificial, but we count the Landau levels on the notion that the total $\nu_T = (n_b + 1/2) + (n_t + 1/2)$, where ν_b is the filling factor one would find for the bottom fan alone and ν_t the same for the top. Each zeroth Landau level, whose energy does not depend on the magnetic field, contributes a filling factor offset of $1/2$. Then, we start counting from $\nu = 1$ above the higher of the two zeroth Landau levels. Because of the single degeneracy of the massless states, the filling factor increases by one as we trace an upward path across a Landau level. The crossings of the Landau levels originating on each surface produce a sequence of odd integer filling factors in the intermediate field (approximately 20 T to 30 T) range, matching the experimental data. The apparent jump of three in the filling factor from $\nu = 5$ to $\nu = 8$ can be seen to be due to the close spacing of the $\nu = 6$ and $\nu = 8$ regions of the map. The scheme of Fig. 4.8a is spelled out in Table 4.1.

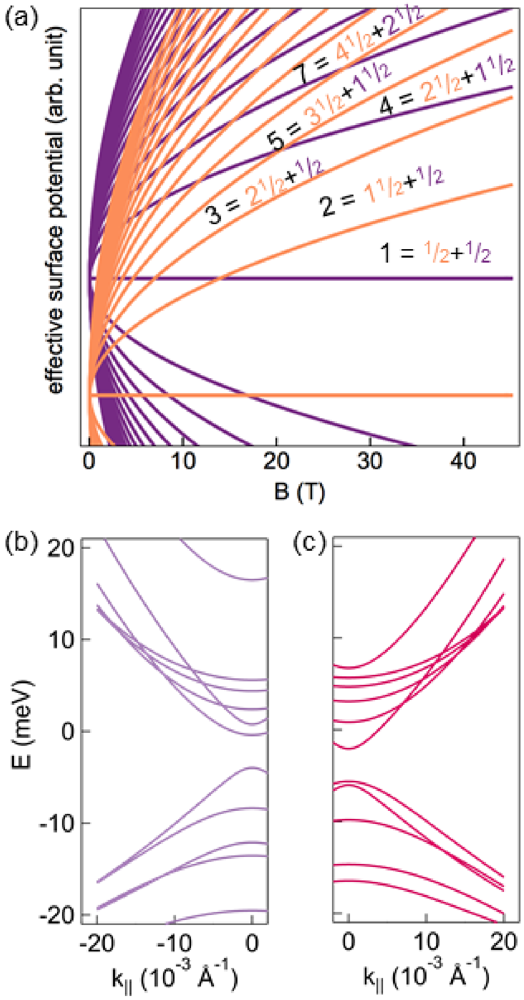

Figure 4.8: (a) Schematic of Landau levels formed by two topological-insulator-like surface states, with filling factors labeled. The energy for two sets of Landau levels is plotted, with the energy of the set originating from the bottom surface scaled relative to the set originating from the top to account for the screening of the gate potential by the top surface. (b) Quantum well subbands formed in a 45-nm-thick Cd_3As_2 film in a simple particle-in-a-box picture. (c) Quantum well subbands formed in a 50-nm-thick film.

Perhaps a less artificial way to count is to start at charge neutrality, i.e. between the two non-dispersing Landau levels, and, as before—what other choice is there?—count in steps of

1 across each Landau level. The drawbacks to this approach are fairly toothless. Chief among them, perhaps, is that the two fans end up appearing—in the accounting of the filling factor—on different footing. As an objection, this may seem aesthetic, but there are some reasons to see it as more substantial. Mainly, consider the following: as a check of the self-consistency of the Landau level assignments to each surface, one can confirm that the carrier densities calculated from the top- and bottom-surface fan diagrams sum to those calculated from the Hall resistance for each gate bias (Fig. 4.9), i.e. whether $\nu_T = \nu_t + \nu_b$.

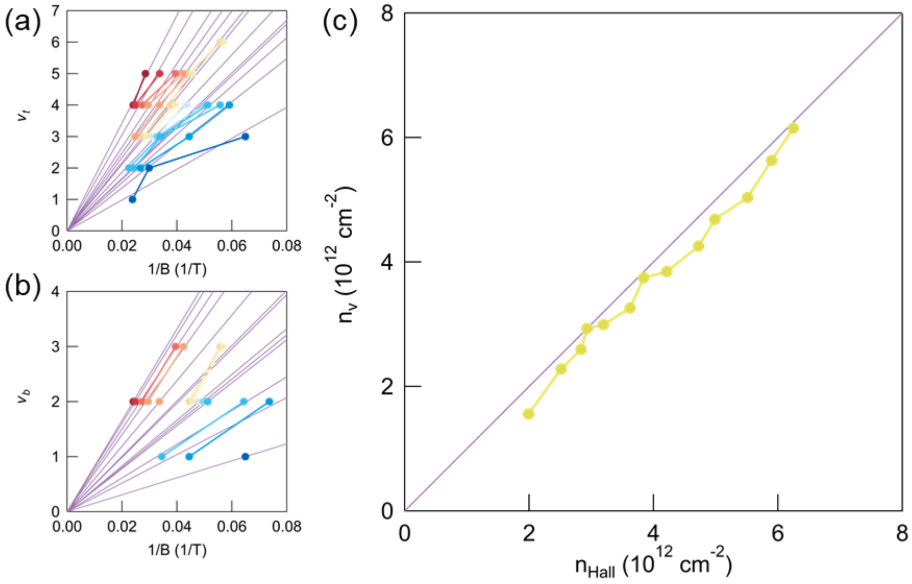

Figure 4.9: Consistency of Landau level indexing scheme with the total carrier density as measured by the Hall resistance. (a) Top-surface and (b) bottom-surface fan diagrams, using the data in Table 4.2, with fits forced to pass through zero. (c) The carrier density, as calculated by summing the slopes of (a) and (b), plotted against the carrier density calculated using the Hall resistance as described below.

Following the scheme in Fig. 4.8 of the main text, we calculate the carrier density by the relation $n_{t,b} = \nu_{t,b}(eB/h)$ and compare the sum $n_T = n_t + n_b$ to the carrier density deduced from the slope of the Hall resistance as described in the previous section. The check has a few

Table 4.1: Top- and bottom-surface filling factor assignments for $\nu_T = 1$ to $\nu_T = 8$ according to Fig. 4.8a.

ν_T	8	7	6	5	4	3	2	1
ν_t	5.5	4.5	4.5	3.5	2.5	2.5	1.5	0.5
ν_b	2.5	2.5	1.5	1.5	1.5	0.5	0.5	0.5

steps. First, we identify and index peaks in the R_{xx} data (cf. Fig. 4.2c). The peak positions (in magnetic field) are found with MATLAB's `findpeaks` and indexed by hand against the corresponding R_{xy} data. These peaks correspond to transitions between filling factors and occur when the Fermi energy is inside of a Landau level. Each peak thus falls between two plateaus; we index each peak with a number corresponding to halfway between the two plateaus. Explicitly, if the peak falls between, say, a plateau at $\nu_T = 3$ and $\nu_T = 5$, we index the peak as 4, while if the peak falls between $\nu_T = 4$ and $\nu_T = 5$, we index the peak as 4.5. We perform this procedure up to about $\nu_T = 10$. Second, we observe that, in the quantum Hall regime, which is to say above about 15 T, the sequences of filling factors in Fig. 4.4 are uniquely identified by the total filling factor ν_T. That is, while it is technically possible to find a region in Fig. 4.8 where, say, $\nu_T = 3$ is achieved by different combinations of top and bottom filling factors ν_t and ν_b, the entire region of interest in the experimental data corresponds—by accident—to a sequence of total filling factors that correspond to exactly one combination of ν_t and ν_b each. In other words, the model assigns only one top and bottom filling factor to each total filling factor (at the carrier density and field that we measure). This assignment is shown in Table 4.2, along with the corresponding peak positions.

Finally, we plot ν_t and ν_b vs. $1/B$ (Fig. 4.9). In such a plot, the slope is (nh/e), where n is

Table 4.2: Since the filling factors are indexed as $\nu = n + \frac{1}{2}$ on each surface, a transition on one surface corresponds to an integer filling factor on that surface. The field sweeps shown are those that demonstrate more than one Landau level transition.

peak index	9	8.5	8	7.5	7	6.5	6	5.5	5	4.5	4	3.5	3	2.5	2	1.5
top index	6	5.5	5	5	5	4.5	4	4	3.5	3	3	2.5	2.5	2	1.5	1
bot. index	3	3	3	2.5	2	2	2	1.5	1.5	1.5	1	1	0.5	0.5	0.5	0.5
sweep no.																
5				35.05			41.48									
6				29.65			40.16									
7			25.33				36.64									
8			23.63				33.91									
9				22.71			29.75			40.42						
10	17.92				22.41			26.69		36.21						
11	17.60				21.91			25.39		34.55						
12						20.77		22.99		31.95						
13							20.07			30.49						
14							19.50			29.53				44.52		
15						15.54		17.94			28.99			41.12		
16						13.58		16.92			22.49			37.34		
17											15.40			33.47		42.00

the carrier density. In this way, we calculate $n_T = n_t + n_b$ for a range of gate voltages, which we can compare to the carrier density calculated from the Hall resistance, n_H (Fig. 4.9c). If we constrain the fits to pass through zero, we achieve dramatically better agreement than if we allow for nonzero intercepts. Since we have already accounted for the Berry phase by insisting on half-integer values for the filling factors on each surface, an intercept of zero corresponds to agreement with the model of topological insulator-like surface states described in the main text. The fact that the agreement is improved by enforcing the zero intercept is a self-consistency check for the model. Noise does not account for the higher disagreement found when the fan diagram is not forced to pass through zero—the relative error in the peak identification is much better than 1% of the magnetic field value. Neither does an erroneous indexing of the Landau levels on the top and bottom surfaces, because (1) there is no discernible pattern in the values of the fitted intercepts and (2) there is no discernible pattern in the sum of the top and bottom intercepts, either. The disagreement may instead be due to an underlying avoided-crossing-type behavior, as in ref. [34], that causes the Landau level dispersion to bend. If the bending is essentially symmetrical around the avoided crossing, the upwards and downwards bending compensate each other, leading to agreement for the total carrier density.

4.2.1 The other picture

The observed sequences of Landau levels cannot originate from bulk-like subbands formed by confinement in 50-nm-thick quantum wells. Using a four-band $k \cdot p$ model (for simplicity, we excluded the surface states; for more on this, see below and Sections 2.2.1 and 4.3.3) with the model parameters of ref. [13], which were chosen to match first-principle calculations, we estimate the subband spacing to be less than 5 meV for both films (see Figs. 4.8b and 4.8c). For comparison, the Zeeman energy increases as at least 15 meV/T, and the cyclotron energy varies as approximately 4 meV/T. From this consideration, it seems impossible to see filling factors ranging from $\nu = 1$ to $\nu = 10$ in a sequence comparable to the data. A more detailed analysis confirms this (see below and Sections 2.2.1 and 4.3.3). Furthermore, the confinement gap is a strongly varying function of the film thickness, and the subbands that form the gap change character three times every 10 nm (see Fig. 4.10). As a result, the conduction and valence subband in-plane masses oscillate with the well thickness, even with relatively small thickness changes, as can be seen from Figs. 4.8b and 4.8c. Thus, we would not expect the same sequence of filling factors in the two devices. Furthermore, the confinement gap depends non-monotonically on the well thickness, as shown in Fig. 4.10a, and we would not expect to see the nearly identical spectra we see in the two devices (see Fig. 4.4a; compare Fig. 4.10b and Fig. 4.10c). Moreover, the in-plane masses of the subbands forming the gap depend sensitively on the well thickness, even for the 45-nm-thick and 50-nm-thick wells shown in Figs. 4.8b and 4.8c. In contrast, signatures of the surface states are expected to be nearly independent of the film thickness, as long as it is large enough to prevent hybridization, which is consistent with these data. We note that there are reasons to expect a bulk gap in these films that is larger than in Fig. 4, for example, due to a small residual film strain [38] or structural inversion asymmetry, but these do not substantially alter the subband spacings. (The picture persists in thinner films, too, as discussed in Section 4.3 below.)

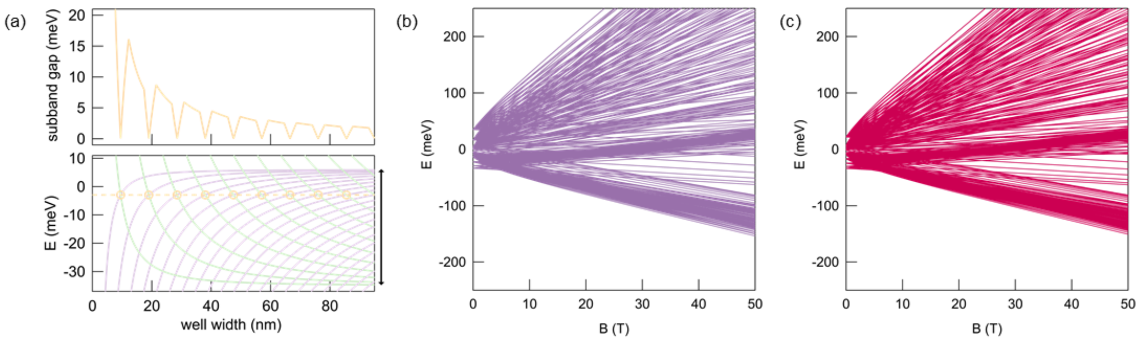

Figure 4.10: (a) Evolution of quantum well confinement gap (upper panel) and subband energy (lower panel) as a function of well thickness. The energy difference between the two bands at Γ that give rise to the Dirac nodes is indicated by the vertical arrow at right. (b) Landau level spectrum for a 45-nm-thick and (c) 50-nm-thick quantum well. The same picture is discussed with more detail in the context of thinner films in Fig. 2.3.

4.2.2 Accidents or clues?

The topological insulator picture does a good job of explaining these data. There are, however, a number of features that fit the picture less satisfactorily:

1. the observed difference in activation energy between the states at even and odd filling factors: while specific sequences of odd versus even plateaus result from the two surfaces and the potential offset with gate voltage, the picture underlying Fig. 4.8 does not result in a strong preference of odd or even plateaus, but in the experiments it seems that the even state filling factors correspond to suppressed activation energies;
2. the absence of a plateau at $\nu = 1$ preceding

3. the highly resistive behavior and breakdown of the quantum Hall effect under large negative bias. This is all the more surprising because we might expect (cf. Fig. 4.8a) that, at $\nu = 1$, the Landau levels are maximally separated, and the plateau is easier to observe, not more elusive.

We speculate that inter-surface Coulomb interactions might be responsible for the departures from the single particle picture. In particular, spatially separated Landau levels of the two surfaces reproduce the scenario in bilayer two-dimensional electron gases of conventional III-V systems [16, 79]. Such systems support a spontaneously coherent phase at $\nu_b = \nu_t = 1/2$, which has been described as a quantum Hall superfluid or exciton condensate. The collapse of the quantum Hall effect just before $\nu = 1$ is reached and the highly resistive state observed here resemble the formation of such a condensate. In 3D TIs, correlation effects can give rise to minima in R_{xx} at even filling factors, which have $\nu_b \neq \nu_t$ [79], and it would be interesting to understand if they can explain the observed difference in activation energies of even and odd plateaus. Measurements that contact the two surfaces separately, which are more challenging and are beyond the scope of this work, should shed light on the nature of this state [39] and should be especially intriguing at parameters corresponding to $\nu = 0$. Such experiments would also allow us to rule out other possible explanations for the highly resistive state, such as electron-hole puddles [63].

Finally, at the risk of repetition, let us contrast the quantum Hall effect of the (001) films with that of (112) oriented films of similar thickness. The quantum Hall effect in (112) films is characterized by a degeneracy factor of two that is lifted in fields above approximately

10 T, leading to quantum Hall plateaus of $\nu = 1, 2, 3, 4, 5, 6$ [67]. Gated magnetotransport and quantum capacitance revealed ambipolar quantum transport, a zero-energy Landau level, and linear dispersion in the 2D states giving rise to the quantum Hall effect [22]. Here, by contrast, states are non-degenerate and integer filling factors of $\nu = 2$ to $\nu = 10$ follow a pattern caused by crossing Landau levels. In the (112)-plane films, the Landau levels do not cross and exhibit a rather conventional fan diagram [22]. Moreover, in the (112)-plane films, $\nu = 1$ and a holelike $\nu = -1$ are readily observed [22], in sharp contrast to what is observed here. The salient difference seems that (112) surfaces host double Fermi arcs, which can give rise to two Fermi pockets near the center of the surface Brillouin zone [36, 35] and thus an even integer Hall effect until their degeneracy is lifted by the magnetic field. In other words, a single (112) surface can give rise to an integer quantum Hall effect with the observed characteristics. In contrast, the (001) transport involves the participation of both surfaces to obtain an integer Hall effect [18]. Although we do not know why the bottom surface does not appear to contribute in an obvious way to the quantum Hall effect of (112) films, this also provides a natural explanation for the absence of the interaction effects, allowing the development of plateaus at $\nu = -1$ and $\nu = 1$.

4.3 Model tests: thinner films and lower fields

This Section of Chapter 4 reproduces and adapts portions of the unpublished manuscript provisionally titled, "The three-dimensional topological insulator picture in (001)-oriented Cd_3As_2 films," which contains work performed in collaboration with my coauthors, Manik Goyal, Tyler Pardue, and Susanne Stemmer.

We refer to transport measurements on four samples, A, B, C, and D. The salient difference between the samples is the thickness of the $\mathrm{Cd_3As_2}$ layer, which is 12 nm for sample A, 14 nm for sample B, 18 nm for sample C, and 24 nm for sample D. The data discussed are summarized in Table 4.3.

Capped (001)-oriented cadmium arsenide films were grown by molecular beam epitaxy and fabricated into gated Hall bar devices. The samples consist of a (100) GaSb substrate, cut 3° toward (111)B, onto which was grown a buffer layer of $\mathrm{In}_x\mathrm{Al}_{1-x}\mathrm{Sb}$, a $\mathrm{Cd_3As_2}$ layer, and finally a thin GaSb cap. Where noted, an $\mathrm{Al_2O_3}$ gate dielectric was deposited *ex situ* using atomic layer deposition after the as-grown devices were first measured. Low-frequency ac Hall measurements were performed using a 1 μA excitation and detected with a lock-in amplifier in a Quantum Design PPMS Dynacool. Raw resistance data were binned and interpolated before being symmetrized (R_{xx}) or antisymmetrized (R_{xy}) with respect to B. The thickness of each sample was determined from cross-sections using transmission electron microscopy (thanks to Tyler Pardue).

Before we compare these four samples, we examine one in detail. Figure 4.11 shows mag-

Table 4.3: Summary of thin film, low-field data for Section 4.3. As discussed in Section 4.3.1, data are fit to $\nu = n_{\mathrm{fan}}(h/eB) + \gamma$ and plotted in Fig. 4.13. Fit error (one standard deviation) for the intercept γ is reported here as $\Delta\gamma$ and discussed below.

sample	thickness (nm)	n_{Hall} (10^{11} cm^{-2})	n_{fan} (10^{11} cm^{-2})	γ	$\Delta\gamma$
A	12	6.42	6.91	−0.061	0.099
B	14	6.92	8.25	−0.217	0.048
		14.0	14.7	0.017	0.562
		15.0	15.4	0.024	0.423
C	18	7.62	6.10	0.012	0.024
		14.5	15.7	0.655	0.308
		15.0	15.8	0.861	0.225
D	24	6.52	6.80	−0.066	0.036
		6.95	7.20	−0.049	0.060
		13.6	14.2	0.795	0.577

netotransport data from sample D. At magnetic fields weaker than about 5 T, the plateaus are weak and proceed according to an apparent degeneracy of two, i.e. the filling factor ν steps from 10 to 8 to 6, as can be seen from panel b, with no hint of other dips in the longitudinal magnetoresistance, R_{xx}, that might reveal the missing filling odd factors (panel a). Around 6 T, the peak in magnetoresistance is, however, clearly split, corresponding to a suppressed (that is, not observed) plateau at $\nu = 5$, and the plateau at $\nu = 3$ is more clearly recorded. The resolution of these odd-numbered plateaus at higher field is in a sense the main feature of these data, and, except for a different background, it is repeated in samples A, B, and C, as discussed below.

The most illuminating comparisons between these samples are made at fixed carrier density. This is illustrated in Figure 4.12, in which different carrier densities are achieved by

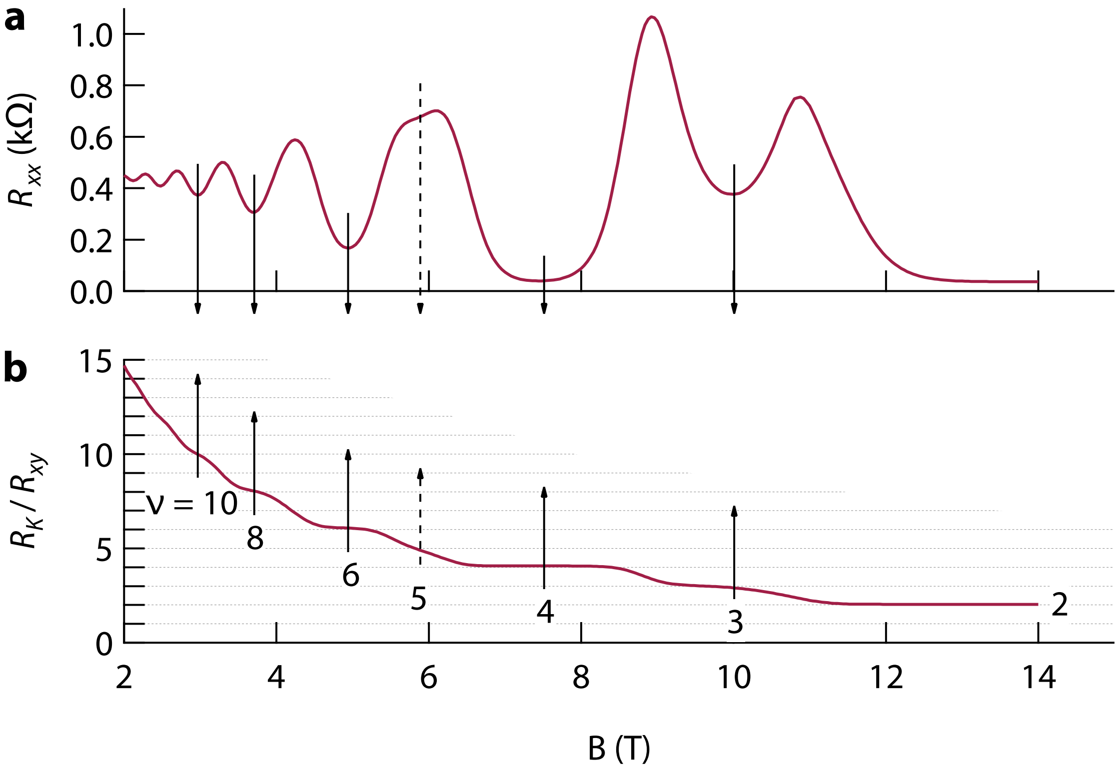

Figure 4.11: The quantum Hall effect in sample D ($n_{\mathrm{Hall}} = 6.95 \times 10^{11}$ cm^{-2}). (a) The longitudinal magnetoresistance, R_{xx}, acquired at 2 K, is plotted against magnetic field. Minima are indicated with arrows, matching those in the lower panel. (b) The von Klitzing constant ($R_K = h/e^2$) is divided by R_{xy}, the Hall resistance, to show that the plateaus match integer filling factors ν, indicated by the dashed lines and labels. Arrows correspond to the minima in R_{xx}, shown in panel (a). Though a plateau does not form with $\nu = 5$, a corresponding minimum in R_{xx} is nevertheless visible (dashed arrows).

adjusting the top gate bias, though some of the traces shown were measured from the films as grown, i.e. without the deposition of a gate dielectric. As grown, the carrier density varies across the samples. It has been demonstrated elsewhere that, in devices such as these, the carrier density and mobility depend strongly on the chemical condition of the sample surface [21]. While the thin GaSb cap, only several atomic monolayers thick, mitigates this, what determines the as-grown carrier density and whether it is extrinsic or intrinsic remain research questions. Here, the as-grown carrier density in samples A and B is nearly the same (about 6.5×10^{11} cm^{-2}), while it is highest in sample C (2.4×10^{12} cm^{-2}), and sample D's falls in between (1.4×10^{12} cm^{-2}).

The Hall data from sample D, shown in Figure 4.12a, demonstrate that the same spectrum is relevant across the range of carrier density shown, which is equal to about 25% the total—there are no qualitative changes. In other words, the evolution of the longitudinal magnetoresistance with the magnetic field, $R_{xx}(B)$ is nearly identical across this range of carrier density. Starting at 14 T and tracing $R_{xx}(B)$ toward $B = 0$, two pairs of peaks in the magnetoresistance are evident, as discussed above in the context of Figure 4.11, resulting in an apparent degeneracy factor of two at low field. As the carrier density differs, so does the shape of the double peak that obscures $\nu = 5$, being essentially a single peak for the lowest-density trace and most clearly two overlapping peaks in the highest-density one. That subtle change represents the sensitivity limit for understanding the data shown in panels b through d: two $R_{xx}(B)$ traces that differ in such a way cannot be understood to result from different underlying spectra. So while in each panel the carrier density has been chosen to be similar, based on the Hall effect at low field, the match is rarely exact.

A comparison between samples A, B, and D is shown in Figure 4.12b, for a density of about 6.5×10^{11} cm^{-2}. Sample D, the thickest, exhibits the longest classical and quantum scattering times, in spite of the fact that the traces for samples A and B in this panel (and panel c) were acquired without a gate deposited, which, as a result of the temperature of the gate deposition, has a tendency to damage the film and reduce the mobility somewhat.

This is reflected mainly in the zero field resistance, which is much higher for the two thinner samples, as well as the onset of the quantum oscillations and the width of the oscillating features. Ignoring the difference in the magnetoresistance background and the broadening of the oscillations, all three traces exhibit the same behavior. Following all three $R_{xx}(B)$ traces from high to low field, a double peak is visible around 11 T, more or less resolved according to the oscillation width, followed by another around 6 T, which in samples A and B is hardly resolved at all in R_{xx}, but slightly clearer in R_{xy}. The sequence of filling factors appears to be identical, and the oscillations match modulo the difference in carrier density.

The same picture is visible in Figures 4.12c and d. Panel c shows the same two traces for samples A and B against two slightly higher density traces for samples C and D. The apparent phase shift of the R_{xx} oscillations in sample C versus those in sample D is clearly due to the difference in carrier density: the same sequence of filling factors is seen in R_{xy}. (See also the discussion of the fan diagrams below.) Sample C, whose thickness is intermediate between that of samples B and D, has a mobility comparable to that of sample D.

Panel d shows traces at high density (approximately 2× that of panels b and c) for samples B, C, and D. (Data for sample D were acquired without the deposition of a gate.) Once again, while the background between the three samples varies significantly, and can be largely attributed to differences in the scattering times that are typically characteristic of the thickness of each sample, the quantum oscillations for each sample reveal the same underlying Landau level spectrum. Here, steps in two of the filling factor are only resolved in R_{xx} in sample C.

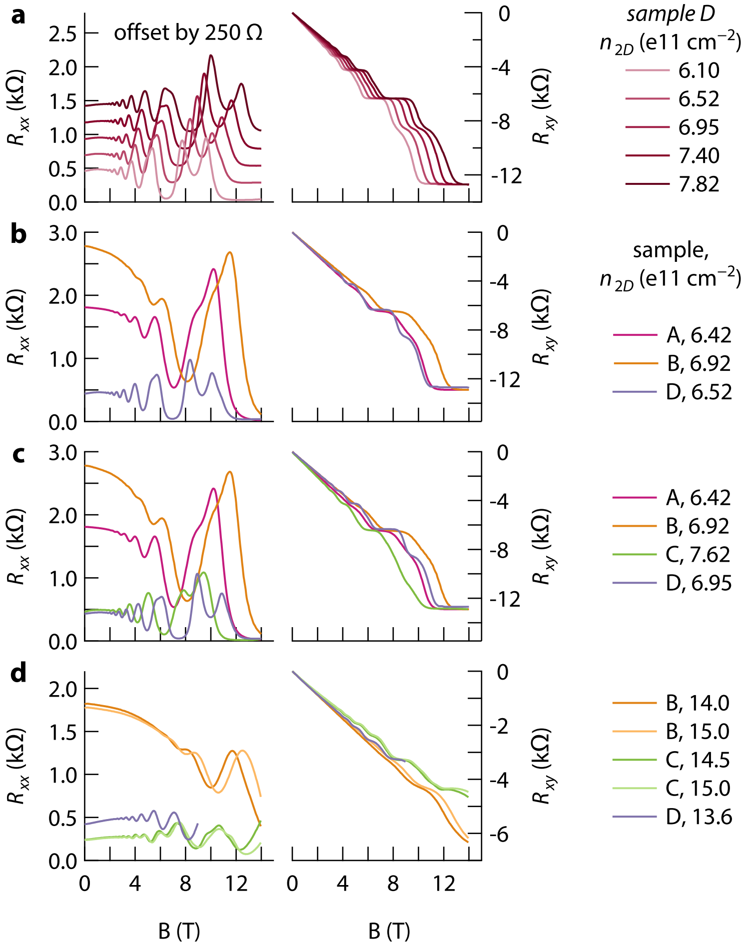

Figure 4.12: A comparison of the quantum Hall effect, measured at 2 K, in samples A, B, C, and D at various values of carrier density. (a) The quantum Hall effect in sample D. Each trace is recorded under a different top gate bias, corresponding to a different carrier density. For legibility, the traces in the longitudinal magnetoresistance, R_{xx}, are offset from each other by sequential multiples of 250 Ω. The true values are all comparable to the lowest resistance trace, which is not offset. (b)-(d) A comparison of the quantum Hall effect in samples A, B, C and D at carrier densities of (b) about 6.5×10^{11} cm^{-2} (traces for samples A and B acquired prior to gate deposition), (c) about 7×10^{11} cm^{-2} (traces for samples A and B acquired prior to gate deposition), and (d) about 1.5×10^{12} cm^{-2} (trace for sample D acquired prior to gate deposition). No offsets have been added in panels (b)-(d).

4.3.1 Fan diagrams

From the R_{xx} data shown in Figure 4.12, we can draw the appropriate fan diagrams. These—organized by sample, not by carrier density—are shown in Figure 4.13, where we have plotted minima in $R_{xx}(B)$, indexed by the concurrent filling factor $\nu = R_K/R_{xy}$, against the inverse of the magnetic field values where they occur. In each panel, we have also plotted a linear fit to these data. For each trace, the carrier density n_{fan} can be calculated from the slope, according to slope $= n_{\text{fan}}/(e/h)$, where e is the magnitude of the electron charge and h is Planck's constant. These values are tabulated in Table 4.3, and agree with those calculated from the Hall effect. To the extent there are discrepancies, they reflect the limitations to using the low-field Hall density apparent already in Figure 4.12.

The intercepts found from these linear fits are mostly consistent with zero, with the exception of the approximately 1.5×10^{12} cm^{-2} traces from samples C and D, as opposed to the low-density traces and the 1.5×10^{12} cm^{-2} traces from sample B. For these traces, the intercept γ seems to lie between $0.5 \leq \gamma \leq 1$, though the error is comparatively large. The fitted values of γ and the fit error are also tabulated in Table 4.3.

4.3.2 Fourier analysis of Shubnikov–de Haas oscillations

One way we can check the 3D TI model against these data, i.e. the reasoning of Section 2.2.3, is by examining the frequency of the quantum oscillations. The oscillating part of the magnetoresistance against $1/B$ has a frequency F that is proportional to the area of the orbit in

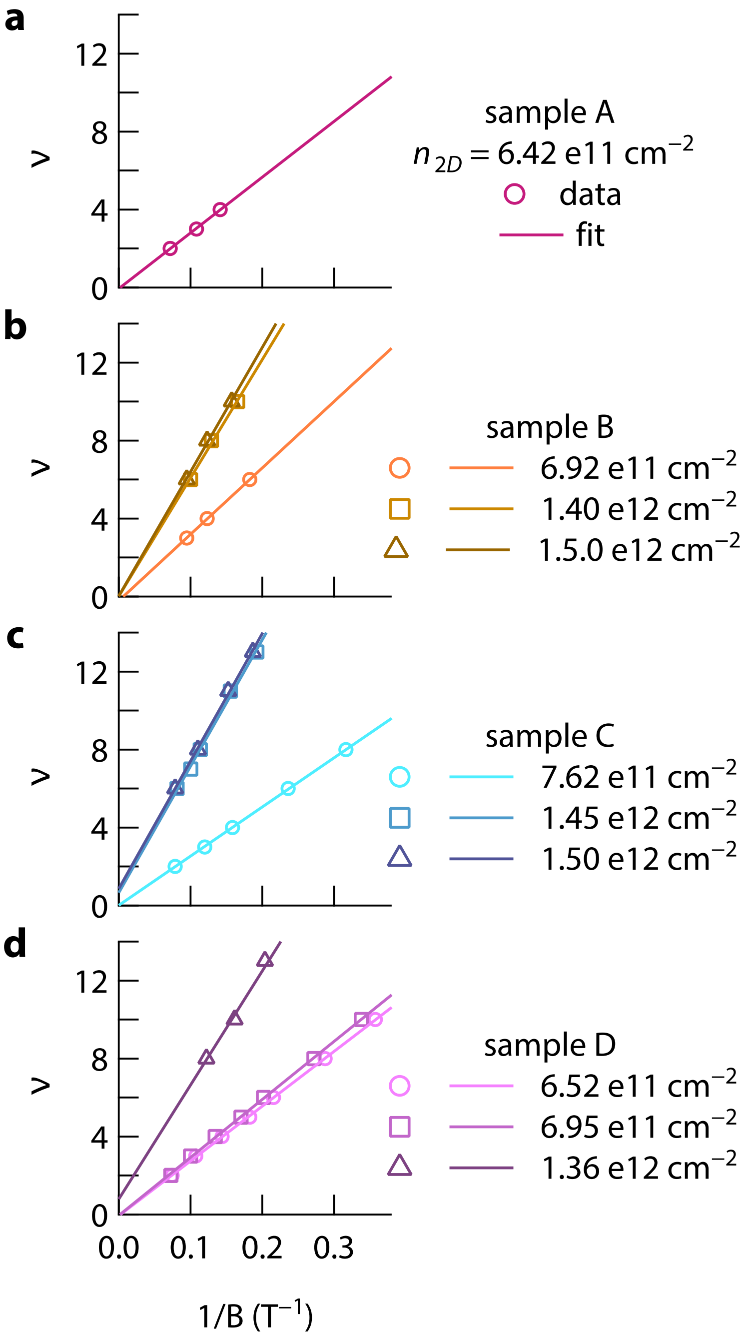

Figure 4.13: Fan diagrams extracted from the magnetoresistance data shown in Fig. 4.12. Each minimum in $R_{xx}(B)$ is indexed by the concurrent value of $\nu = R_K/R_{xy}$, rounded to an integer, which is plotted against the value of the inverse of the magnetic field B where the minimum occurs. Linear fits are shown.

reciprocal space, according to $F = (\hbar/2\pi e)A_k$, where e is the electron charge and A_k is the area of the orbit [71]. A circular orbit, for example, has $A_k = \pi k_F^2$; k_F is the magnitude of the Fermi wavevector. Only extremal orbits contribute. If there are multiple extremal orbits originating with different areas, then multiple frequencies can be visible.

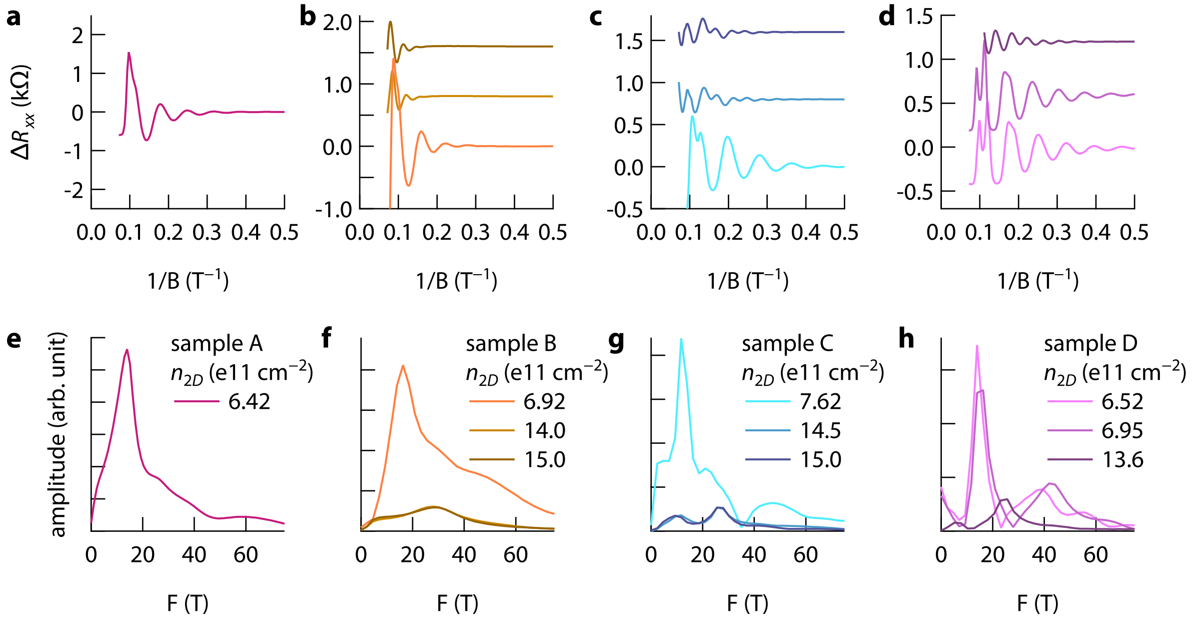

Figure 4.14: Frequency analysis of quantum oscillations. (a)-(d) Oscillations with background subtracted for samples A through D, respectively. (e)-(h) Fourier transform of the background-subtracted data from samples A through D, respectively.

We separate the oscillating part of the magnetoresistance from the slowly varying (classical) magnetoresistance. As can be seen from the raw data (Fig. 4.12) and discussed briefly, the non-oscillating background differs greatly between the four samples. Practically speaking, the best results for the background subtraction are obtained differently for the different backgrounds. In all cases, though, the procedure is to interpolate the raw data on a grid in $1/B$. Then a low-order polynomial is fit with weights to a subset of the interpolated data and then subtracted. The results are shown in Fig. 4.14a-d. The Fourier transform of these traces reveals the frequency components of the oscillations. These are shown in Fig. 4.14e-h.

The Fourier transforms at low carrier density all resemble each other; those at high carrier density are likewise similar to each other. At low carrier density, the Fourier transform for all samples appears to have a single large peak between about 15 and 20 Tesla. A high-frequency peak (40 to 60 T, depending on the sample), is also visible—it is most prominent in sample D (panels d and h). It is not a higher harmonic of the fundamental frequency. Instead, it is due to the resolution of the two fans that appears at high field, which appears as a doubled peak at high field. If the Fourier transform is applied only to the lower-field data—if we window out the high-field doubled peak—the high frequency peak disappears (not shown).

At high carrier density, the Fourier transforms all consist of two comparable-magnitude peaks, at around 10 and 30 T (all samples). The similarity is in spite of the oscillations looking fairly different. A reasonable question is whether the low-frequency (approx. 10 T) peak is spurious, i.e. introduced by an incomplete (or overzealous) background subtraction. The test of the background subtraction is whether the oscillations vary around zero, as it can be seen to in panels a-d. The other check is the number: the Fourier transform of the subtracted polynomial has a low-frequency component if it oscillates (its derivative has zeroes) on the scale of the data. One can estimate that a fourth-order polynomial (the highest degree used here) has at most one full peak or dip in the positive half of the number line. If that were to fall in the range 2 T to 14 T (the plotted and Fourier-transformed range in Fig. 4.14), we would register a peak in the Fourier transform with a maximum of one half period per 12 T, i.e. a frequency in $1/B$ terms corresponding to about 6 T. By contrast, the lowest-frequency peaks seen here, at a frequency of 10 T, would register in the background-subtracted data as having two peaks separated by 0.1 T^{-1}. Both of these factors, the success of the background subtraction and the size of the frequency relative to that characteristic of the background, lead us to conclude that the low frequency peaks measured are not spurious.

4.3.3 Surface states rather than subbands

Recall the subband model introduced in Section 2.2.1. For the range of thickness represented by samples A, B, C, and D, the number of conduction bands relevant to the transport increases from one to two or three (depending on E_F). As a result, the model predicts a commensurate increase in the complexity and/or apparent degeneracy of the Landau level spectrum and sequence of filling factors. The essential feature of the data, however, is no change of the filling factor sequence with thickness: the subband picture cannot be said to agree with the experiment.

Consider now the 3D TI model discussed in Section 2.2.3. The relevant question in applying it to these data is how the Landau level spectrum, which depends on Δ_h, Δ_i, and the carrier density n, should evolve under the influence of experimental parameters like thickness and gate voltage, which do not appear in the model Hamiltonian (eq. 2.4). The term Δ_h, which couples the two surfaces, is relevant when there is appreciable spatial overlap between the states on each surface, which we expect occurs only in very thin films. Heuristically, if the length scale for the Dirac state goes as $\hbar v_F/\Delta$, a Fermi velocity v_F of 8×10^5 m/s and a gap Δ of 100 meV suggest that hybridization of the surface states should occur in films thinner than about 6 nm. The hybridization gap Δ_h, accordingly, should be negligible for films thicker than that. In other words, the Δ_h may be a strong function of film thickness when the film is only a few nanometers thick, but, in the regime studied here, Δ_h is small and unchanging as the thickness is varied. In addition, as discussed in the exposition of Figure 2.5e, the Landau level spectrum is essentially insensitive to modest changes in Δ_h as long as the Fermi energy lies outside of the gap, meaning that, even if Δ_h did vary substantially for films 12 to 24 nm thick, our experiments would likely not detect its influence.

The Fourier transform gives us one consistency check between the TI surface state model (Section 2.2.3) and our data. In the model, finite values of Δ_i should result in two frequencies for quantum oscillations, whose difference increases as a function of increasing E_F or carrier density. This can be seen heuristically by considering the case where $\Delta_h = 0$ and $\Delta_i > 0$. Then the dispersion looks like two offset Dirac cones (see Fig. 2.5b). This results in two extremal orbits: one around the higher-energy Dirac cone, which has a smaller radius, and the other around the lower-energy cone, which has a larger radius. Because the dispersion is linear, the difference in radius is constant. But, if the radii are $k_0 - \Delta k$ and $k_0 + \Delta k$, then the difference in area is $\Delta A = 4\pi k_0 \Delta k$. Since $k_0 \propto E_F$, as E_F increases, it is clear that $\Delta A \propto \Delta F$ increases, where ΔF is the difference in quantum oscillation frequency. Using the dispersion relation in Eqn. 2.5, with $\Delta_h = 0$, one can calculate that the difference between the two frequencies for quantum oscillations is

$$\Delta F = \frac{E_F \Delta_i}{e\hbar v_F^2}. \tag{4.1}$$

As a sanity check, note that a difference in frequency of 20 T, suggests that the quantity $E_F \Delta_i \approx (60\text{ meV})^2$, assuming $v_F = 5 \times 10^5$ m/s. This is indeed what is observed in Fig. 4.14.

4.4 Some pessimism (and a little optimism) regarding rotating the magnetic field

The TI model makes fairly straightforward predictions in a rotating field. What's more, there is a clear rationale for pursuing this line of inquiry. Ultimately, the odd-integer quantum Hall effect, which has never really been observed as such [93]. In the language of the model, this is because $\Delta_i = 0$ is not enforced in a real sample. (The contrasting case is in graphene, where the spin and valley degeneracies are, under the right circumstances, an exact degeneracy. Here, the degeneracy of the top versus bottom surface rests on the presence of inversion symmetry—we can call τ_z an inversion operator—which is not a symmetry of the system.) Without the odd-integer quantum Hall effect as a smoking gun for the TI, a reasonable replacement seems to be the rotation experiment, which, as discussed below, is recommended by the distinctive field dependence of the zeroth Landau level.

Leaving questions of the model's applicability aside, there are still substantial challenges to drawing quantitative conclusions from the quantum Hall effect in a rotating field, even just to discriminate a Dirac dispersion versus a parabolic one. One wishes to observe the crossing of the lowest Landau level, which disperses linearly, as $|B|$, and higher ones, which additionally depend on $|B|\cos\theta$. But, under normal circumstances, this only occurs at low filling factors, and, at correspondingly low density, there are few crossings that can be observed. Since one hopes to plot the locations of the crossings to discern the distinctive zeroth Landau level, one needs at least three such crossings, which typically requires extremely large fields.

The other issue is the broadening of the quantum oscillations, which can obscure the location of the true crossing. If the errors were random in identifying the critical fields and angles, the problem could be solved by fitting the appropriate function to enough data points. In sum, then, the need to find crossings doubly exceeds our expectations for doing so.

The large amount of data we have, as a result, rises only to the level of suggestion. We summarize our inconclusive results to finish this chapter.

Figure 4.15 shows rotated-field data for the sample studied in Fig. 4.3. The rotation axis is parallel to the current path, and the angle θ is measured between $\vec{B}$ and the film normal, [001]. The top row shows one (low) carrier density; the bottom shows a slightly higher one. What is heartening is that the rotation does appear to generate a number of crossings. What is disheartening, however, is that they mainly appear in the higher-density plot, even out to 45 T. Considering that there are at least four model parameters that play a role in determining the crossings, two data points, even if we had a dense enough grid of angles, are hardly enough to check them.

Figures 4.16 and 4.17 show data for a different (capped) film with no usable gate. The data should be more finely spaced and the R_{xx} have sharper features in order to more clearly identify the coincidence angles. That said, a number of transitions are visible. In Fig. 4.15b and 4.15c, the emergence of a plateau at $\nu = 4$ is clearly visible under rotation; in Fig. 4.15e and 4.15f, a plateau at $\nu = 5$ disappears and one at $\nu = 6$ appears under rotation. In Fig. 4.17, at higher density, the changes are less dramatic (most visible is the emergence of a plateau at $\nu = 9$) but clearly the field dependence of the quantum oscillations is not given by only the

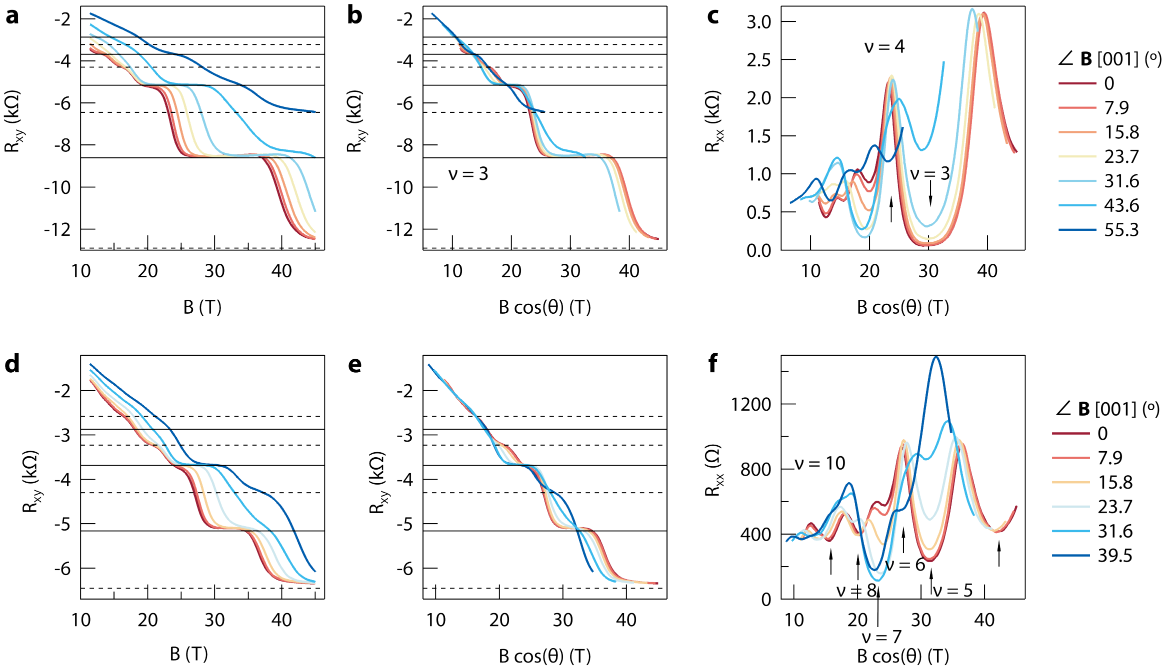

Figure 4.15: Rotation experiment in a Hall bar at high field. The rotation axis is parallel to the current path. Landau levels cross, but the data are not densely spaced enough in angle to identify where, limiting the scope of our conclusions. (Prior to these field sweeps, the sample was damaged slightly by aggressive testing with the gate. Though this is the same sample as in Fig. 4.3, the quantization is slightly worse following the gate testing.) (a) R_{xy} shown at low density. Seven field sweeps at different angles are plotted against field, clearly showing that a plateau emerges at $\nu = 4$ as the sample is rotated. (b) The same data as (a), plotted against the out-of-plane component of $\vec{B}$. (c) Corresponding R_{xx} data for the same sweeps, with dips corresponding to the plateaus at $\nu = 3$ and $\nu = 4$ labeled. The legend at right goes for panels (a)-(c). (d)-(f) The same plots as (a)-(c) at higher carrier density (lower gate bias); the angles are slightly different. More crossings are visible, particularly at $\nu = 5, 6, 7, 8, 10$, as indicated in panel (f).

out-of-plane component of B. Further developing these experiments as a more quantitative probe of the 2D transport in these films is, in spite of the challenges outlined above, an exciting direction for future study.

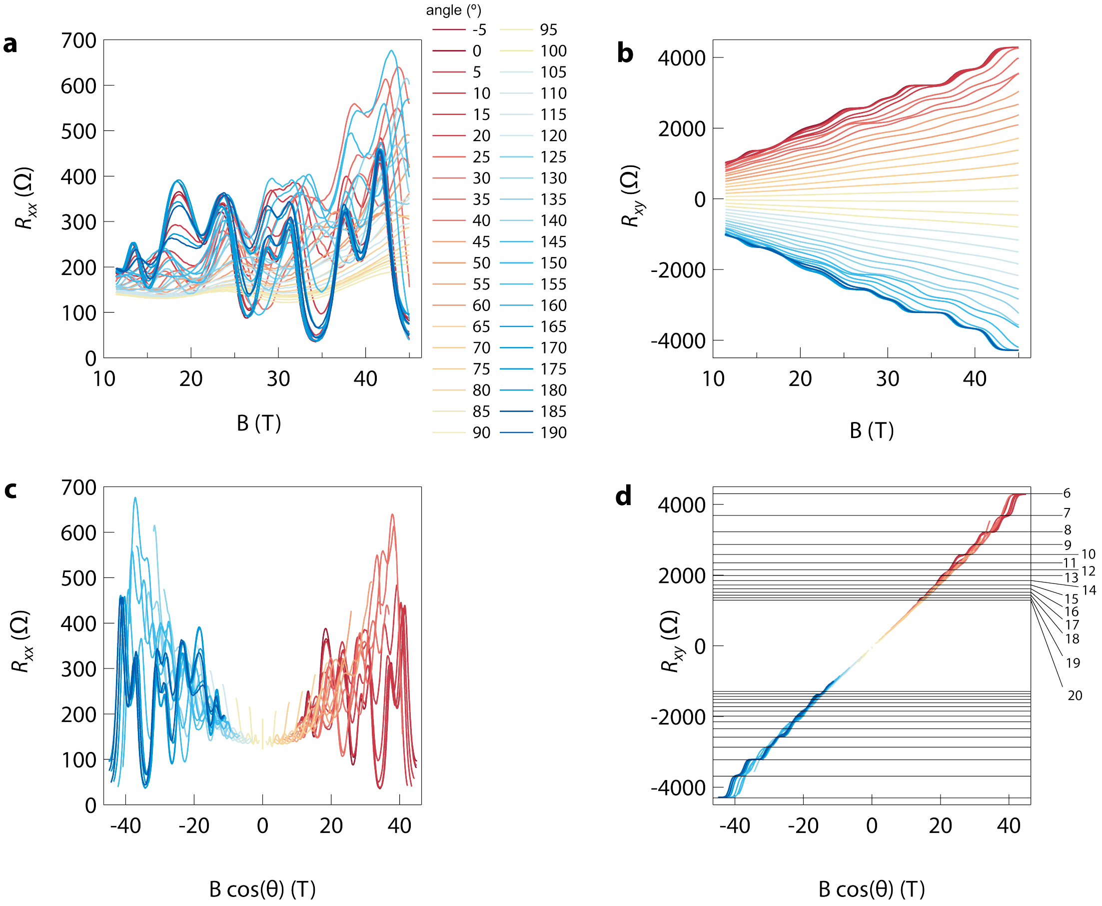

Figure 4.16: Rotation experiment in a Hall bar at high field, with many more angles measured. This (001) sample is capped (cf. the results of Section 4.3). As in Fig. 4.15, the current is injected along the rotation axis. (a) R_{xx} and (b) R_{xy} versus B; (c) and (d) are the same as (a) and (b) but plotted against $B\cos\theta$, θ defined as above, the angle between the film normal and B. Many of the same phenomena are visible as an Fig. 4.15. The gate was unresponsive on this sample, and the carrier density is high; the Hall bar reaches $\nu = 6$ at 45 T.

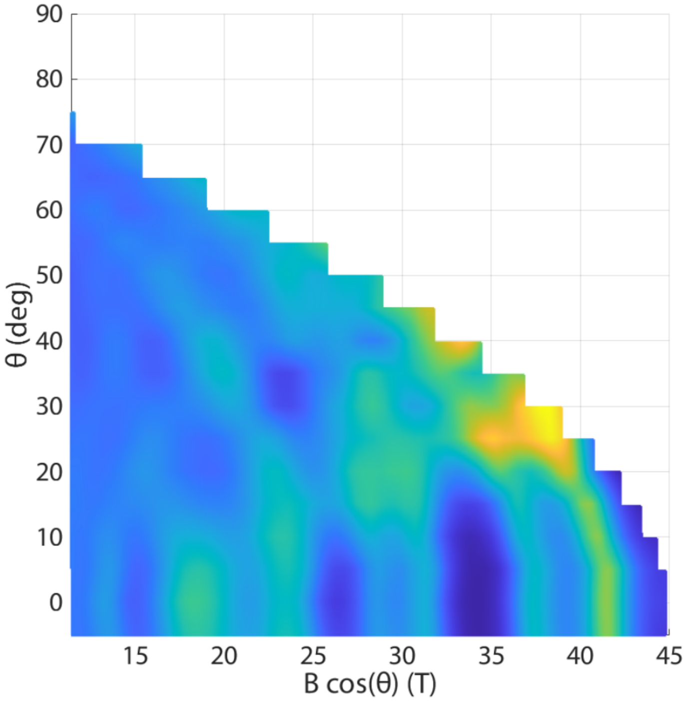

Figure 4.17: Another look at the data in Fig. 4.16c, i.e. R_{xx} (colored) vs. $B\cos\theta$ and θ. Blue is low, and yellow is high. The angle spacing still limits the resolution—features appear more smeared out in angle than the five degree increment—and so does the width of the oscillations. Consider that a crossing can be unambiguously determined by a low-high-low trajectory of R_{xx} along θ (which can be checked against R_{xy}) and that all such features here are broadened over more than ten degrees and/or 5 T.

Chapter 5

Future directions

In this last Chapter, we outline a few challenges for the future in (001) films. The main challenges are related to growth and device processing. While some iteration is sure to progress these two, and particularly the latter, we restrict our attention to suggesting avenues for researech that differ from those currently being undertaken.

5.1 Frontiers for growth

The three main frontiers for growth, beyond the development of a suitable lattice-matched substrate (such as an alloy of CdTe and ZnTe, or even large crystals of Cd_3As_2), are the development of (1) a back-gated structure, (2) capping materials, and (3) exciting heterostructures.

5.1.1 Back gating

Back gates provide a valuable experimental control for testing the TI picture discussed in Section 2.2.3 and will be critical for realizing the quantum spin Hall insulator if it is possible. On top of that, they are critical to expanding the possibilities of the capacitance measurement detailed in Section A.

The AlGaSbAs buffer layer in current use (see Section 3.4) is in principle resistive enough to act as a back gate dielectric, but reports of devices using it effectively as such are scarce Our limited testing (less than five growths) suggests that the layer is prohibitively leaky. But, given the well-known tendency of the III–Sb compounds to form conductive oxides, it is still an open question whether our structures are limited by the growth (extended defects, percolation through the alloy) or by device processing (sidewall conduction). Perhaps the most likely answer is both, but the picture would be clearer after measuring the area versus perimeter dependence of the leakage current for mesas of various sizes.

On the side of growth, there are certainly problems to solve, many of which have already been answered by others. But a suitable growth recipe needs to be developed for a doped layer close to the top layer, with the diffusion of the dopants limited so that there is a sufficiently insulating layer in between the doped layer, to be used as a back contact, and the film. Hard horizontal barriers can perhaps be engineered using a digital-alloying approach to deter percolation and extended defects from traveling (vertically) through the structure. The InAs literature is, doubtless, a valuable resource for this project.

5.1.2 Capping

The addition of a capping layer grown by migration-enhanced epitaxy (a shuttered growth technique that can extend the window for good growth down to 200 °C or, one hopes, lower) has been a breakthrough in our understanding of the mobility-reducing mechanisms in Cd_3As_2, as well as increased the reproducibility of our highest-quality films to the point that consistent comparisons can be drawn from very low numbers of samples. The tradeoff has been, however, that the gate response of the gated samples is much worse. (Obviously, a back gate would remedy some of this.) There are two causes. One is the dielectric constant, which can benefit from only so much engineering. The other is the "quality," which can stand much improvement. Progress on this will almost certainly require improvements to the low-temperature thermometry and careful flux calibration. A perfect, epitaxial cap would enable, besides improved mobilities (preliminary data suggest that world-record thin film mobility can be achieved with a thicker cap), growth of more complicated heterostructures with multiple Cd_3As_2 layers.

5.1.3 Heterostructures

Diverse heterostructures are possible. The idea in general is to use the well-known III–V platform as a probe for studying the Cd_3As_2 layer. Here are three such proposals:

- Heterostructures such as the ones mentioned in the previous section represent completely unexplored terrain for Cd_3As_2 growth. More control can be achieved by combining a back-gated structure with such a structure.

- An less ambitious direction involves the use of bandstructure engineering and coupled quantum wells to develop more quantitative understandings of the properties of the so-called topological heterostructure.
- Further, the condensation of bosons in coupled quantum well systems provides a blueprint for exploring correlations in quantum wells (such as InAs, which is grown as part of one of our buffer structures) with Cd_3As_2 layers.

5.2 Zooming out a little

Perhaps an obvious source of excitement is the prospect of realizing topological superconductivity in a proximitized Cd_3As_2 film. While there are obviously technical challenges for device processing and optimization, it seems like the largest challenge will be establishing—as a field—a consensus on the signatures of the different topological phases that achieved in the absence of superconductivity. One way to achieve this is by marrying spectroscopy to electronic transport, providing an opportunity to synthesize understanding of the last years' sometimes seemingly disparate results (some are reviewed in ref. [15]).

One is also encouraged by the recent work in our group and elsewhere to understand what kinds of crystalline symmetries can be broken by epitaxial strain in Cd_3As_2. These studies, combined with band structure calculations, are a key ingredient in engineering topological phases in Cd_3As_2 beyond the 3D topological insulator.

Appendix A

Capacitance measurements in the Dynacool

The work in this Appendix was performed with the substantial assistance of Daniel Ohara and Paola Perez in designing, drawing, and constructing the probe. I am also grateful for the generous guidance of Alexey Suslov, who showed me the artful rf probes of his own laboratory, which have been influential in the design of our second-generation probe. The original idea for this work and the first design of this probe, along with just about everything else I know about probe construction, comes from Robert Kealhofer and his patient help.

The basic premise of the capacitance measurement is that, in a capacitor where one plate is, say, a Cd_3As_2 film, the dielectric some deposited Al_2O_3, and the other plate a metal electrode, the total capacitance is the capacitance that the geometry provides less a series contribution proportional to the density of states in the Cd_3As_2 film [50]. While the resistance also depends on the density of states, it is usually obscured by scattering processes, making capacitance a relatively more direct method for measuring the density of states. It has also been claimed to be, in topological insulators, more sensitive to one surface compared to the other, which suggests possibilities for experiments that might elucidate some of the complicated spectra observed in, say, Fig. 4.4 [42].

In this Appendix, we discuss a method for measuring capacitance in the Quantum Design Dynacool PPMS and present some preliminary data that, though promising, is marred by thermalization problems with the first version of the probe, which can be attributed to damage prior to loading.

A.1 Primary considerations

When measuring the resistance of a, say, Hall bar, it is common knowledge that one can avoid measuring the resistance of the leads by employing a so-called four-point measurement, where the current is injected across two leads, and the resulting voltage drop is measured across two different leads. Since no appreciable amount of current flows across the voltage leads, the potential difference is not reduced by their resistance—in fact, it is ideal if the voltage leads have as high of resistance as possible.

Such a four-point measurement is impossible for a capacitor. But perhaps, one might hope, the capacitance of the lines is negligible. If one proposes to measure the capacitance of a capacitor in the PPMS Dynacool, however, one must overcome the fact that the lines are organized in twisted pairs. (For resistance measurements, twisting pairs is a low-tech way of accomplishing “common-mode rejection” of random variations in magnetic field. Your leads form a giant loop, and emf is induced like $\mathrm{d}\phi_B/\mathrm{d}t$ from power lines, vibration, and the rest of the outside world. Much can be rejected by twisting the wires into pairs, which gives one negative loop area per one positive loop area, by one way of looking at it, or induces currents, approximately equal and of opposite direction, by the other way of looking at it.) Twisted pairs have a large capacitance per unit length. So the capacitance of the lines is a problem.

It’s not much fixed by the hope that the true quantum capacitance will, in a magnetic field, vary as $1/B$, unlike the line capacitance, because one ends up needing to measure a small change in capacitance on top of a giant capacitance from the lines. The clever solution is to use a virtual ground to zero the parallel capacitance to ground, which allows one to isolate the capacitance of interest; this requires coaxial cables. (For a synopsis, see ref. [73].)

Since there are not coaxial cables on the PPMS, a simple solution is to route your own down to the bottom. The main design issue one faces is that this increases the heat load substantially, which will necessitate some design tradeoffs. We allude to these in the next section, where we simply describe our probe design.

A.2 Shortcuts

Jack Ekin’s book, its appendix particularly, is an invaluable resource for probe construction [17].

The cryostat of the PPMS Dynacool is made of stainless steel, except for the last four inches, which are made of copper. Our probe matches this, consisting of a stainless steel tube to which is attached a copper sample platform. A printed circuit board is attached to the platform with sockets for a standard 8 DIP connector, allowing the sample to be bonded to the DIP carrier and loaded without complicated manipulations of the four-foot-long probe. In the second version of the probe, the sample is shielded by sitting in a nearly-sealed copper cup to which are soldered SMA jacks; in the original version, an unshielded line is patched from the coaxial cable ends to the PCB. The probe end is attached to a "universal puck," available from Quantum Design (this is what was broken, causing the thermalization issue in Fig. A.1). The top of the probe is based on a KF cross. Hermetic SMA feedthroughs are mounted on blank KF flanges that are isolated from the rest of the probe by a plastic centering ring and clamp. The seal between the tube and probe head is accomplished with a Swagelok fitting.

A.3 Preliminary data (first probe)

Before loading the first probe, the universal puck was separated from the probe end; the probe was loaded, but the sample was not in firm contact with the bottom of the cryostat, meaning that the sample temperature was unknown and likely several degrees higher than the cryostat temperature.

The data show clear quantum oscillations, with two frequencies, as the magnetic field is swept. The double dip feature in the zero-field sweep (Fig. A.1a) is intriguing and does not match the CV measurement on the (112) surface [22].

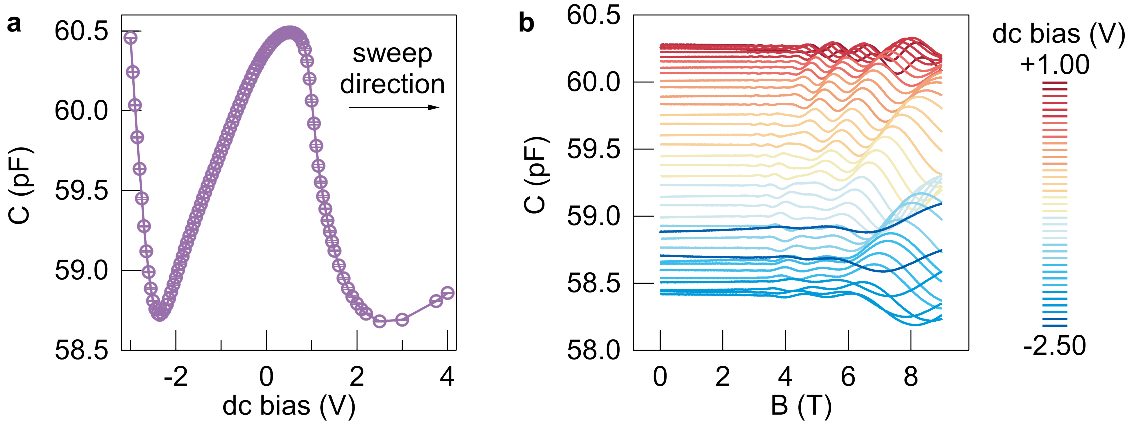

Figure A.1: Capacitance measurements using the first probe. (a) The dc bias is swept with a 25 mV ac test excitation. The geometric capacitance is not subtracted. (b) The dc bias is stepped in between field sweeps. The traces are not offset, and the geometric capacitance is not subtracted.

Appendix B

My devices don’t work

This Appendix is directed particularly to beginning graduate students, who I hope can forgive the obviousness of some of these suggestions. But the guiding principle is to include some practices, which, though they may not be the best way to proceed, are ones I wish I had known when I started growing films.

B.1 Where did I go wrong?

Be alive to the possibility that you could have gone wrong from before you began your experiment. I am bound, here, to write something general—but you do not have that luxury. So, while I might write that *Your devices don’t work*, that’s not to say that you should frame your research question that way, i.e. work vs. not work. In other words, plan out your research question so that you consider the possible outcomes in detail, and, in the spirit of the scientific method, ask yourself what you will learn if the different possible results come to pass. More to the point, don't plan an experiment whose only outcomes are Glory or Nothing—the outcome will always be Nothing. When it comes to Glory, the process by which you find it will not involve The first time, That thing you don't quite understand not being important, or Getting lucky.

OK, but you're sitting at the cryostat and things are going badly with your Cd_3As_2 film. Where did you go wrong? This Appendix is falsely advertised because it's probably impossible to debug your problem once and for all. But the following sections contain some reminders for how the growth should work, a couple suggestions for how to keep your calibrations current, and a baseline suggestion for some cleanroom practices that have had acceptable yields, which maybe you can use, too.

B.2 Temperature

Check Section 3.3.2, which describes some of the main issues with the thermometry.

If your III–V layer is bad, it is almost always the temperature at fault. I say that because, unless you're growing InSb, the growth isn't that sensitive to the flux ratio. That said, check the BFMs first, because it's easy to. If the BFMs look good ($<$ 10% different from normal; growth temperature differences of a couple of degrees at most), start in on the temperature. (The main reason they would not look good: the "pipe" in the Sb cell is clogged.)

Next check that the pyrometer is well aligned to the window, and that the emissivity is set to 1 ("uncorrected"). There is actually an onboard memory for the pyrometer that stores the last emissivity setting, so if you sent the pyrometer away on loan to another system, it's possible that it could be a surprising number. If the pyrometer is not occluded by the viewport, and the emissivity is 1, proceed to testing whether the viewport is the problem.

The fastest version is to check the oxide off temperature and then clean the pyrometer viewport by baking it out, and then check the oxide off temperature again. Suppose you're growing GaSb. Desorb the oxide manually, by first setting the Sb flux to a few $\times 10^{-7}$ torr BEP, something like 525 °C, and heating up the substrate; you can go fast at first, but slow down for the last 100 °C or so. Start RHEED and look for the oxide to come off (bright spots or streaks).

If the oxide off temperature is good to start with, then there's likely no serious problem with the thermometry. If the oxide off temperature is off by more than a few degrees, then continue to baking off the viewport. This is a tricky business, because the feedback on too cold is nothing and the feedback on too hot is venting the system. Using the dedicated viewport heater-cleaner (a cartridge heater wearing a hat and wrapped in a burrito) and a Variac set to the mid-20s, bake the window all night. (A real boon here would be some vacuum testing of an old viewport. Even better would be a re-design of the whole viewport. There is a scheme for doing a hot "screen," but you would need (1) a custom feedthrough and (2) a clever solution to the problem that the deposition tends to short the heater element. I am optimistic that two sapphire disks clamped on either side of a sturdy piece of wire would be up to the task, but you would need to figure out the best way to relieve strain and support the assembly besides actually testing the maximum temperature.) In the morning, turn off the heater and wait for it to cool. Check your oxide off temperature again.

If your oxide off temperature is still not normal, you have options. One is to simply adjust your recipes relative to the oxide off temperature. The best way to do this is with a couple of calibration points. I suggest using the (3×1)–(5×2) transition under Sb flux, particularly if you have a good baseline on the normal pyrometer temperature at which this occurs (given your Sb flux!), alongside the oxide off temperature. What's advantageous about this is that the only resource it really uses is your patience. With two temperature calibration points by which to set your growth that you can check during the first post-desorption planarization step, you should be able to continue growing without opening the system, your other option.

What you may notice, however, is that a good flux calibration can save your bad thermometry. Read on.

B.3 Flux calibration

This is actually pretty fun the first time, and it's not a bad introduction to MBE for a new student. Also, since it gets less fun each time, there's a lot of incentive to get others to help you out with it—maybe that's how you'll get someone their first growths. But this is also perhaps the best example of a task that lets you get to know the system—a phrase you hear probably too much among growers—because it gives you a lot of information that can be checked in different ways. I think this sort of check makes sense to be done every couple months, and after openings. It's also a thing to check (just do it for Ga) if things seem to be different on the system. I usually budget a few hours for each element (just about all active time). (Ordinary BFMs, which give you BEP vs. source temperature, can be run every week under normal conditions. More is overkill, less is risky.)

I haven't cited anything in this section because it constitutes common practice and some kind of basic MBE knowledge—and yet it's difficult to find in the scientific literature. What is written here I learned from groupmates as well as Faebian Bastiman's excellent, though no longer updated, MBE blog, which can be accessed at `https://faebianbastiman.wordpress.com`. If that link is broken, you can surely find an archived version such as this one: `https://web.archive.org/web/20201024054645/https://faebianbastiman.wordpress.com/`. So, while I've tried to detail my own practices in what lies below, I recommend reading this blog to every new student and MBE enthusiast.

B.3.1 Group III sources

On the whole the group III sources are fairly forgiving and work as designed. They are all melts, so the exposed surface area is constant, at least for a month at a time, and so the flux is very stable. As an aside, they are unforgiving at their freezing point: freezing Al will almost definitely destroy the crucible because of how it adheres to PBN (have a replacement handy if you are opening the system—order this well in advance); freezing Ga can, so keep it hot if possible, but it supercools readily, and the melting point is less than 40 °C, so it's rarely an issue except during long openings; freezing In is kind of in the middle—the melting point is under 300 °C—and to be avoided if possible; unlike, Al, though, it doesn't wet the crucible effectively.

Calibrating the flux for Ga, In, and Al is all manageable on a GaSb(001) substrate—the same one, if you're careful enough. The desired calibration is a conversion between BEP and absolute flux (atoms/cm^2/s). This is accomplished by monitoring RHEED oscillations. The cartoon version of how this works is that, as single layers form by adatoms forming islands and coalescing, the sample surface gets rougher, roughest at half coverage, and smoother, smoothest at full coverage. This causes the reflected intensity of the RHEED beam to oscillate with a period equal to the number of atomic monolayers deposited per second. You're also roughening the surface slowly as you go, because at some other rate, lower than the formation of monolayers, you're growing two layers at a time, or, in other words, atoms are being deposited on the partial monolayers. This causes a decaying envelope on top of the oscillations; eventually the oscillations cease. Since you're under continuous, excess group V flux, you can close the group III shutter and things will smooth out a little. But after a while it usually becomes necessary to smooth things out by increasing the temperature and growing "step-flow," which does a better job of planarizing the growing surface.

Starting with a flat GaSb layer (planarize your desorbed substrate as normal), reduce the temperature by 40 °C or so, and open the Ga shutter while monitoring the specular intensity vs. time. A good-enough number of oscillations is anything more than 10. When the oscillations become fainter, close the shutter and wait a couple minutes. Take the time to count the number of oscillations. Each full period is a single atomic monolayer; there are two per unit cell; you can look up the area per unit cell (a^2), and you know how many Ga atoms because they make up an fcc lattice. Measure again if you like, or change the Ga cell temperature and do it all again. You'll only need four or five data points (Ga cell temperatures) to get a good calibration.

If the oscillations weaken, you either need to change the substrate temperature or the Sb flux. The way to find the right temperature is to start at your step-flow growth temperature and reduce (no Ga flux) until you start seeing oscillations when you open the shutter. Don't grow too much because you are slowly making the surface rougher. (It's reversible—you can always go back up to the step-flow growth temperature and grow a little to smooth things back out.) The Sb flux is important because, to some approximation, increasing Sb flux beyond what is necessary for the growth limits the surface diffusion length of the Ga atoms—they bump into the Sb ones—which can suppress the layer-by-layer growth you seek. We've found that the right number changes a little, but a decent guess for the Sb BEP is about 3× the Ga BEP.

Once you're happy with the Ga calibration, you can do the same for Al, by growing AlSb layer by layer, which will have a slightly different best temperature. Note that if you grow a strained layer (monitor the linescan), you have the same planar density of Al atoms as Ga atoms in GaSb.

The same procedure is possible for InAs, but, if you haven't worked out the best way to get a clean, abrupt interface with the switchover of Sb to As flux (I haven't), you will typically need to grow a little InAs on the GaSb to have a smooth starting surface before growing. Once again, keep track of whether you're growing a coherently strained film.

B.3.2 Group V sources

First, figure out why you need group V flux calibration. I think the most compelling reason is to back up the thermometry. That means you are trying to get a correspondence between, say, the Sb flux BEP and the pyrometer reading for later debugging of one against the other. So that suggests you map the static reconstructions (no growth) of a smooth GaSb surface against the pyrometer, recording Sb flux BEP. As long as the Sb flux BEP versus absolute flux (unmeasured) is stable, you can avoid the following.

But suppose you want to know your absolute flux, as for the group III source. (Keep in mind that the relative incorporation of As vs. Sb in an alloy has its own temperature dependence, and iterating growth with feedback from RSMs might be the easier task—you can run the RSM at night and sleep while you calibrate.) The way to get absolute Sb flux is to match the flux of the Sb to that of the Ga by ramping up the Ga until the growth is limited by the Sb supply. Start with a known (absolute) Ga flux and excess Sb. (Note that you won't be in the ideal conditions for layer-by-layer growth, particularly if you are calibrating at a substrate temperature more relevant to your alloy growth than GaSb growth.) Monitor the growth rate from the RHEED oscillations. Close the Ga shutter, planarize if you need to, and then continue. Increase the Ga flux, and check the growth rate again and again.

Eventually you will find that the growth rate no longer tracks the Ga flux because you're now limited by the Sb flux. Unfortunately, you can only grow thin layers in Ga-rich conditions before you irreversibly ruin the substrate. So go slow, particularly around 1:1 (though you might not have a good guess of where that is), and be ready to close the shutter if the RHEED pattern changes.

Note as well that, unlike in the group III calibration, which amounts to fitting a line to relate BEP to absolute flux, here you are finding a point. There, if you monitor for long enough, you can have very low errors on the growth rates, and also in determining the final relationship—usually you will find that you're bumping up against the temperature stability (0.1 °C) of the source. But here, your error in finding 1:1 Ga:Sb is mostly set by the spacing of the Ga flux you choose near the corner, i.e. where you cross over to Ga-rich conditions. In principle, you could zero in on the point iteratively, but that's assuming you don't at some point ruin your substrate with Ga droplets.

The same procedure can be performed for As on a GaAs substrate, but take care, before starting, to get a good practical understanding of the As valve. First, the flux as a function of valve position is roughly sigmoid: there's a region from about 25% open to 75% open that is linear, and the flux isn't that responsive outside of that. So if the flux you want corresponds to a valve position that is nearly closed, reduce the crucible temperature so that you have real control of the flux with the valve. (Do this the day or at least two hours before your measurement: the crucible contains a huge amount of arsenic and takes a long time to stabilize.) Second, open the shutter and open the valve halfway, and see how long it takes the BFM gauge to read consistently (consider using the pressure logger and leaving for twenty minutes). Note the pressure in the growth chamber when you're done. Then blip the shutter and change the valve position, and watch the response.

The essential problem is that the As, more than any other element on the system, lingers in the chamber. So, if you want to use the As BEP, know how long it takes the reading on the BFM gauge to plateau (it usually creeps up for minutes). Perhaps one way to proceed is to throw out the hopes for measuring the BEP vs. valve position curve. Instead, one calibrates valve vs. absolute flux in the manner as for Sb. (I would do this for only a couple of interesting points.) While this means that you never have to rely on the BEP vs. valve position curve, it's not a bad idea to try to measure it anyway, not least to be able to do quick checks of backlash and clogging. The question is how to reproducibly define the stable flux from the source for your BFM recipe. How long to wait, and how to subtract the background, which is more substantial than for the other sources? I think the best way of proceeding is to take a slow measurement of the flux from high to low. That way, after some equilibration, the background, though high, is being reduced at a steady rate.

The main advantage of working with the As source is that it cleans up the growth chamber a little bit by burying material on the walls, but it will also coat any unprotected viewport.

B.4 Cadmium arsenide in the clean room

The slightly embarrassing truth about Cd_3As_2 in the clean room is that we don't know a lot of the chemistry involved in its degradation. What we know for sure is that water and heat degrade it, and both together are deadly. Too bad, then, that our best-performing gates involve steam for three hours at 120 °C.

The general guidance is to iterate on an existing recipe, but do be open-minded. Keep the bakes as cool as you dare (I haven't noticed problems on the 115 °C hot plates), don't sonicate your films unless you need to, don't run the ALD hotter than 120 °C without expecting to lower your mobility, and be gentle while etching. (Note that some have reported a phase transition in Cd_3As_2 at 220 °C, though the thermometry is probably not reliably compared to the cleanroom hot plates or the table in the ALD chamber [62].)

A few tips based on mistakes I've made:

- If you think your film is not defective enough, you can always squeeze it with the tweezers and watch as a plaid pattern forms on the surface. This can be seen under the yellow lights in the lithography bays, under the microscope by the development benches, in the SEM, or, most commonly, at the bottom of a beaker after your sample breaks into pieces. Gentle handling is a must, but dropping your sample is to be avoided at all costs. It's better practice to grow on larger pieces and write off a region for handling with tweezers, rather than try to save the whole piece and destroy the sample with risky grips.

- Don't sonicate your GaSb samples unless you've exhausted all other options, and don't clean them if they're not dirty.

- You can plasma ash your samples. But if you're ashing for two seconds (it's not as uncommon as you might think, per the log, and I've seen such a recipe on printed paper—and followed it), the process will be different every time because the main ashing will be during the ignition. In any event, one should be hesitant to put one's sample in the asher if a three second run, rather than two second one, will destroy it. Consider reducing the power and/or oxygen flow to make your process reproducible.

- We have had no success in improving the top gate by increasing the temperature in the ALD system, though hope springs eternal. The main effect is a reduction in the sample mobility, as side-by-side experiments have shown.
- Paranoia is the prime evil. Tools, chemicals, and processes that have been unfairly and publicly accused (of murdering Cd_3As_2 films) in my time in graduate school: Dehydration bake, photoresist spin-on, soft bake, hard bake, development, NMP rinse, plasma ashing, ion milling, e-beam deposition of Ti, Pt, Au, sputter deposition, ALD, and surely others. This isn't to say that you can't destroy a film with these tools or chemicals, or that I've never thrown around wild accusations. But the goal is to develop a consistent practice that allows you to reduce the suspects in new and mysterious film deaths to a manageable number.
- A minor evil is the bizarre ritual. To some extent one must learn to live with these, but not if they're ruining your devices. One example is the practice of blowing off the paper in the spinner with nitrogen after HMDS spin-on. The inventor of this practice was presumably aware of the negative interaction between HMDS vapor (not the adsorbed molecules on the film surface) and photoresist and, following spin-on of the HMDS, wished to *clear the air.* I think it's fairly clear that this practice is more likely to spray whatever detritus has fallen onto the spinner paper—crumbly bits of compound

semiconductor, photoresist splatter, what have you—into your sticky photoresist, to be trapped, forever, like a bug in amber, than it is to add that extra 5% to your development. And yet one sees it. Better to let the fume hood do its job.

- It's good to work off-cycle, but it's bad to fall asleep in the cleanroom.

And a few suggestions for comparatively unexplored processes:

- Give another look to the SiON gate dielectric, which can be sputtered in the ion beam deposition tool. A good process was developed by Luca Galletti—perhaps you can find it on the hard drive. There was some skepticism at the time that this gate was an improvement over the ALD Al_2O_3 gate that has been our workhorse. There was also some suggestion that the Fermi level had been pinned as a result of the gate's deposition; the evidence for that was weak, and seems a weaker reason to avoid it still in that the effect doesn't seem to be greater than in the Al_2O_3 gate. The two distinct advantages of the SiON gate are (1) the process temperature is low (the sample is heated from the exposure to the plasma, but in some sense it is performed at room temperature) and the environment is dry and (2) the nitrogen plasma treatment, shown to reduce water at the film surface (and the water seems to be an issue vis-à-vis interface traps), can be performed, prior to deposition, *in situ* [21]. The main drawback to working on this is that the ion beam deposition system suffers from lots of unexpected maintenance and downtime, compared to the ALD chamber and the sputter systems.

- Do some broad-minded testing on gentle wet etches, preferably non-aqueous. Anything that is used to etch more than microscopic quantities of GaAs will etch Cd_3As_2 uncontrollably fast. Some promising work has been done by Omor Shoron in this vein with buffered HF. Even for mesa isolation in large devices, the ion mill, our current main etching tool, is sub-ideal because the etch rate is high and variable, and the resulting surface is neither smooth nor passivated.

- Improvements to the contact resistance may be necessary for transistors and also pulsed field measurements. For measuring Hall bars in pulsed field, it's important that the two-point resistance of the device be low enough that current will pass through the sample. The issue is the big capacitance from the cables down the cryostat that are in parallel with your current leads. This point is kind of moot in the quantum Hall regime—the resistance is too high to support the 100 kHz ac signal anyway, so you're stuck measuring at dc (which works, but barely and slowly). But there's interesting stuff to be done at higher density where the resistance is in a kind of intermediate regime where factor of two improvements to the contact resistance will open up the range of viable measurements.

Bibliography

[1] Sadao Adachi. *Properties of group-IV, III-V and II-VI semiconductors.* Wiley series in materials for electronic and optoelectronic applications. John Wiley & Sons, Chichester, England; Hoboken, NJ, 2005.

[2] A. Akrap, M. Hakl, S. Tchoumakov, I. Crassee, J. Kuba, M. O. Goerbig, C. C. Homes, O. Caha, J. Novák, F. Teppe, W. Desrat, S. Koohpayeh, L. Wu, N. P. Armitage, A. Nateprov, E. Arushanov, Q. D. Gibson, R. J. Cava, D. van der Marel, B. A. Piot, C. Faugeras, G. Martinez, M. Potemski, and M. Orlita. Magneto-optical signature of massless Kane electrons in Cd_3As_2. *Phys. Rev. Lett.*, 117:136401, Sep 2016.

[3] Mazhar N. Ali, Quinn Gibson, Sangjun Jeon, Brian B. Zhou, Ali Yazdani, and R. J. Cava. The crystal and electronic structures of Cd_3As_2, the three-dimensional electronic analogue of graphene. *Inorganic Chemistry*, 53(8):4062–4067, 2014. PMID: 24679042.

[4] N. P. Armitage, E. J. Mele, and Ashvin Vishwanath. Weyl and Dirac semimetals in three-dimensional solids. *Rev. Mod. Phys.*, 90:015001, Jan 2018.

[5] Grigory Bednik. Surface states in Dirac semimetals and topological crystalline insulators. *Phys. Rev. B*, 98:045140, Jul 2018.

[6] G. R. Bell and C. F. McConville. Atomic hydrogen cleaning of GaSb(001) surfaces. *Applied Physics Letters*, 69(18):2695–2697, 1996.

[7] F.A.P. Blom, J.W. Cremers, J.J. Neve, and M.J. Gelten. Anisotropy of the electronic g^*-factor in cadmium arsenide. *Solid State Communications*, 33(1):69 – 73, 1980.

[8] Jan Bodnar. Band structure of Cd_3As_2 from Shubnikov–de Haas and de Haas–van Alphen effects, in *Proceedings of the 3rd International Conference on Physics of Narrow-Gap Semiconductors.* PWN, Warsaw, 1977, p.311. See also arXiv:1709.05845.

[9] Sergey Borisenko, Quinn Gibson, Danil Evtushinsky, Volodymyr Zabolotnyy, Bernd Büchner, and Robert J. Cava. Experimental realization of a three-dimensional Dirac semimetal. *Phys. Rev. Lett.*, 113:027603, Jul 2014.

[10] A.S Bracker, M.J Yang, B.R Bennett, J.C Culbertson, and W.J Moore. Surface reconstruction phase diagrams for InAs, AlSb, and GaSb. *Journal of Crystal Growth*, 220(4):384 – 392, 2000.

[11] C. Brüne, C. X. Liu, E. G. Novik, E. M. Hankiewicz, H. Buhmann, Y. L. Chen, X. L. Qi, Z. X. Shen, S. C. Zhang, and L. W. Molenkamp. Quantum Hall effect from the topological surface states of strained bulk HgTe. *Phys. Rev. Lett.*, 106:126803, Mar 2011.

[12] A. A. Burkov and Leon Balents. Weyl semimetal in a topological insulator multilayer. *Phys. Rev. Lett.*, 107:127205, Sep 2011.

[13] Jennifer Cano, Barry Bradlyn, Zhijun Wang, Max Hirschberger, N. P. Ong, and B. A. Bernevig. Chiral anomaly factory: Creating Weyl fermions with a magnetic field. *Phys. Rev. B*, 95:161306, Apr 2017.

[14] Bing Cheng, Natsuki Kanda, Tatsuhiko N. Ikeda, Takuya Matsuda, Peiyu Xia, Timo Schumann, Susanne Stemmer, Jiro Itatani, N. P. Armitage, and Ryusuke Matsunaga. Efficient terahertz harmonic generation with coherent acceleration of electrons in the dirac semimetal Cd_3As_2. *Phys. Rev. Lett.*, 124:117402, Mar 2020.

[15] I. Crassee, R. Sankar, W.-L. Lee, A. Akrap, and M. Orlita. 3D Dirac semimetal Cd_3As_2: A review of material properties. *Phys. Rev. Materials*, 2:120302, Dec 2018.

[16] J. P. Eisenstein and A. H. MacDonald. Bose-Einstein condensation of excitons in bilayer electron systems. *Nature*, 432(7018):691–694, 2004.

[17] Jack Ekin. *Experimental Techniques for Low-Temperature Measurements: Cryostat Design, Material Properties and Superconductor Critical-Current Testing*. Oxford University Press, Oxford, 2006.

[18] Liang Fu and C. L. Kane. Topological insulators with inversion symmetry. *Phys. Rev. B*, 76:045302, Jul 2007.

[19] Liang Fu and C. L. Kane. Superconducting proximity effect and majorana fermions at the surface of a topological insulator. *Phys. Rev. Lett.*, 100:096407, Mar 2008.

[20] Liang Fu, C. L. Kane, and E. J. Mele. Topological insulators in three dimensions. *Phys. Rev. Lett.*, 98:106803, Mar 2007.

[21] Luca Galletti, Timo Schumann, Thomas E. Mates, and Susanne Stemmer. Nitrogen surface passivation of the Dirac semimetal Cd_3As_2. *Phys. Rev. Materials*, 2:124202, Dec 2018.

[22] Luca Galletti, Timo Schumann, Omor F. Shoron, Manik Goyal, David A. Kealhofer, Honggyu Kim, and Susanne Stemmer. Two-dimensional dirac fermions in thin films of Cd_3As_2. *Phys. Rev. B*, 97:115132, Mar 2018.

[23] Zihao Gao, Meng Hua, Haijun Zhang, and Xiao Zhang. Classification of stable Dirac and Weyl semimetals with reflection and rotational symmetry. *Phys. Rev. B*, 93:205109, May 2016.

[24] C. Gerardi, C. Giannini, L. De Caro, L. Tapfer, Y. Rouillard, B. Jenichen, L. Däweritz, and K. H. Ploog. Secondary-ion-mass spectrometry and high-resolution x-ray diffraction analyses of gasb–algasb heterostructures grown by molecular beam epitaxy. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, 19(3):836–842, 2001.

[25] H.-J. Gossmann, G. P. Schwartz, B. A. Davidson, and G. J. Gualtieri. Strain and critical thickness in GaSb(001)/AlSb. *Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena*, 7(4):764–766, 1989.

[26] Manik Goyal, Luca Galletti, Salva Salmani-Rezaie, Timo Schumann, David A. Kealhofer, and Susanne Stemmer. Thickness dependence of the quantum Hall effect in films of the three-dimensional Dirac semimetal Cd_3As_2. *APL Materials*, 6(2):026105, 2018.

[27] Manik Goyal, Salva Salmani-Rezaie, Tyler N. Pardue, Binghao Guo, David A. Kealhofer, and Susanne Stemmer. Carrier mobilities of (001) cadmium arsenide films. *APL Materials*, 8(5):051106, 2020.

[28] M. Hakl, S. Tchoumakov, I. Crassee, A. Akrap, B. A. Piot, C. Faugeras, G. Martinez, A. Nateprov, E. Arushanov, F. Teppe, R. Sankar, Wei-li Lee, J. Debray, O. Caha, J. Novák, M. O. Goerbig, M. Potemski, and M. Orlita. Energy scale of Dirac electrons in Cd_3As_2. *Phys. Rev. B*, 97:115206, Mar 2018.

[29] E Hall and H Kroemer. Surface morphology of GaSb grown on (111)B GaAs by molecular beam epitaxy. *Journal of Crystal Growth*, 203(3):297 – 301, 1999.

[30] M. Z. Hasan and C. L. Kane. Colloquium: Topological insulators. *Rev. Mod. Phys.*, 82:3045–3067, Nov 2010.

[31] Conyers Herring. Accidental degeneracy in the energy bands of crystals. *Phys. Rev.*, 52:365–373, Aug 1937.

[32] R. Jackiw. Fractional charge and zero modes for planar systems in a magnetic field. *Phys. Rev. D*, 29:2375–2377, May 1984.

[33] Sangjun Jeon, Brian B. Zhou, Andras Gyenis, Benjamin E. Feldman, Itamar Kimchi, Andrew C. Potter, Quinn D. Gibson, Robert J. Cava, Ashvin Vishwanath, and Ali Yazdani. Landau quantization and quasiparticle interference in the three-dimensional Dirac semimetal Cd_3As_2. *Nature Materials*, 13:851, Sep 2014.

[34] Matija Karalic, Susanne Mueller, Christopher Mittag, Kiryl Pakrouski, QuanSheng Wu, Alexey A. Soluyanov, Matthias Troyer, Thomas Tschirky, Werner Wegscheider, Klaus Ensslin, and Thomas Ihn. Experimental signatures of the inverted phase in InAs/GaSb coupled quantum wells. *Phys. Rev. B*, 94:241402, Dec 2016.

[35] Mehdi Kargarian, Yuan-Ming Lu, and Mohit Randeria. Deformation and stability of surface states in Dirac semimetals. *Phys. Rev. B*, 97:165129, Apr 2018.

[36] Mehdi Kargarian, Mohit Randeria, and Yuan-Ming Lu. Are the surface Fermi arcs in Dirac semimetals topologically protected? *Proceedings of the National Academy of Sciences*, 113(31):8648–8652, 2016.

[37] David A. Kealhofer, Luca Galletti, Timo Schumann, Alexey Suslov, and Susanne Stemmer. Topological insulator state and collapse of the quantum Hall effect in a three-dimensional Dirac semimetal heterojunction. *Phys. Rev. X*, 10:011050, Feb 2020.

[38] David A. Kealhofer, Honggyu Kim, Timo Schumann, Manik Goyal, Luca Galletti, and Susanne Stemmer. Basal-plane growth of cadmium arsenide by molecular beam epitaxy. *Phys. Rev. Materials*, 3:031201, Mar 2019.

[39] M. Kellogg, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Vanishing Hall resistance at high magnetic field in a double-layer two-dimensional electron system. *Phys. Rev. Lett.*, 93:036801, Jul 2004.

[40] Honggyu Kim, Manik Goyal, Salva Salmani-Rezaie, Timo Schumann, Tyler N. Pardue, Jian-Min Zuo, and Susanne Stemmer. Point group symmetry of cadmium arsenide thin films determined by convergent beam electron diffraction. *Phys. Rev. Materials*, 3:084202, Aug 2019.

[41] Earl J Kirkland. *Advanced computing in electron microscopy*. Springer, New York, 2nd ed. edition, 2010.

[42] D. A. Kozlov, D. Bauer, J. Ziegler, R. Fischer, M. L. Savchenko, Z. D. Kvon, N. N. Mikhailov, S. A. Dvoretsky, and D. Weiss. Probing quantum capacitance in a 3D topological insulator. *Phys. Rev. Lett.*, 116:166802, Apr 2016.

[43] H. A. Kramers. Théorie générale de la rotation paramagnetique dans les cristaux. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, 33(6-10):959–972, 1930.

[44] Herbert Kroemer. Polar-on-nonpolar epitaxy. *Journal of Crystal Growth*, 81(1):193 – 204, 1987.

[45] Herbert Kroemer. The 6.1 Å family (InAs, GaSb, AlSb) and its heterostructures: a selective review. *Physica E: Low-dimensional Systems and Nanostructures*, 20(3):196 – 203, 2004. Proceedings of the 11th International Conference on Narrow Gap Semiconductors.

[46] Tian Liang, Quinn Gibson, Mazhar N. Ali, Minhao Liu, R. J. Cava, and N. P. Ong. Ultrahigh mobility and giant magnetoresistance in the Dirac semimetal Cd_3As_2. *Nature Materials*, 14(3):280–284, 2015.

[47] Jae S. Lim. *Two-dimensional signal and image processing*. Prentice-Hall signal processing series. Prentice Hall, Englewood Cliffs, N.J., 1990.

[48] Chao-Xing Liu, HaiJun Zhang, Binghai Yan, Xiao-Liang Qi, Thomas Frauenheim, Xi Dai, Zhong Fang, and Shou-Cheng Zhang. Oscillatory crossover from two-dimensional to three-dimensional topological insulators. *Phys. Rev. B*, 81:041307, Jan 2010.

[49] Z. K. Liu, J. Jiang, B. Zhou, Z. J. Wang, Y. Zhang, H. M. Weng, D. Prabhakaran, S.-K. Mo, H. Peng, P. Dudin, T. Kim, M. Hoesch, Z. Fang, X. Dai, Z. X. Shen, D. L. Feng, Z. Hussain, and Y. L. Chen. A stable three-dimensional topological Dirac semimetal Cd_3As_2. *Nature Materials*, 13(7):677–681, 2014.

[50] Serge Luryi. Quantum capacitance devices. *Applied Physics Letters*, 52(6):501–503, 1988.

[51] David M. Mahler, Julian-Benedikt Mayer, Philipp Leubner, Lukas Lunczer, Domenico Di Sante, Giorgio Sangiovanni, Ronny Thomale, Ewelina M. Hankiewicz, Hartmut Buhmann, Charles Gould, and Laurens W. Molenkamp. Interplay of Dirac nodes and Volkov-Pankratov surface states in compressively strained HgTe. *Phys. Rev. X*, 9:031034, Aug 2019.

[52] R. E. Marsh. Centrosymmetric or noncentrosymmetric? *Acta Crystallographica Section B*, 42(2):193–198, Apr 1986.

[53] J.W. Matthews and A.E. Blakeslee. Defects in epitaxial multilayers: I. Misfit dislocations. *Journal of Crystal Growth*, 27:118 – 125, 1974.

[54] Koichi Momma and Fujio Izumi. *VESTA3* for three-dimensional visualization of crystal, volumetric and morphology data. *Journal of Applied Crystallography*, 44(6):1272–1276, Dec 2011.

[55] J. E. Moore and L. Balents. Topological invariants of time-reversal-invariant band structures. *Phys. Rev. B*, 75:121306, Mar 2007.

[56] Adriano Mosca Conte, Olivia Pulci, and Friedhelm Bechstedt. Electronic and optical properties of topological semimetal Cd_3As_2. *Scientific Reports*, 7:45500, Apr 2017.

[57] Yusuke Nakazawa, Masaki Uchida, Shinichi Nishihaya, Markus Kriener, Yusuke Kozuka, Yasujiro Taguchi, and Masashi Kawasaki. Structural characterisation of high-mobility Cd_3As_2 films crystallised on $SrTiO_3$. *Scientific Reports*, 8(1):2244, February 2018.

[58] A. Narayanan, M. D. Watson, S. F. Blake, N. Bruyant, L. Drigo, Y. L. Chen, D. Prabhakaran, B. Yan, C. Felser, T. Kong, P. C. Canfield, and A. I. Coldea. Linear magnetoresistance caused by mobility fluctuations in n-doped Cd_3As_2. *Phys. Rev. Lett.*, 114:117201, Mar 2015.

[59] Madhab Neupane, Su-Yang Xu, Raman Sankar, Nasser Alidoust, Guang Bian, Chang Liu, Ilya Belopolski, Tay-Rong Chang, Horng-Tay Jeng, Hsin Lin, Arun Bansil, Fangcheng Chou, and M. Zahid Hasan. Observation of a three-dimensional topological Dirac semimetal phase in high-mobility Cd_3As_2. *Nature Communications*, 5(1):3786, 2014.

[60] H.B. Nielsen and Masao Ninomiya. The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal. *Physics Letters B*, 130(6):389 – 396, 1983.

[61] O.A. Pankratov, S.V. Pakhomov, and B.A. Volkov. Supersymmetry in heterojunctions: Band-inverting contact on the basis of $Pb_{1-x}Sn_xTe$ and $Hg_{1-x}Cd_xTe$. *Solid State Communications*, 61(2):93 – 96, 1987.

[62] A. Pietraszko and K. Łukaszewicz. Thermal expansion and phase transitions of Cd_3As_2 and Zn_3As_2. *physica status solidi (a)*, 18(2):723–730, 1973.

[63] L. A. Ponomarenko, R. Yang, R. V. Gorbachev, P. Blake, A. S. Mayorov, K. S. Novoselov, M. I. Katsnelson, and A. K. Geim. Density of states and zero landau level probed through capacitance of graphene. *Phys. Rev. Lett.*, 105:136801, Sep 2010.

[64] Andrew C. Potter, Itamar Kimchi, and Ashvin Vishwanath. Quantum oscillations from surface Fermi arcs in Weyl and Dirac semimetals. *Nature Communications*, 5:5161, Oct 2014.

[65] Arthur J. Rosenberg and Theodore C. Harman. Cd_3As_2—A noncubic semiconductor with unusually high electron mobility. *Journal of Applied Physics*, 30(10):1621–1622, 1959.

[66] Adriaan M. J. Schakel. Relativistic quantum Hall effect. *Phys. Rev. D*, 43:1428–1431, Feb 1991.

[67] Timo Schumann, Luca Galletti, David A. Kealhofer, Honggyu Kim, Manik Goyal, and Susanne Stemmer. Observation of the quantum Hall effect in confined films of the three-dimensional Dirac semimetal Cd_3As_2. *Phys. Rev. Lett.*, 120:016801, Jan 2018.

[68] Timo Schumann, Manik Goyal, David A. Kealhofer, and Susanne Stemmer. Negative magnetoresistance due to conductivity fluctuations in films of the topological semimetal Cd_3As_2. *Phys. Rev. B*, 95:241113, Jun 2017.

[69] Timo Schumann, Manik Goyal, Honggyu Kim, and Susanne Stemmer. Molecular beam epitaxy of Cd_3As_2 on a III-V substrate. *APL Materials*, 4(12):126110, 2016.

[70] Dexi Shao, Jiawei Ruan, Juefei Wu, Tong Chen, Zhaopeng Guo, Haijun Zhang, Jian Sun, Li Sheng, and Dingyu Xing. Strain-induced quantum topological phase transitions in Na_3Bi. *Phys. Rev. B*, 96:075112, Aug 2017.

[71] D. Shoenberg. *Magnetic Oscillations in Metals*. Cambridge Monographs on Physics. Cambridge University Press, 1984.

[72] Omor F. Shoron, Manik Goyal, Binghao Guo, David A. Kealhofer, Timo Schumann, and Susanne Stemmer. Prospects of terahertz transistors with the topological semimetal cadmium arsenide. *Advanced Electronic Materials*, 6(10):2000676, 2020.

[73] Lee Stauffer. C-V measurement tips, tricks, and traps. Technical report, Keithley Instruments, Inc., Cleveland, OH, 2011.

[74] G. A. Steigmann and J. Goodyear. The crystal structure of Cd_3As_2. *Acta Crystallographica Section B*, 24(8):1062–1067, Aug 1968.

[75] G.B. Stringfellow. Miscibility gaps in quaternary III/V alloys. *Journal of Crystal Growth*, 58(1):194 – 202, 1982.

[76] Takeyoshi Sugaya and Mitsuo Kawabe. Low-temperature cleaning of GaAs substrate by atomic hydrogen irradiation. *Japanese Journal of Applied Physics*, 30(Part 2, No. 3A):L402–L404, Mar 1991.

[77] S. Tchoumakov, V. Jouffrey, A. Inhofer, E. Bocquillon, B. Plaçais, D. Carpentier, and M. O. Goerbig. Volkov-Pankratov states in topological heterojunctions. *Phys. Rev. B*, 96:201302, Nov 2017.

[78] C. Thomas, A. T. Hatke, A. Tuaz, R. Kallaher, T. Wu, T. Wang, R. E. Diaz, G. C. Gardner, M. A. Capano, and M. J. Manfra. High-mobility InAs 2DEGs on GaSb substrates: A platform for mesoscopic quantum transport. *Phys. Rev. Materials*, 2:104602, Oct 2018.

[79] Dagim Tilahun, Byounghak Lee, E. M. Hankiewicz, and A. H. MacDonald. Quantum Hall superfluids in topological insulator thin films. *Phys. Rev. Lett.*, 107:246401, Dec 2011.

[80] Sandip Tiwari and David J. Frank. Empirical fit to band discontinuities and barrier heights in III–V alloy systems. *Applied Physics Letters*, 60(5):630–632, 1992.

[81] Masaki Uchida, Yusuke Nakazawa, Shinichi Nishihaya, Kazuto Akiba, Markus Kriener, Yusuke Kozuka, Atsushi Miyake, Yasujiro Taguchi, Masashi Tokunaga, Naoto Nagaosa, Yoshinori Tokura, and Masashi Kawasaki. Quantum Hall states observed in thin films of Dirac semimetal Cd_3As_2. *Nature Communications*, 8(1):2274, 2017.

[82] M. v. Stackelberg and R. Paulu. Untersuchungen an den Phosphiden und Arseniden des Zinks und Cadmiums. Das Zn_3P_2-gitter. *Zeitschrift für Physikalische Chemie*, 28B(1):427 – 460, 01 Feb. 1935.

[83] T.D Veal, M.J Lowe, and C.F McConville. HREELS and photoemission study of GaSb(100)-(1×3) surfaces prepared by optimal atomic hydrogen cleaning. *Surface Science*, 499(2):251 – 260, 2002.

[84] Antonio Basanes Villaflor and Morihiko Kimata. Controlling plastic relaxation using graded superlattice buffer layers. *Japanese Journal of Applied Physics*, 28(Part 2, No. 3):L496–L499, Mar 1989.

[85] B.A. Volkov and O.A. Pankratov. Two-dimensional massless electrons in an inverted contact. *JETP Letters*, 42:178, 1985.

[86] R.J. Wagner, E.D. Palik, and E.M. Swiggard. Interband magnetoabsorption in $Cd_xZn_{3-x}As_2$. *Physics Letters A*, 30(3):175 – 176, 1969.

[87] P. R. Wallace. Electronic $g*$-factor in Cd_3As_2. *Physica Status Solidi (b)*, 92(1):49–55, 1979.

[88] Xiangang Wan, Ari M. Turner, Ashvin Vishwanath, and Sergey Y. Savrasov. Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates. *Phys. Rev. B*, 83:205101, May 2011.

[89] An-Qi Wang, Xing-Guo Ye, Da-Peng Yu, and Zhi-Min Liao. Topological semimetal nanostructures: From properties to topotronics. *ACS Nano*, 14(4):3755–3778, 2020. PMID: 32286783.

[90] Mei-Xiao Wang, Canhua Liu, Jin-Peng Xu, Fang Yang, Lin Miao, Meng-Yu Yao, C. L. Gao, Chenyi Shen, Xucun Ma, X. Chen, Zhu-An Xu, Ying Liu, Shou-Cheng Zhang, Dong Qian, Jin-Feng Jia, and Qi-Kun Xue. The coexistence of superconductivity and topological order in the Bi_2Se_3 thin films. *Science*, 336(6077):52–55, 2012.

[91] Zhijun Wang, Hongming Weng, Quansheng Wu, Xi Dai, and Zhong Fang. Three-dimensional Dirac semimetal and quantum transport in Cd_3As_2. *Phys. Rev. B*, 88:125427, Sep 2013.

[92] Su-Yang Xu, Ilya Belopolski, Nasser Alidoust, Madhab Neupane, Guang Bian, Chenglong Zhang, Raman Sankar, Guoqing Chang, Zhujun Yuan, Chi-Cheng Lee, Shin-Ming Huang, Hao Zheng, Jie Ma, Daniel S. Sanchez, BaoKai Wang, Arun Bansil, Fangcheng Chou, Pavel P. Shibayev, Hsin Lin, Shuang Jia, and M. Zahid Hasan. Discovery of a Weyl fermion semimetal and topological Fermi arcs. *Science*, 349(6248):613–617, 2015.

[93] Yang Xu, Ireneusz Miotkowski, Chang Liu, Jifa Tian, Hyoungdo Nam, Nasser Alidoust, Jiuning Hu, Chih-Kang Shih, M. Zahid Hasan, and Yong P. Chen. Observation of topological surface state quantum Hall effect in an intrinsic three-dimensional topological insulator. *Nature Physics*, 10(12):956–963, 2014.

[94] Zhongbo Yan, Zhigang Wu, and Wen Huang. Vortex end Majorana zero modes in superconducting Dirac and Weyl semimetals. *Phys. Rev. Lett.*, 124:257001, Jun 2020.

[95] Bohm-Jung Yang and Naoto Nagaosa. Classification of stable three-dimensional Dirac semimetals with nontrivial topology. *Nature Communications*, 5(1):4898, September 2014.

[96] Hemian Yi, Zhijun Wang, Chaoyu Chen, Youguo Shi, Ya Feng, Aiji Liang, Zhuojin Xie, Shaolong He, Junfeng He, Yingying Peng, Xu Liu, Yan Liu, Lin Zhao, Guodong Liu, Xiaoli Dong, Jun Zhang, M. Nakatake, M. Arita, K. Shimada, H. Namatame, M. Taniguchi, Zuyan Xu, Chuangtian Chen, Xi Dai, Zhong Fang, and X. J. Zhou. Evidence of topological surface state in three-dimensional Dirac semimetal Cd_3As_2. *Scientific Reports*, 4(1):6106, 2014.

[97] S. M. Young, S. Zaheer, J. C. Y. Teo, C. L. Kane, E. J. Mele, and A. M. Rappe. Dirac semimetal in three dimensions. *Phys. Rev. Lett.*, 108:140405, Apr 2012.