Association for Women in Mathematics Series

Bahar Acu · Catherine Cannizzo · Dusa McDuff · Ziva Myer · Yu Pan · Lisa Traynor *Editors*

Research Directions in Symplectic and Contact Geometry and Topology

Springer

Association for Women in Mathematics Series

Volume 27

Series Editor
Kristin Lauter
Facebook
Seattle, WA, USA

Association for Women in Mathematics Series

Focusing on the groundbreaking work of women in mathematics past, present, and future, Springer's Association for Women in Mathematics Series presents the latest research and proceedings of conferences worldwide organized by the Association for Women in Mathematics (AWM). All works are peer-reviewed to meet the highest standards of scientific literature, while presenting topics at the cutting edge of pure and applied mathematics. Since its inception in 1971, The Association for Women in Mathematics has been a non-profit organization designed to help encourage women and girls to study and pursue active careers in mathematics and the mathematical sciences and to promote equal opportunity and equal treatment of women and girls in the mathematical sciences. Currently, the organization represents more than 3000 members and 200 institutions constituting a broad spectrum of the mathematical community, in the United States and around the world.

Titles from this series are indexed by Scopus.

More information about this series at http://www.springer.com/series/13764

The Virtual Series on Symplectic Geometry

More information about this series at http://www.springer.com/series/16019

Bahar Acu • Catherine Cannizzo • Dusa McDuff
Ziva Myer • Yu Pan • Lisa Traynor
Editors

Research Directions in Symplectic and Contact Geometry and Topology

Springer

AWM
ASSOCIATION FOR WOMEN IN MATHEMATICS

Editors
Bahar Acu
Department of Mathematics
ETH Zürich
Zürich, Switzerland

Catherine Cannizzo
Simons Center for Geometry and Physics
State University of New York
Stony Brook, NY, USA

Dusa McDuff
Department of Mathematics, Barnard College
Columbia University
New York, NY, USA

Ziva Myer
Department of Mathematics
Bryn Mawr College
Bryn Mawr, PA, USA

Yu Pan
Center of Applied Mathematics
Tianjin University
Tianjin, Tianjin, China

Lisa Traynor
Department of Mathematics
Bryn Mawr College
Bryn Mawr, PA, USA

ISSN 2364-5733 ISSN 2364-5741 (electronic)
Association for Women in Mathematics Series
ISBN 978-3-030-80978-2 ISBN 978-3-030-80979-9 (eBook)
https://doi.org/10.1007/978-3-030-80979-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

About This Volume

This volume highlights the fruits of the mathematical research collaborations initiated at the 2019 Women in Symplectic and Contact Geometry and Topology (WiSCon) Workshop. This event was held at the Institute for Computational and Experimental Research in Mathematics (ICERM) in Providence, Rhode Island on July 22–26, 2019. The WiSCon Workshop is a Research Collaboration Conference for Women (RCCW) in the fields of contact and symplectic geometry, topology, and related areas of low-dimensional topology. RCCWs are week-long conferences, held at mathematics institutes, where women and nonbinary mathematicians at various career stages in certain mathematical areas collaborate in groups on projects designed and guided by leaders in the field. This peer-reviewed volume is the proceedings of the first WiSCon workshop and features seven submissions from nine project groups featured in the WiSCon Workshop.

About the 2019 WiSCon Workshop

The 2019 WiSCon Workshop was organized by Bahar Acu, Catherine Cannizzo, Dusa McDuff, Ziva Myer, Yu Pan, and Lisa Traynor. The event was attended by 68 participants at various career stages including faculty, postdoctoral fellows, and graduate students. These participants came to Providence from all over the world, including from 14 states in the USA, Australia, Canada, China, England, France, Germany, Hungary, Pakistan, Portugal, Scotland, Serbia, and Sweden.

Based on their research interests and backgrounds, each participant was assigned to one of the nine project groups led by leading researchers in the field of symplectic and contact geometry and topology. In addition to intense research time, the event featured:

- A plenary talk by Dusa McDuff titled "Symplectic and Contact Geometry Today"
- Project introductions by project leaders
- Five research talks by junior participants
- Presentations reporting the progress by group spokespersons
- A career panel about funding opportunities; postdoc/tenure-track job applications
- A panel discussion about mentoring, diversity, equity, and inclusion and imposter syndrome

In this volume, you will find seven peer-reviewed research outputs from these nine projects whose detailed descriptions can be found in the Introduction section of this volume (Table 1).

For the full workshop schedule, please follow the workshop link below: https://icerm.brown.edu/topical_workshops/tw19-4-wiscon/

Table 1 Project titles and project leaders as they appeared in the WiSCon workshop

Project titles and leaders	
Title	Leaders
Applications of Heegaard Floer homology to low-dimensional topology	Jen Hom and Allison Moore
Khovanov homology and related invariants: local and global approaches	Radmila Sazdanović and Christine Ruey Shan Lee
Bordered invariants in contact manifolds	Ina Petkova and Vera Vértesi
Polyfold Laboratory	Katrin Wehrheim
Mirror symmetry and symplectic geometry	Chiu-Chu Melissa Liu and Ailsa Keating
Homological invariants, braids, transverse links, and surfaces	Eli Grigsby, Diana Hubbard, Keiko Kawamuro, and Olga Plamenevskaya
Weinstein Kirby calculus and Fukaya categories	Emmy Murphy and Laura Starkston
Lagrangian cobordisms between Legendrian submanifolds	Yu Pan and Lisa Traynor
Symmetry and moment maps in symplectic geometry and topology	Tara Holm and Ana Rita Pires

Acknowledgments

We would like to express our gratitude to the project leaders and contributing authors who submitted their manuscript to this proceedings volume, and to the referees who carefully reviewed the submissions and provided valuable feedback and suggestions to the authors. The first Women in Symplectic and Contact Geometry and Topology (WiSCon) Workshop was made possible with generous financial support and infrastructure from the host institution, ICERM, with partial support from the AWM ADVANCE grant funded by the National Science Foundation, and with the organizational guidance and encouragement by the leadership of AWM. The organizers of WiSCon, the editors, and the authors in this volume are all tremendously grateful for the support for this unique opportunity for research collaboration.

Finally, the editors would like to thank Kristin Lauter, the editor of the AWM-Springer Series, for the opportunity to publish this volume. We look forward to many more WiSCon research collaboration workshops in the future.

Zürich, Switzerland — Bahar Acu
Stony Brook, NY, USA — Catherine Cannizzo
New York, NY, USA — Dusa McDuff
Durham, NC, USA — Ziva Myer
Tianjin, Tianjin, China — Yu Pan
Bryn Mawr, PA, USA — Lisa Traynor
April 2021

Contents

A Polyfold Proof of Gromov's Non-squeezing Theorem 1
Franziska Beckschulte, Ipsita Datta, Irene Seifert, Anna-Maria Vocke, and Katrin Wehrheim

Infinite Staircases for Hirzebruch Surfaces 47
Maria Bertozzi, Tara S. Holm, Emily Maw, Dusa McDuff, Grace T. Mwakyoma, Ana Rita Pires, and Morgan Weiler

Action-Angle and Complex Coordinates on Toric Manifolds 159
Haniya Azam, Catherine Cannizzo, and Heather Lee

An Introduction to Weinstein Handlebodies for Complements of Smoothed Toric Divisors 217
Bahar Acu, Orsola Capovilla-Searle, Agnès Gadbled, Aleksandra Marinković, Emmy Murphy, Laura Starkston, and Angela Wu

Constructions of Lagrangian Cobordisms 245
Sarah Blackwell, Noémie Legout, Caitlin Leverson, Maÿlis Limouzineau, Ziva Myer, Yu Pan, Samantha Pezzimenti, Lara Simone Suárez, and Lisa Traynor

On Khovanov Homology and Related Invariants 273
Carmen Caprau, Nicolle González, Christine Ruey Shan Lee, Adam M. Lowrance, Radmila Sazdanović, and Melissa Zhang

Braids, Fibered Knots, and Concordance Questions 293
Diana Hubbard, Keiko Kawamuro, Feride Ceren Kose, Gage Martin, Olga Plamenevskaya, Katherine Raoux, Linh Truong, and Hannah Turner

Photographs 325

Introduction

The first notions of symplectic geometry date back to the mid-nineteenth century when Hamilton formulated the notion of a symplectic form in connection with his study of classical mechanics and the dynamics of energy conserving flows. Much work was done during the next century, by many including Poincaré, Brouwer, Arnold, and Moser, to understand the special properties of these flows. These were easiest to understand in two dimensions where they are simply area-preserving flows; many special features were discovered, for example, the existence in many situations of unexpectedly many fixed points.

The modern theory was inspired by the great Russian mathematician Arnold, who, motivated by results known in two dimensions, made far-reaching conjectures about analogous properties that should hold in higher dimensions. His conjectures concern not only the dynamical aspects of these flows but also the geometric properties of the phase space in which the flow takes place. Early work by Eliashberg on wave fronts in high dimensions and Thurston and Bennequin on flows in three manifolds showed that these new geometric structures were very rich. Another major achievement was the proof by Conley–Zehnder in 1984 of the Arnold conjectures about the fixed points of Hamiltonian flows on the standard $2n$-torus, which they proved by purely analytic methods. The key geometric breakthrough occurred very soon after, with the publication in 1985 of Gromov's paper about pseudo-holomorphic curves in symplectic manifolds. These "curves" have two real dimensions (equivalently, complex dimension one) and are the analog of geodesics in symplectic manifolds. They provide a key tool with which to analyze the structure of a symplectic manifold and its special submanifolds—for example, its Lagrangian submanifolds, which are manifolds of half the dimension of the ambient manifold and on which the symplectic form vanishes.

Gromov's paper contained several key results, for example, the nonsqueezing theorem, stating that a round ball cannot be symplectically embedded into a thinner cylinder, and the fact that there are no closed, exact Lagrangian submanifolds in standard symplectic Euclidean space $\mathbb{R}^{2n}$. These are instances of the *hard* aspect of symplectic geometry. Other results of Gromov's (e.g., the existence of symplectic structures on any noncompact manifold that satisfies a mild topological condition on

its tangent bundle) are instances of *softness*. These days, this hard/soft dichotomy is often phrased in terms of rigidity versus flexibility, and trying to pinpoint exactly where the boundary lies between these two regimes remains a central topic of investigation. Papers 1 and 2 in this volume directly stem from aspects of Gromov's work and explore the line between rigidity and flexibility.

The odd dimensional analog of a symplectic structure is a contact structure, a field of maximally non-integrable hyperplanes on a manifold X. The interesting geometric interplay between these two kinds of structure is illustrated in Paper 5. This paper, as well as many of the others presented here, focuses on four-dimensional phenomena, in part because this dimension is large enough to be interesting while being small enough for there to be extra topological tools that enable detailed calculations. Note that although dimension four is very special in smooth topology and many of the structures seen there either become simpler or simply disappear in high dimensions, the same simplification is not expected to happen (or at least not to the same extent) in symplectic/contact geometry.

One major advance in the 1990s was the reformulation of gauge theory (which at present exists only in four dimensions) in terms of the Seiberg–Witten equations, and the subsequent discovery by Taubes that Gromov's pseudo-holomorphic curves are limits of solutions of a suitably perturbed Seiberg–Witten equation. This created a very important link between smooth and symplectic geometry in dimensions four and below. It also led to further refinements of the Seiberg–Witten equations, most importantly to Ozsváth and Szabó's definition of a new class of gauge-theoretic invariants for three-manifolds and the curves (knots and links) that they contain. After much work, it is understood that the resulting knot invariants may be defined purely combinatorially and are closely related to other new invariants such as those coming from Khovanov homology. Thus, in low dimensions, there is a strong connection between symplectic and smooth topology, with similar structures being investigated and with many similarities among the tools used.

Today, the field of symplectic and contact geometry is extremely broad. The projects undertaken in the workshop give a sampling of this breadth. Here is a more detailed description of each of the papers in this volume.

Paper 1: *A Polyfold Proof of Gromov's Nonsqueezing Theorem*

Gromov's nonsqueezing theorem is a foundational result of symplectic topology, and it can be proved in many different ways. Gromov's original proof uses pseudo-holomorphic curves. These curves form the foundation for building many systems of invariants such as the Fukaya category and symplectic field theory. However, as these theories become more sophisticated, the analysis needed to define them rigorously becomes much more complicated, mostly because standard implicit function theorems are inadequate to deal with the bubbling and splitting of these curves. Polyfold theory is a completely new way of doing this analysis that has been developed over many years by Hofer–Wysocki–Zehnder, and is still rather

little known and understood. This paper explains how this new theory can be used to give a fully detailed and rigorous proof of the nonsqueezing theorem. Thus, it is a short and accessible introduction to this very important theory.

Paper 2: *Infinite Staircases for Hirzebruch Surfaces*

The Gromov width of a symplectic manifold X detects the largest radius of a standard ball that admits a symplectic embedding into X. This idea has been generalized to embeddings of ellipsoids: the ellipsoidal embedding capacity function $c_X(\lambda)$ for a symplectic four manifold X specifies exactly how much one needs to expand X in order for it to contain a symplectically embedded standard ellipsoid of eccentricity λ. This function is the maximum of the volume constraint and a piecewise linear function, and if it has infinitely many nonsmooth points, the target X is said to have a staircase. When X is a four-ball (or, equivalently, the projective plane), this function has been fully calculated, and was found to have a staircase with numerics determined by the Fibonacci numbers. The present paper studies the capacity function as X ranges over the family of Hirzebruch surfaces formed by blowing up the projective plane by a ball of varying size b, paying particular attention to finding values of b for which there is a staircase. This problem was explored by a fruitful combination of computer exploration with theory, and was found to have much more structure than might at first be imagined.

Paper 3: *Action-Angle and Complex Coordinates on Toric Manifolds*

A symplectic toric manifold is a $2n$-dimensional symplectic manifold M that admits a Hamiltonian action of the n-torus T. These manifolds have a simple explicit description in terms of the quotient M/T which (at least when M is compact) is a convex polytope. Because they have such a simple global description, they form important examples in which, for example, one can test ideas in mirror symmetry. Toric manifolds can also be thought of as a (partial) compactification of a complex n-dimensional torus. This paper explores the relation between the two natural systems of coordinates that are defined on this complex torus, writing down explicit formulas for the key objects of interest. It then describes various explicit examples that illustrate some key results in the theory of homological mirror symmetry.

Paper 4: *An Introduction to Weinstein Handlebodies for Complements of Smooth Toric Divisors*

A Weinstein domain is a special kind of symplectic manifold that can be built from handles. Because these domains have such an explicit description, they form a useful class of examples in which to explore the intricacies of the rigid/flexible boundary, or the properties of mirror symmetry. In four dimensions, they can be encoded diagrammatically by Legendrian fronts, with different diagrams related by explicit moves coming from "Weinstein Kirby calculus." This paper describes, and illustrates, an algorithm to produce Weinstein handle diagrams for the complements of certain (partially) smoothed toric divisors in toric four-manifolds.

Paper 5: *Constructions of Lagrangian Cobordisms*

A well-known result of Gromov is that there are no closed exact Lagrangian submanifolds in standard Euclidean space. However, in a symplectic manifold that is equivalent to standard Euclidean space, there do exist compact, exact Lagrangians with Legendrian boundary. These so called Lagrangian cobordisms at times mimic the flexibility seen in the world of topology, and at other times exhibit rigidity phenomena. There are many powerful techniques, coming from pseudo-holomorphic curves and Heegaard Floer theory, that have been employed to find obstructions to the existence of Lagrangian cobordisms in four dimensions. To complement these obstructions, there are some known combinatorial and geometric constructions, which are explained in this survey article. A basic question roughly asks whether the existence of a connected Lagrangian cobordism between two Legendrian knots is equivalent to the existence of a "decomposable" Lagrangian cobordism, which is one that can be constructed by stacking "elementary" cobordisms between intermediate Legendrians. This survey article gathers together a number of theorems that give potential candidates for Legendrians that are connected by a Lagrangian cobordism but not by one that is decomposable. In this way, this survey outlines strategies for finding a non-decomposable Lagrangian cobordism through explicit constructions from combinatorial and geometrical moves.

Paper 6: *On Khovanov Homology and Related Invariants*

In 2006, Plamenevskaya defined an invariant of a link transverse to the standard contact structure in the three-sphere using Khovanov homology. This unexpected discovery established a connection between Khovanov homology, a combinatorial invariant of links considered up to regular isotopy, and the contact geometry of transverse links, considered up to isotopy preserving the transversality of the

link to the standard contact structure of the three-sphere. Related works have similarly defined transverse link invariants using $\mathfrak{sl}(n)$ homology, a generalization of Khovanov homology. This paper surveys recent applications of Khovanov-type theories to low-dimensional topology and then derives new applications: a new obstruction to ribbon concordance, and new bounds to the alternating number and Turaev genus of a knot in S^3. This paper fits into the broader goal of understanding these link homology theories and their Plamenevskaya-type invariants.

Paper 7: *Braids, Fibered Knots, and Concordance Questions*

To understand three-dimensional contact manifolds, a number of techniques from low-dimensional topology have been employed and adapted. For example, there is a bijection between the set of oriented contact structures on a three-dimensional manifold M up to isotopy and the set of open book decompositions of M up to positive stabilization. Furthermore, it is known that a contact structure is tight if and only if the monodromy of each compatible open book is "right-veering." This paper explores different notions of monodromy that can be associated to a knot in S^3: a knot has braid representations, which can in turn be thought of as diffeomorphisms of punctured disks. Second, in the case that the knot is fibered, it can also be associated to the monodromy of its fibration. This article proves new results as well as poses conjectures and questions exploring how properties of these monodromies relate to four-dimensional properties of the corresponding knots. The primary objects of study in this paper—braids, fibered knots, open book decompositions, and the fractional Dehn twist coefficient—are all interesting in their own right both in the context of smooth low-dimensional topology and in the context of contact geometry. A deeper understanding of the relationships of these objects to each other helps to illuminate the connections between smooth and contact geometry.

A Polyfold Proof of Gromov's Non-squeezing Theorem

Franziska Beckschulte, Ipsita Datta, Irene Seifert, Anna-Maria Vocke, and Katrin Wehrheim

1 Introduction

The non-squeezing theorem, proven by Mikhail Gromov in 1985, essentially excludes nontrivial symplectic embeddings between balls B_R and cylinders Z_r of radius $R, r > 0$ given by

$$B_R := B^n_R := \left\{(x_i, y_i)_{i=1,\dots,n} \in \mathbb{R}^{2n} \;\middle|\; \textstyle\sum_{i=1}^n x_i^2 + y_i^2 \le R^2\right\}, \text{ and}$$
$$Z_r := B^2_r \times \mathbb{R}^{2n-2} := \left\{(x_i, y_i)_{i=1,\dots,n} \in \mathbb{R}^{2n} \;\middle|\; x_1^2 + y_1^2 \le r^2\right\},$$

in any dimension $2n \ge 4$. More precisely, we equip both the closed balls and closed cylinders above with the standard symplectic form $\omega_{\text{st}} = \sum_{i=1}^n dx_i \wedge dy_i$ as subsets of $\mathbb{R}^{2n}$ with coordinates $x_1, y_1, \dots, x_n, y_n$. Then it is easy to see that for any choice of R and r there are volume preserving embeddings $B_R \hookrightarrow Z_r$, due to the infinite length of the cylinder. If we only consider symplectic embeddings $\varphi : B_R \hookrightarrow Z_r$

F. Beckschulte
Fakultät für Mathematik, Ruhr-Universität Bochum, Bochum, Germany
e-mail: Franziska.Beckschulte@ruhr-uni-bochum.de

I. Datta
Department of Mathematics, Stanford University, Stanford, California, USA
e-mail: ipsi@stanford.edu

I. Seifert · A.-M. Vocke
Mathematical Institute, Heidelberg University, Heidelberg, Germany
e-mail: iseifert@mathi.uni-heidelberg.de; avocke@mathi.uni-heidelberg.de

K. Wehrheim (✉)
Department of Mathematics, University of California, Berkeley, CA, USA
e-mail: wehrheim@berkeley.edu

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_1

with $\varphi^*\omega_{\text{st}} = \omega_{\text{st}}$, there are trivial embeddings for $R \leq r$. However, symplectic embeddings cannot exist for $R > r$, as was shown by Gromov [7]—with various more detailed proofs published subsequently, e.g. [8, 11, 15, 20].[1]

Theorem 1.1 *If there is a symplectic embedding $\varphi : B_R \hookrightarrow Z_r$, then $R \leq r$.*

The idea of the proof is to construct an almost complex structure J_1 on the cylinder that pulls back to the standard complex structure $\varphi^* J_1 = J_{\text{st}}$ on the ball, and find a non-constant J_1-holomorphic curve $C_1 \subset Z_r$ passing through $\varphi(0)$ with symplectic area at most πr^2. Then the pullback $\varphi^{-1}(C_1) \subset B_R$ is a J_{st}-holomorphic curve, and thus a minimal surface with respect to the standard metric on $\mathbb{R}^{2n}$. As it passes through the center of the ball $0 \in B_R$, comparison with the flat disk of area πR^2 implies $R \leq r$ via a monotonicity lemma. To find such a J_1-holomorphic curve, one observes that the disk cross-section of Z_r at the height of $\varphi(0)$ has the required properties, except that it is holomorphic with respect to the standard complex structure J_0. The idea is then to establish the existence of J_t-holomorphic curves for a path J_t of almost complex structures connecting J_0 to J_1. This requires subtle geometric analysis that is best performed by studying spheres in the closed symplectic manifold $\mathbb{CP}^1 \times T^{2n-2}$ as described in Sect. 2.1. In this setting, the main work is to find a pseudoholomorphic curve in the homology class $[\mathbb{CP}^1 \times \{\text{pt}\}]$ through a fixed point. Here and throughout we denote the $(2n-2)$-dimensional torus by T and note that the following existence of pseudoholomorphic spheres can be generalized to other closed symplectic manifolds T (see Remark 1.3).

Theorem 1.2 *Given any point p_0 and compatible almost complex structure J on $\mathbb{CP}^1 \times T$, there exists a J-holomorphic sphere $u : S^2 \to \mathbb{CP}^1 \times T$ with $p_0 \in u(S^2)$ and homology class $u_*[S^2] = [\mathbb{CP}^1 \times \{\text{pt}\}]$.*

Assuming this existence, Sect. 2.3 proves Theorem 1.1 by applying a monotonicity lemma for pseudoholomorphic maps that we discuss in Appendix 3.7.

Remark 1.3 We will prove Theorem 1.2 for any compact symplectic manifold (T, ω_T) with $\omega_T(\pi_2(T)) = 0$, which excludes bubbling in the given homology class of minimal symplectic area. The torus satisfies this assumption since $\pi_2(T) = 0$. For more general symplectic manifolds, our line of argument still applies but the polyfold setup would require the inclusion of bubble trees. Moreover, this only proves the result with a possibly nodal J-curve.

Classical non-squeezing proofs establish Theorem 1.2 only for "generic" J and require a delicate analysis of linearized Cauchy-Riemann operators to show that their surjectivity can be achieved alongside the condition $\varphi^* J = J_{\text{st}}$. Our proof of this more general result demonstrates what proofs of geometric statements look

[1] Theorem 1.1, stated for closed balls and cylinders, implies the analogous statement for open balls and cylinders as follows: Assume there is a symplectic embedding φ of the open ball of radius R into the open cylinder of radius r, where $R > r$. Then there is $\varepsilon > 0$ such that $R - \varepsilon > r + \varepsilon$, and φ restricts to a symplectic embedding of the closed ball of radius $R - \varepsilon$ into the closed cylinder of radius $r + \varepsilon$. This is a contradiction to Theorem 1.1.

like when they can build on abstract polyfold theory [9] and an existing polyfold description of the relevant moduli spaces, like for Gromov–Witten spaces given in [10]. When analytical difficulties are outsourced to polyfold theory, the proof of Theorem 1.2 becomes a transparent geometric argument:

- The space $\mathcal{M}(J)$ of solutions in Theorem 1.2 modulo reparameterization is the zero set of a Fredholm section $\sigma_J : \mathcal{B} \to \mathcal{E}_J$ as described in Sect. 3.
- For the standard complex structure $J = J_0$ we show in Lemma 2.2 that $\mathcal{M}(J_0) = \{[u_0]\}$ consists of a unique solution $u_0(z) = (z, \pi_T(p_0))$. Theorem 3.17 moreover shows that the linearized operator $\mathrm{D}_{[u_0]}\sigma_{J_0}$ is surjective.
- Given any other $J = J_1$, Sect. 2.4 explains that a smooth family (J_t) of compatible almost complex structures connecting it to J_0 gives rise to a compact family of moduli spaces $\bigsqcup_{t\in[0,1]} \mathcal{M}(J_t)$. Section 3 identifies it with the zero set of the Fredholm section $\sigma(t, [u]) = \sigma_{J_t}([u])$ over $[0, 1] \times \mathcal{B}$.
- If we assume $\mathcal{M}(J_1)$ to be empty, this implies transversality of σ over the boundary $\{0, 1\} \times \mathcal{B}$. Then Sect. 2.5 uses the polyfold regularization scheme explained in Sect. 1.1 to construct a compact 1-dimensional cobordism $(\sigma + p)^{-1}(0)$ between $\mathcal{M}(J_0) = \{[u_0]\}$ and $\mathcal{M}(J_1) = \emptyset$. This proves Theorem 1.2 by contradiction.

Remark 1.4 The exact meaning of 'Fredholm section' in the first step determines the class in which the last step provides a contradiction. Minimal work in the first step would be to cite [4, 10] for a general description in which $\mathcal{B}$ is a polyfold (possibly containing nodal curves or curves with nontrivial isotropy). However, this would force us to work with multivalued perturbations p and discuss weighted branched orbifolds in the last step.[2] Instead, we show in Sect. 3 that the specific choice of homology class rules out isotropy, so that we are working with an M-polyfold $\mathcal{B}$. Then p is single-valued, and the perturbed solution set $(\sigma + p)^{-1}(0)$ is a manifold, contradicting the fact that compact 1-manifolds have an even number of boundary points.

The above outline uses polyfold theory entirely as a "black box" with two features: (a) it describes compactified moduli spaces as zero sets of 'Fredholm sections'; (b) such 'Fredholm sections' can be perturbed to regularize the moduli space. To demystify feature (a), Sect. 3 gives an introduction to the rather technical polyfold description of general Gromov-Witten moduli spaces by going through the details of [10] for our specific case. Since we consider curves of minimal positive symplectic area, we can exclude nodal curves arising from bubbling so that our ambient space $\mathcal{B}$ is the space of maps $S^2 \to \mathbb{CP}^1 \times T$ (of Sobolev class $W^{3,2}$) modulo Möbius transformations of S^2 that fix a marked point. Moreover, we show that this space has trivial isotropy, that is, none of these maps—holomorphic or

[2] The perturbation p is generally a 'multisection functor' resulting in $\widetilde{\mathcal{M}}^p := (\sigma + p)^{-1}(0)$ being a weighted branched orbifold. The contradiction to its boundary (in appropriate orientation) $\partial\widetilde{\mathcal{M}}^p = \mathcal{M}(J_0)$ being a single point (of trivial isotropy and weight 1) then arises from Stokes' Theorem $0 = \int_{\widetilde{\mathcal{M}}^p} \mathrm{d}(1) = \int_{\partial\widetilde{\mathcal{M}}^p} 1 = 1$ which holds in this context by Hofer et al. [9, §9.5].

otherwise—is invariant under reparameterization with a nontrivial Möbius transformation. This gives the ambient space $\mathcal{B}$ the structure of a sc-Hilbert manifold. We describe this notion in the following section as part of a brief introduction to polyfold theory. This section also demystifies feature (b) by stating the perturbation theorem in the trivial isotropy case that is relevant for the non-squeezing proof.

1.1 Polyfold Notions and Regularization Theorems

Polyfold theory was developed by Hofer, Wysocki, and Zehnder (see [9] and the citations therein) as a general solution to the challenge of regularizing compactified moduli spaces of pseudoholomorphic curves. The expectation is that any compact moduli space $\overline{\mathcal{M}}$ that is described as zero set of a section can be regularized by appropriate perturbations of the section. For smooth sections in finite dimensions this is proven in e.g. [6, ch.2].

Theorem 1.5 (Finite Dimensional Regularization) *Let $E \to B$ be a smooth finite dimensional vector bundle, and let $s : B \to E$ be a smooth section such that $s^{-1}(0)$ is compact. Then there exist arbitrarily small, compactly supported, smooth perturbation sections $p : B \to E$ such that $s + p$ is transverse to the zero section, and hence $(s + p)^{-1}(0)$ is a manifold of dimension* $\dim(B) - \operatorname{rank}(E)$.

Moreover, the perturbed zero sets $(s + p')^{-1}(0)$ and $(s + p)^{-1}(0)$ of any two such perturbations $p, p' : B \to E$ are cobordant.

A direct generalization of this theorem applies when B has boundary giving rise to perturbed zero sets with boundary $\partial(s + p)^{-1}(0) = (s + p)^{-1}(0) \cap \partial B$. Unfortunately, moduli spaces of pseudoholomorphic curves generally do not have natural descriptions to which this theorem applies. There are several reasons:

1. The space of pseudoholomorphic maps $u : S \to M$, with fixed domain S and target M, is the zero set of a Fredholm section of a Banach bundle. But due to 'bubbling and breaking' phenomena, its compactification contains maps defined on different domains.[3] To put a topology on the resulting set of maps from varying domains, one uses 'pregluing constructions' to transfer maps from 'nodal or broken' domains to nearby smooth domains. However, these constructions do not provide homeomorphisms to open subsets of Banach spaces (see e.g. [3, p.10]), so they do not yield local charts for any classical generalization of (topological or smooth) manifolds.
2. The moduli spaces typically arise as quotients of spaces of pseudoholomorphic maps by groups of reparameterizations of their domains. If these groups act with nontrivial isotropy, then we expect an orbifold structure on any ambient space that contains the moduli space.

[3] The domains have fixed genus and number of punctures/ends, but vary in nodal structure. Some 'SFT neck stretching' moduli spaces also vary the target space M.

3. If we wish for a differentiable structure on this ambient space, then we need to ensure that reparameterizations act differentiably. However, this is not the case for the classical Banach manifold structures on spaces of maps. As a result, while there are local charts for maps-modulo-reparameterization with fixed domain (constructed as local slices to the group action), the transition maps between different charts are nowhere differentiable.

The classical regularization constructions for moduli spaces of pseudoholomorphic curves, such as [15], work around these problems by finding geometric perturbations that achieve transversality for spaces of maps with fixed domain. Once they achieve finite dimensional spaces of perturbed solutions, this requires further steps to take quotients and compactify. Whether or not there are sufficiently many geometric perturbations that are both equivariant and compatible with gluing constructions depends on the particular geometric setting. The classical proof of the non-squeezing theorem makes use of the geometric setting of 'least energy' to rule out (1) nodal curves as well as (2) isotropy (due to multiple covers), so that only (3) the differentiability challenge is present. The latter is resolved by finding a regular choice of J with $\varphi^* J = J_{\mathrm{st}}$. This requires showing that holomorphic maps in the desired homology class all pass injectively through a part of $(\mathbb{CP}^1 \times T) \setminus \varphi(B_R)$, where we can freely vary the almost complex structure. While this approach yields a rigorous proof, it requires—even in the simplest case—sophisticated analysis combined with specific geometric properties of the holomorphic curves.

Polyfold theory, on the other hand, uses the above challenges as a guide to generalize the notion of a section in Theorem 1.5 so that the abstract perturbation theory applies to the desired moduli spaces, and no further case-specific analysis or geometric properties are needed. The main features are as follows:

1. The 'pregluing constructions' generalize open subsets of Banach spaces to images of retractions as new local models. For example, the neighbourhood of a nodal sphere is described by an open subset of the model space $\operatorname{im}\rho = \bigcup_{a\in\mathbb{C}}\{a\} \times \pi_a(V)$ that arises from a family of projections $\rho : \mathbb{C} \times V \to \mathbb{C} \times V, (a, v) \mapsto (a, \pi_a v)$ on a Banach space V, which are centered at $\pi_0 = \mathrm{id}$ and $\operatorname{im}\pi_a \subsetneq V$ for $a \neq 0$. For further details see e.g. [3, 2.3].
2. The orbifold structure is captured by formulating the notion of an atlas as a groupoid. This provides a nonsingular structure as in (1) on the object space (where e.g. perturbations are constructed), with isotropy appearing in the morphisms (e.g. forcing the perturbations to become multivalued).
3. Differentiability of transition maps between different local charts (resp. the structure maps in the groupoid) is achieved by defining a new notion of scale-differentiability for maps between Banach spaces equipped with an additional scale structure. These notions are obtained by formalizing the differentiability features of reparameterization maps between Sobolev spaces into a notion that satisfies a chain rule (see [3, 2.2]).

Restricted to finite dimensional Banach spaces, the retractions in (1) are trivial, and (3) coincides with classical differentiability, so (2) reproduces the notion

of an orbifold being represented by a proper étale groupoid (see e.g. [13] for an introduction). In infinite dimensions, these generalizations yield the following new notions. Here and throughout, we restrict the notions of [9] to metrizable topologies[4]and sc-Hilbert spaces.[5] In each case, sc-compatibility means that the transition maps are scale-smooth.

- A **sc-Hilbert manifold** is a metric space equipped with sc-compatible local homeomorphisms to open subsets of sc-Hilbert spaces.
- An **M-polyfold** is a metric space equipped with sc-compatible local homeomorphisms to open subsets of scale-smooth retracts in sc-Hilbert spaces.
- A **polyfold** is a metric space equipped with sc-compatible local homeomorphisms to finite quotients of open subsets of scale-smooth retracts in sc-Hilbert spaces. These domains with group actions and lifts of transition maps form a proper groupoid whose object and morphism spaces are M-polyfolds, and whose structure maps are local sc-diffeomorphisms.
- A sc-Hilbert manifold/M-polyfold/polyfold $\mathcal{B}$ **with boundary**[6] is a space with compatible charts as before but allowing for open subsets of $[0, \infty) \times H$, where H is a sc-Hilbert space. Its boundary $\partial\mathcal{B}$ is the union of all preimages of $\{0\} \times H$.

Remark 1.6 Every manifold M (with boundary) is a sc-Hilbert manifold: It is locally homeomorphic to open subsets of $\mathbb{R}^n$ (or $[0, \infty) \times \mathbb{R}^{n-1}$). Here each $\mathbb{R}^k$ is a sc-Hilbert space with trivial scale structure; see [3, Ex.4.1.8].

With this language in place, the application of polyfold theory to a given moduli space $\overline{\mathcal{M}}$—for example the moduli space in Theorem 1.2—has two steps:

(a) Describe $\overline{\mathcal{M}} \cong \sigma^{-1}(0)$ as the zero set of a section $\sigma : \mathcal{B} \to \mathcal{E}$ over a polyfold or M-polyfold $\mathcal{B}$ with $\mathcal{E}$ a 'strong bundle' and σ 'scale-Fredholm' as defined in [3, 9]. Intuitively, $\mathcal{B}$ is the same space of possibly-nodal-maps-modulo-reparameterization as $\overline{\mathcal{M}}$ but allows for general maps in some Sobolev space, with the pseudoholomorphic condition encoded in the section $\sigma([u]) = [\overline{\partial}_J u]$ of an appropriate bundle $\mathcal{E}$. This step is best achieved by combining existing polyfold descriptions such as [10] with general construction principles such as restrictions [4], pullbacks [17], quotients [21]. For our example, we describe this in detail in Sect. 3.

(b) Apply the corresponding regularization theorem to draw the desired conclusions. For general polyfolds (with boundary) this is [9, Thm.15.4 (15.5)] and

[4]Metrizability of polyfolds is guaranteed by paracompactness and [9, Thm.7.2].

[5]Hilbert spaces equipped with scale structures automatically admit scale-smooth bump functions by Hofer et al. [9, §5.5]. These are crucial for the existence of transverse perturbations.

[6]For an introduction to the notion of corners in polyfold theory see [3, §5.3]. Note, however, that corners usually appear together with coherence conditions—in which perturbations on boundary strata need to coincide with the perturbations of other moduli spaces which are identified with these boundary strata. The construction of coherent perturbations then requires not just a workable notion of boundary and corner strata, but an ordering of the moduli spaces that prevents circular coherence conditions when constructing perturbations.

involves multivalued perturbations. In our example, the M-polyfold versions in Theorem 1.7 and Remark 1.8 below suffice.

The following generalization of the finite dimensional regularization Theorem 1.5 is proven in [9, Theorems 3.4, 5.5, 5.6]. It resolves the challenges (1) and (3) above, hence covering the case required for the non-squeezing proof.

Theorem 1.7 (M-Polyfold Regularization) *Let $\mathcal{E} \to \mathcal{B}$ be a strong M-polyfold bundle, and let $\sigma : \mathcal{B} \to \mathcal{E}$ be a scale-smooth Fredholm section such that $\sigma^{-1}(0)$ is compact. Then there exists a class of perturbation sections $p : \mathcal{B} \to \mathcal{E}$ supported near $\sigma^{-1}(0)$ such that $(\sigma + p)^{-1}(0)$ carries the structure of a smooth compact manifold of dimension* $\operatorname{index}(\sigma)$ *with boundary $\partial(\sigma+p)^{-1}(0) = (\sigma+p)^{-1}(0) \cap \partial\mathcal{B}$.*

Moreover, for any other such perturbation $p' : \mathcal{B} \to \mathcal{E}$, there exists a smooth cobordism between $(\sigma + p')^{-1}(0)$ and $(\sigma + p)^{-1}(0)$.

Remark 1.8 Suppose that the section σ in Theorem 1.7 restricts to a transverse section on the boundary, i.e. $\sigma|_{\partial\mathcal{B}} : \partial\mathcal{B} \to \mathcal{E}|_{\partial\mathcal{B}}$ has surjective linearizations at all points in $\sigma^{-1}(0) \cap \partial\mathcal{B}$. Then we can choose the perturbation section p to be supported in the interior, i.e. $p|_{\partial\mathcal{B}} \equiv 0$. As a result, $(\sigma + p)^{-1}(0)$ has boundary $\partial(\sigma + p)^{-1}(0) = (\sigma + p)^{-1}(0) \cap \partial\mathcal{B} = (\sigma|_{\partial\mathcal{B}})^{-1}(0)$.

This can be proven by following the proof of the regularization theorem in [9]. It is explicitly stated and proven in the last part of [5, Thm.A9]. In our case, the map $e : \mathcal{X} \to \emptyset$ and submanifolds $C_i = \emptyset$ are trivial, and the polyfold $\mathcal{X} = \mathcal{B}$ has trivial isotropy. So, the 'multisection' λ will be represented by a perturbation section $p : \mathcal{B} \to \mathcal{E}$. Our transversality assumption on the boundary means that the 'trivial multisection λ^δ representing' $p|_{\partial\mathcal{B}} \equiv 0$ yields an 'admissible ... multisection in general position to the perturbed zero set in the boundary' $\{x \in \partial\mathcal{B} \,|\, \sigma(x) = 0\}$. The conclusion is the existence of a perturbation section p with $p|_{\partial\mathcal{B}} \equiv 0$ so that $\sigma + p$ is 'admissible' and in 'general position', as required for the conclusions of Theorem 1.7.

For readers interested in regularization theorems that resolve the challenge (2) of nontrivial isotropy, we recommend the brief overview [3, Rmk.2.1.7] and the in-depth discussion of the finite dimensional case [12] before diving into the technicalities of [9] or their summary in [17]. Despite a lot of notational overhead, the general polyfold regularization theorem [9, Thm.15.4] can be understood as a direct combination of the regularization theorems for sections over M-polyfolds and finite dimensional orbifolds.

2 Outline of the Proof

Let us consider a symplectic embedding $\varphi : B_R \hookrightarrow Z_r$ for radii $R > 0$ and $r > 0$. We will prove $R \leq r$ by showing that $R' \leq r + \varepsilon$ for any choice of $0 < R' < R$ and $\varepsilon > 0$. These choices are needed for constructions in the following section.

2.1 Compactifying the Target Space

The proof uses the theory of pseudoholomorphic curves. Since the analytic setup is simpler for closed manifolds, we prefer to work with a compact target space. For that purpose we fix an $\varepsilon > 0$. Then we can understand φ as an embedding

$$\varphi : B_R \hookrightarrow \mathring{Z}_{r+\varepsilon}$$

into the slightly larger open cylinder $\mathring{Z}_{r+\varepsilon} = \mathring{B}^2_{r+\varepsilon} \times \mathbb{R}^{2n-2}$. The first factor of this cylinder compactifies to a $\mathbb{CP}^1$. The standard symplectic form on $\mathring{B}^2_{r+\varepsilon}$ descends to a symplectic (and thus, area) form $\omega_{\mathbb{CP}^1}$ such that $\mathbb{CP}^1$ has area

$$\int_{\mathbb{CP}^1} \omega_{\mathbb{CP}^1} = \pi(r+\varepsilon)^2.$$

So we may view φ as a symplectic embedding

$$\varphi\colon (B_R, \omega_{\mathrm{st}}) \hookrightarrow (\mathbb{CP}^1 \times \mathbb{R}^{2n-2}, \omega_{\mathbb{CP}^1} \oplus \omega_{\mathrm{st}}).$$

Now, we want to compactify the second factor of the cylinder as well. Remember that B_R is the closed ball. So, the projection to $\mathbb{R}^{2n-2}$ of its image under the continuous map φ is compact. This means that we can choose $N > 0$ sufficiently large such that $\varphi(B_R) \subset \mathbb{CP}^1 \times (-\frac{1}{2}N, \frac{1}{2}N)^{2n-2}$. Then we can view $\varphi(B_R)$ as a subset of the $(2n-2)$-dimensional torus $T := \mathbb{R}^{2n-2}/N\mathbb{Z}^{2n-2}$ with standard symplectic form ω_T induced from ω_{st} on $\mathbb{R}^{2n-2}$. This means we get a symlplectic embedding (again denoted by) φ,

$$\varphi : (B_R, \omega_{\mathrm{st}}) \hookrightarrow (\mathbb{CP}^1 \times T, \omega := \omega_{\mathbb{CP}^1} \oplus \omega_T).$$

The proof now proceeds by studying pseudoholomorphic curves in $\mathbb{CP}^1 \times T$. Here we wish to work with an almost complex structure J_1 on $\mathbb{CP}^1 \times T$ so that the pullback of J_1-holomorphic maps under the embedding φ yields pseudoholomorphic maps to B_R with respect to the standard complex structure J_{st} on $B_R \subset \mathbb{R}^{2n}$. This is crucial for the last step of the non-squeezing proof in Sect. 2.3 which uses monotonicity with respect to the standard metric $\omega_{\mathrm{st}}(\cdot, J_{\mathrm{st}}\cdot)$ on B_R. To do this rigorously, we need to shrink the ball slightly to interpolate between almost complex structures.

Lemma 2.1 *For any $0 < R' < R$ there is an almost complex structure J_1 on $\mathbb{CP}^1 \times T$ that is compatible with $\omega = \omega_{\mathbb{CP}^1} \oplus \omega_T$ and satisfies $\varphi^* J_1|_{B_{R'}} = J_{\mathrm{st}}$.*

Proof The basic idea is to define $J_1 = \varphi_* J_{\mathrm{st}}$ on the image of φ and to set $J_1 = J_0$ outside a neighbourhood of the image. But as a weighted sum of two almost complex structures will in general not be an almost complex structure, we can not directly interpolate between these. Instead, we interpolate between the

corresponding Riemannian metrics $g_0 := \omega(\cdot, J_0\cdot)$ on $\mathbb{CP}^1 \times T$ and $g_{\varphi_* J_{\mathrm{st}}} := \omega(\cdot, \varphi_* J_{\mathrm{st}}\cdot)$ on $\varphi(B_R)$. To do this, we choose a partition of unity $\psi_0 + \psi_1 = 1$ subordinate to the cover $\mathbb{CP}^1 \times T = U_0 \cup U_1$ where $U_0 := \mathbb{CP}^1 \times T \setminus \varphi(B_{R'})$ and $U_1 := \varphi(\mathring{B}_R)$. (These are open subsets because φ, being an embedding, maps open/closed subsets to open/closed subsets.) Since ψ_i is supported in U_i and $\varphi(B'_R) \cap U_0 = \emptyset$, we have $\psi_1|_{\varphi(B'_R)} \equiv 1$ and thus obtain a metric g_1 with $g_1|_{\varphi(B'_R)} = g_{\varphi_* J_{\mathrm{st}}}$ by interpolating with this partition of unity,

$$g_1 := \psi_0 \cdot g_0 + \psi_1 \cdot g_{\varphi_* J_{\mathrm{st}}}.$$

Finally, a pair of a Riemannian metric g and a symplectic form ω determine an almost complex structure J compatible with ω and if g was of the form $g = \omega(\cdot, J\cdot)$, then the determined almost complex structure is in fact the same J, see [14, Prop. 2.50 (ii)]. Thus, g_1 and ω together determine an almost complex structure J_1 that has the required properties. □

There are two more properties of pullbacks $C_0 := \varphi^{-1}(C_1) \subset B_{R'}$ of J_1-holomorphic curves $C_1 \subset \mathbb{CP}^1 \times T$ that are required for their to prove the non-squeezing result $R' \leq r + \varepsilon$. First, we need C_1 to pass through the point $p_0 := \varphi(0) \in \mathbb{CP}^1 \times T$, so that $C_0 \subset B_{R'}$ passes through the center 0 of the ball. Second, we wish to bound the symplectic area $\int_{C_0} \omega_{\mathrm{st}} \leq \int_{C_1} \omega \leq \pi(r+\varepsilon)^2$. The latter is achieved by prescribing the homology class $[C_1] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}]$ since this determines the integral of the closed symplectic form ω,

$$\int_{C_1} \omega = \int_{\mathbb{CP}^1 \times \{\mathrm{pt}\}} \omega_{\mathbb{CP}^1} \oplus \omega_T = \int_{\mathbb{CP}^1} \omega_{\mathbb{CP}^1} + \int_{\{\mathrm{pt}\}} \omega_T = \pi(r+\varepsilon)^2 + 0.$$

Ultimately, we will find a not necessarily embedded curve $C_1 = u(S^2)$ by studying J_1-holomorphic maps $u : (S^2, i) \to (\mathbb{CP}^1 \times T^{2n-2}, J_1)$ with a point constraint $u(z_0) = p_0$ in the homology class $[u] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}]$. Their existence is stated, for general J on a product manifold $\mathbb{CP}^1 \times T$, in Theorem 1.2. The proof starts with the existence of a unique J_0-holomorphic map for a specific J_0 described in Sect. 2.2, and is completed in Sect. 2.5 based on the polyfold constructions in Sect. 3.

2.2 The Unique J_0-Holomorphic Curve

This section begins our study of pseudoholomorphic curves in $\mathbb{CP}^1 \times T$ by considering a split almost complex structure $J_0 = i \oplus J_T$ on $\mathbb{CP}^1 \times T$. Here (T, ω_T) can be any compact symplectic manifold, though its topology will be restriced in following sections. For the nonsqueezing proof, there is a standard complex structure J_T on the torus $T = \mathbb{R}^{2n-2}/N\mathbb{Z}^{2n-2}$; in general, we choose any ω_T-compatible almost complex structure J_T on T. This ensures that J_0 is compatible with $\omega = \omega_{\mathbb{CP}^1} \oplus \omega_T$, meaning $g_0 = \omega(\cdot, J_0\cdot)$ is a Riemannian metric. In the

following, we view S^2 as a Riemann surface by identifying it with $\mathbb{CP}^1$ and using the standard complex structure i on $\mathbb{CP}^1$. Then we find a J_0-holomorphic sphere passing through any given point $p_0 = (z_0, m_0) \in \mathbb{CP}^1 \times T$ by combining the identification $S^2 \cong \mathbb{CP}^1$ with a constant map to T,

$$u_0 : (S^2, i) \to (\mathbb{CP}^1 \times T, J_0), \quad z \mapsto (z, m_0).$$

The symplectic area of this sphere—a quantity that only depends on the homology class—is

$$E(u_0) = \int_{S^2} u_0^* \omega = \int_{\mathbb{CP}^1} \omega_{\mathbb{CP}^1} = \pi(r + \varepsilon)^2. \tag{1}$$

The next lemma shows that, up to reparameterization, this is the only sphere with these properties in its homology class.

Lemma 2.2 *Assume $u : (S^2, i) \to (\mathbb{CP}^1 \times T, J_0)$ is J_0-holomorphic, passes through p_0, and represents the class $[\mathbb{CP}^1 \times \{\text{pt}\}] \in H_2(\mathbb{CP}^1 \times T; \mathbb{Z})$. Then, there is a biholomorphism $\psi : (S^2, i) \to (S^2, i)$ such that $u \circ \psi = u_0$.*

Proof Since u is $(J_0 = i \oplus J_T)$-holomorphic, its composition with projection to each factor yields holomorphic maps $f := \text{pr}_{\mathbb{CP}^1} \circ u : S^2 \to \mathbb{CP}^1 \cong S^2$ and $g := \text{pr}_T \circ u : S^2 \to T$. Moreover, the homology condition specifies $g_*[S^2] = [\{\text{pt}\}] = 0 \in H_2(T)$ and $f_*[S^2] = [S^2] \in H_2(S^2)$. The energy identity $\int g^*\omega = \int \frac{1}{2}|\mathrm{d}g|^2$ (see [15, Lemma 2.2.1]) then implies $\int |\mathrm{d}g|^2 = 0$. So, g must be constant. Since u passes through $p_0 = (z_0, m_0)$, this means $g(z) = \text{pr}_T(u(z)) = m_0$. Moreover, $f_*[S^2] = [S^2] \in H_2(S^2)$ implies that f is neither constant nor a multiple cover of another holomorphic map. Thus, $\psi := f^{-1}$ exists and is a biholomorphism of S^2. Then we obtain the claim as $(u \circ \psi)(z) = \big(f(f^{-1}(z)),\, g(\psi(z)\big) = (z, m_0) = u_0$ for all $z \in S^2$. □

Remark 2.3 It is part of both the classical and our polyfold proof to show that the curve u_0 is transversely cut out of the space of all curves in its homology class passing through p_0. This statement will be made precise in Theorem 3.17.

2.3 Using the Monotonicity Lemma

This section finishes the proof of the nonsqueezing Theorem 1.1 assuming that we have found a J_1-holomorphic map $u_1 : S^2 \to \mathbb{CP}^1 \times T$ with the same properties as the unique J_0-holomorphic curve in Lemma 2.2, except that J_1 is an almost complex structure as in Lemma 2.1 with $\varphi^* J_1 = J_{\text{st}}$. The existence of u_1 follows from Theorem 1.2, proven in Sect. 2.5. Given such $u_1 : S^2 \to \mathbb{CP}^1 \times T$, we obtain a J_{st}-holomorphic map on $\widetilde{S} := u_1^{-1}(\varphi(\mathring{B}_{R'})) \subset S^2$,

$$v := \varphi^{-1} \circ u_1 : \widetilde{S} \longrightarrow \mathbb{R}^{2n}.$$

This map passes through the center $\varphi^{-1}(p_0) = 0$ of the ball $B_{R'}$ and has area $\int v^*\omega_{\mathrm{st}} \leq \int u_1^*\omega = \pi(r+\varepsilon)^2$, so comparison with the minimal surface through the center of the ball—the disk of area $\pi(R')^2$—will yield $R' \leq r + \varepsilon$. To deduce this inequality directly from the classical monotonicity formula for minimal surfaces (e.g. [2, Prop.1.12.]), we would have to show that v is an embedding. Instead, we use the monotonicity Lemma A.2 for holomorphic maps to a complex Hilbert space. To ensure accessibility (and establish this result in maximal generality) we include a detailed proof of this "well known result" in the appendix.

Proof of Theorem 1.1 We will apply Lemma A.2 to the map $v : \widetilde{S} \to \mathbb{R}^{2n}$ with $(V, J) := (\mathbb{R}^{2n}, J_{\mathrm{st}})$ and open balls $\mathring{B}_{R_k} := \mathring{B}_{R_k}(0) \subset \mathbb{R}^{2n}$ of radii $R_k \to R'$. The preimage of the center is nonempty, $v^{-1}(\{0\}) = u^{-1}(\{p_0\}) \neq \emptyset$, since u_1 passes through $p_0 = \varphi(0)$. The domain $\widetilde{S}$ of v is an open subset of S^2 because $\varphi(\mathring{B}_{R'}) \subset \mathbb{CP}^1 \times T$ is the image of an open set under an embedding. To apply the lemma we need to restrict v to a compact subdomain $S_k \subset \widetilde{S}$ with smooth boundary such that $\|v(z)\| \geq R_k$ for all $z \in \partial S_k$. For that purpose we consider the smooth function $\rho : \widetilde{S} \to \mathbb{R}, z \mapsto \|v(z)\|^2$. Since its regular values are dense we can find a sequence $0 < R_k < R'$ with limit $\lim_{k\to\infty} R_k = R'$ such that R_k^2 are regular values of ρ. Then $S_k := \{z \in \widetilde{S} \mid \|v(z)\| \leq R_k\}$ is a domain with smooth boundary $\partial S_k = \rho^{-1}(R_k^2)$. It is compact because $S_k = v^{-1}(B_{R_k}) = u_1^{-1}(\varphi(B_{R_k}) \subset S^2$ is a closed subset of the compact S^2. Moreover, $v|_{S_k}$ is nonconstant on each connected component of S_k, since u_1 is nonconstant (as it has positive energy) and its critical points in S^2 are a finite set by McDuff and Salamon [15, Lemma 2.4.1].

So, we can apply Lemma A.2 to $v|_{S_k} : S_k \to V = \mathbb{R}^{2n}$ and the open ball $\mathring{B}_{R_k} = \{q \in \mathbb{R}^n \mid \|q\| < R_k\}$ centered at $p = 0 \in \mathbb{R}^{2n}$ to obtain

$$\pi R_k^2 \leq \int_{v^{-1}(\mathring{B}_{R_k})} v^*\omega_{\mathrm{st}} = \int_{u_1^{-1}(\varphi(\mathring{B}_{R_k}))} u_1^*\omega \leq \int_{S^2} u_1^*\omega = \pi(r+\varepsilon)^2.$$

As $R_k \to R'$, this yields $\pi(R')^2 \leq \pi(r+\varepsilon)^2$ as claimed; and by taking $R' \to R$ and $\varepsilon \to 0$ this proves the non-squeezing $R \leq r$. □

2.4 A Compact Moduli Space

In this and the next subsection, we prove Theorem 1.2, while assuming that the M-polyfold construction in Sect. 3 holds.

Our argument is a special case of proving the independence of Gromov-Witten invariants from the choice of a compatible almost complex structure J. Indeed, in Lemma 2.2 we compute the number of pseudoholomorphic curves in the particular homology class intersecting the given point p_0 to be 1 for $J = J_0$. So, by showing

that this count is independent of J, we can show the existence of a J_1-holomorphic map in Theorem 1.2.

For that, we use the fact that the space of ω-compatible almost complex structures is contractible (see e.g. [14, Prop. 4.1]). Thus, we can choose a smooth path $(J_t)_{t\in[0,1]}$ of ω-compatible almost complex structures from J_0 to J_1. Moreover, we fix a marked point $z_0 \in S^2$ that we require is mapped to p_0 by all the considered maps.

Then, for every $t \in [0, 1]$, we define the moduli space of J_t-holomorphic curves

$$\mathcal{M}_t := \left\{ \begin{array}{c} u : S^2 \to \mathbb{CP}^1 \times T \\ \text{smooth} \end{array} \middle| \begin{array}{c} u(z_0) = p_0, \quad \overline{\partial}_{J_t} u = 0, \\ [u] = [\mathbb{CP}^1 \times \{\text{pt}\}] \end{array} \right\} \Big/ \sim, \tag{2}$$

where $u \sim u'$ iff there is a biholomorphism $\psi : S^2 \to S^2$ such that $u' = u \circ \psi$. Here $\overline{\partial}_{J_t}$ is the Cauchy-Riemann operator for J_t, that is, $\overline{\partial}_{J_t} u = \frac{1}{2}(\mathrm{d}u + J_t \circ \mathrm{d}u \circ i)$.

Now consider the moduli space for the family $(J_t)_{t\in[0,1]}$

$$\mathcal{M} := \{(t, [u]) \mid t \in [0, 1], [u] \in \mathcal{M}_t\}. \tag{3}$$

We will prove Theorem 1.2 by contradiction: Assuming $\mathcal{M}_1 = \emptyset$, in Sect. 2.5 we will show that a perturbation of $\mathcal{M}$ is a compact 1-dimensional cobordism from $\mathcal{M}_0$ to $\mathcal{M}_1$. However, $\mathcal{M}_0 = \{[u_0]\}$ consists of exactly one element (see Lemma 2.2). Therefore, this cobordism contradicts that $\mathcal{M}_1$ is empty.

The first step in completing the proof of Theorem 1.2 is to establish compactness of the unperturbed moduli space. This is a special case of Gromov's compactness theorem, where bubbling is excluded in the given homology class using the restriction of the topology of T from Remark 1.3. For general symplectic manifolds (T, ω_T), we would need to compactify $\mathcal{M}$ by bubble trees. So, the subsequent proof of $\mathcal{M}_1$ being nonempty would only show the existence of (possibly singular) J_1-curves and not necessarily smooth spheres in the required homology class.

Theorem 2.4 *Let (T, ω_T) be a compact symplectic manifold with $\omega_T(\pi_2(T)) = 0$. Then the moduli space $\mathcal{M}$ defined in Eq. (3) is compact with respect to the quotient topology induced by $[0, 1] \times \mathcal{C}^\infty(S^2, M)$.*

Proof This proof follows [8, Chapter 4] and [1]. Let $(t_n, [u_n])$ be a sequence in the moduli space $\mathcal{M}$. In particular, u_n is a sequence of J_{t_n}-holomorphic maps in $\mathcal{M}$. We want to show that there exists a subsequence which converges to an element $(t_\infty, [u_\infty]) \in \mathcal{M}$, meaning that u_∞ is a J_{t_∞}-holomorphic map. Here, by *convergence* we mean the following:

(1) t_n converges to t_∞ in the usual topology of $[0, 1] \subset \mathbb{R}$,
(2) $[u_n]$ converges to $[u_\infty] \in \mathcal{M}_{t_\infty}$ in the *Gromov sense*, that is, there exist biholomorphic maps $\varphi_n : (S^2, j) \to (S^2, j)$ with $\varphi_n(z_0) = z_0$ such that the reparameterized maps $u_n \circ \varphi_n : S^2 \to Q$ converge in C^∞ to u_∞.

To achieve (1) we can choose a subsequence of $t_n \in [0,1]$ with t_n converging to a $t_\infty \in [0,1]$ since the interval is compact. Then, as $\{J_t\}_{t\in[0,1]}$ is a continuous path, we can deduce $\mathcal{C}^\infty$-convergence of the almost complex structures $J_{t_n} \to J_{t_\infty}$. To achieve (2), we consider this subsequence, denoting it again by $(t_n, [u_n])$. The key observation is that the area functional is uniformly bounded on $\mathcal{M}$. Indeed, all J_{t_n}-holomorphic curves u_n represent the same homology class $[u] = [\mathbb{CP}^1 \times \{\text{pt}\}]$, so that

$$E(u_n) = \int_{\mathbb{CP}^1} u_n^*\omega = \omega([\mathbb{CP}^1 \times \{\text{pt}\}]) = \pi(r+\varepsilon)^2$$

is constant and hence bounded. Moreover, S^2 is a closed surface. Thus, by Gromov's compactness theorem (e.g. [8, Chapter V, Thm. 1.2]) there exists a subsequence of $[u_n]$ converging in the Gromov sense to a J_{t_∞}-holomorphic cusp curve $\overline{u}_\infty$ of the same energy $E(\overline{u}_\infty) = \pi(r+\varepsilon)^2$.

Actually, this cusp curve consists of a single J_{t_∞}-holomorphic sphere. Indeed, let $[v_1], \ldots, [v_k]$ be the non-constant components of $\overline{u}_\infty$ of energy $E(v_n) = \int_{\mathbb{CP}^1} v_n^*\omega > 0$ which sum to $E(v_1) + \ldots + E(v_k) = E(\overline{u}_\infty) = \pi(r+\varepsilon)^2$. Since the symplectic form $\omega = \omega_{\mathbb{CP}^1} \oplus \omega_T$ splits, the energies are the sums $E(u_n) = \omega_{\mathbb{CP}^1}(\alpha_n) + \omega_T(\beta_n)$ of symplectic areas of the projections $\alpha_n := [\text{pr}_{\mathbb{CP}^1} \circ u_n]$ and $\beta_n := [\text{pr}_T \circ u_n]$ to the factors $\mathbb{CP}^1$ and T. Here we have $\omega_T(\beta_n) = 0$ because of the assumption $\omega_T(\pi_2(T)) = 0$, and $\omega_{\mathbb{CP}^1}(\alpha_n) \in \mathbb{Z}\pi(r+\varepsilon)^2$ since $H_2(\mathbb{CP}^1)$ is generated by $[\mathbb{CP}^1]$ which has symplectic area $\pi(r+\varepsilon)^2$ by construction. Thus each nontrivial component has energy at least $E(v_n) \geq \pi(r+\varepsilon)^2$, but since the total energy of the bubble tree is $\pi(r+\varepsilon)^2$ this implies $k = 1$. This means that the limit cusp curve $\overline{u}_\infty$ has one non-constant component. Since it has only one marked point (arising from z_0 in the definition of $\mathcal{M}_t$), it cannot have ghost components, and thus $\overline{u}_\infty$ consists of a single J_{t_∞}-holomorphic sphere $u_\infty \colon S^2 \to \mathbb{CP}^1 \times T$.

The meaning of Gromov-convergence $[u_n] \to \overline{u}_\infty = [u_\infty]$ is exactly as stated in (2) above, so we have shown that $\mathcal{M}$ is sequentially compact. Finally, compactness follows from the fact that the Gromov-topology is metrizable; see [15, Theorem 5.6.6.]. □

2.5 Applying the Polyfold Regularization Scheme

This section proves Theorem 1.2. We will use the notation and facts established in Sects. 2.2 and 2.4, the polyfold constructions from Sect. 3, and the polyfold regularization scheme from Sect. 1.1.

Proof of Theorem 1.2 Assume that $\mathcal{M}_1 = \emptyset$. Under this assumption, we will the polyfold regularization scheme in Theorem 1.7 to perturb $\mathcal{M}$ just enough to achieve a smooth structure, while not loosing compactness, or changing its boundary at

$t = 0, 1$. With that, we obtain a compact cobordism between $\mathcal{M}_0 = \{[u_0]\}$ and $\mathcal{M}_1 = \emptyset$. For that purpose we construct the following objects in Sect. 3:[7]

- an ambient M-polyfold $[0, 1] \times \mathcal{B} \supset \mathcal{M}$ modeled on sc-Hilbert spaces (Theorem 3.5),
- a tame strong M-polyfold bundle $\mathcal{E} \to [0, 1] \times \mathcal{B}$ (Theorem 3.10),
- a sc-Fredholm section $\sigma : [0, 1] \times \mathcal{B} \to \mathcal{E}$ such that $\mathcal{M} = \sigma^{-1}(0) \subset [0, 1] \times \mathcal{B}$ (Theorem 3.14).

The use of sc-Hilbert spaces guarantees the existence of sc-smooth bump functions on $\mathcal{B}$ by Hofer et al. [9, §5.5], which is required for Theorem 1.7. We prove that σ is transverse to the zero section at $\{0\} \times \mathcal{B}$ in Theorem 3.17. Moreover, the assumption $\mathcal{M}_1 = \emptyset$ implies transversality of σ at $\{1\} \times \mathcal{B}$ (see Remark 3.18). Now we apply the M-polyfold regularization scheme, see Theorem 1.7. This gives a perturbation section $p : [0, 1] \times \mathcal{B} \to \mathcal{E}$, such that $(\sigma + p)^{-1} =: \mathcal{M}^p$ is a compact 1-dimensional manifold. By Remark 1.8, we can assume p to be supported inside $(0, 1) \times \mathcal{B}$. So, the boundary of $\mathcal{M}^p$ is

$$\partial(\mathcal{M}^p) = \mathcal{M}^p \cap \partial([0, 1] \times \mathcal{B}) = \mathcal{M}^p \cap (\{0, 1\} \times \mathcal{B}) \cong \mathcal{M}_0 \sqcup \mathcal{M}_1 = \{[u_0]\}.$$

Thus, the boundary $\partial(\mathcal{M}^p)$ consists of only one point. Such a manifold does not exist. Therefore, the assumption $\mathcal{M}_1 = \emptyset$ was false and we have proven Theorem 1.2. □

Remark 2.5 If we dropped the condition $u(z_0) = p_0$ in the construction of the moduli spaces $\mathcal{M}_t$ in Eq. (2), then we could directly use the polyfold setup for Gromov-Witten moduli spaces with one marked point from [10]. Then the above arguments would provide a (2n+1)-dimensional cobordism $\mathcal{M}^p$ between the 2n-dimensional manifolds $\mathcal{M}_0$ and $\mathcal{M}_1^p$, where the latter is obtained from $\mathcal{M}_1$ by the perturbation $p|_{\{1\}\times\mathcal{B}}$. One would need to choose the perturbation to be supported away from J_1-curves intersecting the point $p_0 \in \mathbb{CP}^1 \times T$ (which by assumption do not exist), so that the evaluation map on the perturbed moduli space $\mathrm{ev} : \mathcal{M}_1^p \to \mathbb{CP}^1 \times T$ does not contain p_0 in its image and hence has degree 0. Moreover, one would need to formulate Lemma 2.2 in a way that the evaluation map $\mathrm{ev} : \mathcal{M}_0 \to \mathbb{CP}^1 \times T$ is a bijection and thus has degree 1. This difference between degrees is then in contradiction to the fact that the evaluation map $\mathcal{B} \to \mathbb{CP}^1 \times T$ extends both $\mathrm{ev}|_{\mathcal{M}_1^p}$ and $\mathrm{ev}|_{\mathcal{M}_0}$ to a continuous map $\mathrm{ev} : \mathcal{M}^p \to \mathbb{CP}^1 \times T$ on the cobordism.

[7] We introduce a candidate space $\mathcal{B}$ in (4), then we establish the M-polyfold structure on an open subset $\mathcal{B}' \subset \mathcal{B}$, that we rename into $\mathcal{B}$ here for ease of notation. The exact choice of $\mathcal{B}'$ as discussed in Remark 3.6 is immaterial, since the moduli space $\mathcal{M}$ and all its regular perturbations will be automatically contained in $[0, 1] \times \mathcal{B}'$.

3 Polyfold Setup

This section provides the polyfold description of the compact moduli space $\mathcal{M}$ that is the basis of the proof of Theorem 1.2 in Sect. 2.5.

Recall that (T, ω_T, J_T) denotes the torus with a compatible pair of symplectic form and almost complex structure (e.g. the one constructed in Sect. 2.1). We may consider any other compact symplectic manifold with $\omega_T(\pi_2(T)) = 0$, as explained in Remark 1.3. Moreover, we fix a point p_0 and compatible almost complex structure J on $(Q, \omega) = (\mathbb{CP}^1 \times T, \omega_{\mathbb{CP}^1} \times \omega_T)$. We choose a smooth path $(J_t)_{t\in[0,1]}$ of ω-compatible almost complex structures from $J_0 = i \times J_T$ to $J_1 = J$. Then, Sect. 2.4 proves compactness of the family $\mathcal{M} = \{(t, [u]) \mid t \in [0, 1], [u] \in \mathcal{M}_t\}$ of moduli spaces

$$\mathcal{M}_t = \left\{ \begin{array}{c|c} u : S^2 \to \mathbb{CP}^1 \times T & u(z_0) = p_0, \quad \overline{\partial}_{J_t} u = 0, \\ \text{smooth} & [u] = [\mathbb{CP}^1 \times \{\text{pt}\}] \end{array} \right\} \Big/ \sim,$$

with the equivalence relation

$$u \sim v \quad :\Longleftrightarrow \quad \exists \psi : S^2 \to S^2 \text{ biholomorphism with } \psi(z_0) = z_0 \text{ and } u = v \circ \psi.$$

The polyfold setup starts with a choice of an ambient space that contains $\mathcal{M}$ (as a compact zero set of a sc-Fredholm section). For the family $\mathcal{M} = \bigcup_{t\in[0,1]} \mathcal{M}_t$, the natural choice of ambient space is $[0, 1] \times \mathcal{B}$, where $\mathcal{B}$ is an ambient space for each of the moduli spaces $\mathcal{M}_t$, given by

$$\mathcal{B} := \left\{ \begin{array}{c|c} u : S^2 \to \mathbb{CP}^1 \times T & u(z_0) = p_0, \\ \text{of class } W^{3,2} & [u] = [\mathbb{CP}^1 \times \{\text{pt}\}] \end{array} \right\} \Big/ \sim . \tag{4}$$

This uses the same equivalence relation $\sim$ as in the definition of $\mathcal{M}_t$, but we equip $\mathcal{B}$ with the quotient topology induced by the metrizable topology on the Hilbert manifold $H = W^{3,2}(S^2, \mathbb{CP}^1 \times T)$, unlike the smooth topology in Theorem 2.4. The condition $[u] = [\mathbb{CP}^1 \times \{\text{pt}\}]$ specifies some connected component(s) of H, and $u(z_0) = p_0$ cuts out a further submanifold, so that $\mathcal{B}$ is the quotient of a Hilbert manifold. However, the action by reparameterization with biholomorphisms is not differentiable (see e.g. [3, § 2.2]), and so $\mathcal{B}$ does not inherit the smooth structure of a Hilbert manifold. Instead, we will show in Theorem 3.5 that it carries the structure of a sc-Hilbert manifold.[8]

[8] Strictly speaking, our proofs establish the polyfold structures not for $\mathcal{B}$ and $\mathcal{E} \to \mathcal{B}$ as stated, but after restriction to a $W^{3,2}$-open neighbourhood $\mathcal{B}' \subset \mathcal{B}$ of the dense subset of smooth curves $\mathcal{B}_\infty \subset \mathcal{B}$. Additional estimates could prove $\mathcal{B}' = \mathcal{B}$, but applications of the polyfold description yield the results for any $\mathcal{B}_\infty \subset \mathcal{B}' \subset \mathcal{B}$; see Remark 3.6.

We sometimes write an equivalence class as $\alpha \in \mathcal{B}$ instead of $[u] \in \mathcal{B}$, indicating that there is no preferred representative in the class. To build the bundle $\mathcal{E} \to [0,1] \times \mathcal{B}$, we consider for each $(t, \alpha) \in [0,1] \times \mathcal{B}$, the Hilbert space quotient

$$\mathcal{E}_{(t,\alpha)} := \left\{ (u,\xi) \;\middle|\; \begin{array}{c} [u] = \alpha \\ \xi \in \Lambda^{0,1}_{J_t}\big(S^2, u^*\mathrm{T}(\mathbb{CP}^1 \times T)\big) \text{ of class } W^{2,2} \end{array} \right\} \Big/ \sim, \tag{5}$$

where $\Lambda^{0,1}_J(S^2, u^*\mathrm{T}Q)$ denotes the 1-forms on S^2 with values in the pullback bundle $u^*\mathrm{T}Q$ that are complex antilinear with respect to i on S^2 and J on $Q = \mathbb{CP}^1 \times T$. The equivalence relation $\sim$ is given by

$$(u,\xi) \sim (v,\zeta) \quad :\Longleftrightarrow \quad \begin{array}{c} \exists \psi : S^2 \to S^2 \text{ biholomorphism with } \psi(z_0) = z_0 \\ \text{and } u = v \circ \psi \text{ and } \xi = \zeta \circ \mathrm{d}\psi. \end{array}$$

Now the bundle $\mathcal{E} \to [0,1] \times \mathcal{B}$ is given by the total space

$$\mathcal{E} := \left\{ (t, [(u,\xi)]) \;\middle|\; t \in [0,1], \alpha \in \mathcal{B}, [(u,\xi)] \in \mathcal{E}_{(t,\alpha)} \right\} \tag{6}$$

with the projection to $[0,1] \times \mathcal{B}$. This projection is well-defined since $(u,\xi) \sim (v,\zeta)$ implies $u \sim v$. We equip $\mathcal{E}$ with the quotient topology obtained by embedding Q into a Euclidean space and viewing the pairs (u,ξ) as subset of the pairs of $W^{2,3}$-maps and vector-valued 1-forms of class $W^{2,2}$.

Finally, the section

$$\sigma : [0,1] \times \mathcal{B} \to \mathcal{E}, \qquad (t, [u]) \mapsto (t, [(u, \overline{\partial}_{J_t} u)]) \tag{7}$$

cuts out the moduli space $\sigma^{-1}(0) = \mathcal{M}$. To apply the M-polyfold regularization Theorem 1.7, we need to equip the spaces $\mathcal{B}$ and $\mathcal{E}$ with sc-smooth structures such that σ is sc-smooth, and moreover show that σ is a sc-Fredholm section.

3.1 The Gromov-Witten Space of Stable Curves

The sc-smooth structure on the base space $\mathcal{B}$ is obtained in Sect. 3.3 by understanding it as a subset $\mathcal{B} \subset Z$ of a space of stable curves in the manifold $Q = \mathbb{CP}^1 \times T$. The space

$$Z := \left\{ u : S^2 \to Q \text{ of class } W^{3,2} \;\middle|\; [u] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}] \right\} \Big/ \sim \tag{8}$$

does not satisfy the condition $u(z_0) = p_0$, but the marked point z_0 is present in the definition of the equivalence relation $\sim$, that equals the one in (4). Thus, there exists a well defined evaluation map $\mathrm{ev} : Z \to Q$, $\mathrm{ev}([u]) := u(z_0)$.

Then, the base space $\mathcal{B}$ can be viewed as the preimage of the point $p_0 \in Q$ under the evaluation map ev, that is

$$\mathcal{B} \;=\; \{[u] \in Z \mid u(z_0) = p_0\} \;=\; \mathrm{ev}^{-1}(\{p_0\}) \;\subset\; Z. \tag{9}$$

The space Z is a subspace of a polyfold that Hofer, Wysocki and Zehnder construct in [10], which we will denote by Z^{HWZ}. In [10], Hofer, Wysocki and Zehnder also construct a strong polyfold bundle $W \to Z^{\mathrm{HWZ}}$ and a sc-Fredholm section $\overline{\partial} : Z^{\mathrm{HWZ}} \to W$, that cuts out holomorphic curves. More precisely, for any numbers $g, k \in \mathbb{N}_0$ and nontrivial homology class $A \in H_2(Q)$, the polyfold Z^{HWZ} has a component[9] $Z^{\mathrm{HWZ}}_{g,k,A}$ so that

$$\overline{\partial}^{-1}(0) \cap Z^{\mathrm{HWZ}}_{g,k,A} = \overline{\mathcal{M}}_{g,k}(A)$$

is the compactified Gromov-Witten moduli space of (possibly nodal) pseudoholomorphic curves in class A of genus g, with k marked points. In the following, we will explain how considering genus $g = 0$, one marked point, and homology class $A := [\mathbb{CP}^1 \times \{\mathrm{pt}\}] \in H_2(Q)$ will provide an identification $Z \cong Z^{\mathrm{HWZ}}_{0,1,A} \subset Z^{\mathrm{HWZ}}$.

In general, the space $Z^{\mathrm{HWZ}} = \{(S, j, M, D, u) | \ldots\}/\sim$ is defined in [10, Definition 1.4,1.5] as a set, and given a topology in [10, §3.4], as a quotient of

$$\left\{ (S, j, M, D, u) \;\middle|\; \begin{array}{c} (S, j, M, D) \text{ connected nodal Riemann surface,} \\ u \in W^{3,2,\delta}(S, Q),\ \int_C u^*\omega \geq 0 \text{ for each component } C \subset S, \\ \int_C u^*\omega > 0 \text{ for each non-stable component } C \subset S \end{array} \right\}.$$

Here (S, j) is a (not necessarily connected) Riemann surface, $M \subset S$ is a finite set of marked points, and D is a finite set of nodal pairs in S. These nodal pairs are identified in order to obtain a noded Riemann surface—which is required to be connected. The maps $u\colon S \to Q$ are then required to descend to a continuous map from the noded Riemann surface, and they are required to be of weighted Sobolev class $W^{3,2,\delta}$ on the complement $S \setminus D$ of the nodal points.

Two stable maps $(S, \ldots, u) \sim (S', \ldots, u')$ are equivalent if there is a biholomorphism ψ between the corresponding marked noded Riemann surfaces (i.e., compatible with M and D), such that $u' = u \circ \psi$. The conditions on the symplectic area in Z^{HWZ} are known to guarantee that pseudoholomorphic curves of this type are stable, meaning that their isotropy groups are finite. We show in Lemma 3.2 that all maps in Z^{HWZ} have finite isotropy. In fact, Lemma 3.3 proves that all maps in our particular case $Z^{\mathrm{HWZ}}_{0,1,A}$ have trivial isotropy. To begin with, the following Lemma simplifies the description of this space to the description given in Eq. (8).

[9] These components are open and closed but not necessarily connected.

Lemma 3.1 *The polyfold $Z^{HWZ}_{0,1,A}$ for $A = [\mathbb{CP}^1 \times \{\text{pt}\}] \in H_2(Q)$ is naturally identified with Z. The same holds true if we replace the torus T in $Q = \mathbb{CP}^1 \times T$ by any compact symplectic manifold with $\omega_T\,(\pi_2(T)) = 0$.*

Proof First recall that the symplectic area of a map $u : S \to Q$ depends only on its homology class. Thus, for $[u] = [\mathbb{CP}^1 \times \{\text{pt}\}] = A \in H_2(Q)$, we obtain $\int_S u^*\omega = \omega(A) = \pi(r+\varepsilon)^2 > 0$, as computed in Eq. (1). This is the symplectic area of all equivalence classes of (possibly nodal) maps in

$$Z^{\text{HWZ}}_{0,1,A} = \left\{ [(S, j, M, D, u)] \in Z^{\text{HWZ}} \;\middle|\; \begin{array}{c} \text{genus of } (S, j, D) \text{ is } g = 0,\ \#M = 1, \\ [u] = A = [\mathbb{CP}^1 \times \{\text{pt}\}] \end{array} \right\}.$$

We claim that this can be simplified to the formulation given in (8). For that, we first recall that all the nodal surfaces of genus $g = 0$ are trees of spheres. Next, note that the assumption $\omega_T(\pi_2(T)) = 0$ guarantees that all components $S^2 \simeq C \subset S$ have energy $\int_C u^*\omega \in \omega(\pi_2(Q)) = \omega_{\mathbb{CP}^1}(\pi_2(\mathbb{CP}^1)) = \mathbb{Z}\pi(r+\varepsilon)^2$. Recall here that we chose the symplectic form on $\mathbb{CP}^1$ in Sect. 2.1 in such a way that we have $\omega_{\mathbb{CP}^1}([\mathbb{CP}^1]) = \int_{\mathbb{CP}^1} \omega_{\mathbb{CP}^1} = \pi(r+\varepsilon)^2$. As in the proof of Theorem 2.4, the homology condition $[u] = [\mathbb{CP}^1 \times \{\text{pt}\}]$ implies that S can only have one component, on which u is non-constant. Indeed, the total symplectic area $\int_S u^*\omega = \omega([\mathbb{CP}^1 \times \{\text{pt}\}]) = \pi(r+\varepsilon)^2$ is the sum of non-negative areas of all components, but each non-constant component has energy $\int_C u^*\omega \geq \pi(r+\varepsilon)^2$. Moreover, S cannot have so-called ghost components $C \subset S$ with $\int_C u^*\omega = 0$ because stability of such components would require at least three special points, while there is only one marked point from $k = 1$ and at most one nodal point connecting C to the unique non-constant component. Thus all nodal surfaces (S, j, D), that are needed for the component $Z^{\text{HWZ}}_{0,1,A} \subset Z^{\text{HWZ}}$, are single spheres, i.e. $D = \emptyset$.

The absence of nodes, i.e. $D = \emptyset$, also explains why we do not need to consider weighted Sobolev spaces but can directly work with maps of Sobolev class $W^{3,2}$. Moreover, the topology specified in [10, §3.4] simplifies in this setting to the quotient topology coming from the space of $W^{3,2}$-maps.

Finally, we will use the fact that any compact genus 0 Riemann surface (S, j) without nodes is biholomorphic to (S^2, i), so that for each point in $Z^{\text{HWZ}}_{0,1,A}$ we can choose representatives $[(S^2, i, M, \emptyset, u)]$. The remaining equivalence relation is then by biholomorphisms of (S^2, i), which we can use to fix the marked point $M = \{z_0\}$ and reduce the equivalence relation to reparameterization with biholomorphisms $\psi : S^2 \to S^2$ that fix the marked point $\psi(z_0) = z_0$ as in Eq. (4). This identifies the polyfold given in [10] with Z as defined in Eq. (8) in the following way:

$$Z^{\text{HWZ}}_{0,1,A} = \left\{ (S, j, M, \emptyset, u) \;\middle|\; \begin{array}{c} u \in W^{3,2}(S, Q) \\ \dots \\ [u] = [\mathbb{CP}^1 \times \{\text{pt}\}] \end{array} \right\} \Big/ \begin{array}{c} (S, j, M, u) \\ \sim (S', \psi^* j, \psi^{-1}(M), u \circ \psi) \end{array}$$

$$\cong \left\{ (S^2, i, \{z_0\}, \emptyset, u) \,\middle|\, \begin{array}{c} u \in W^{3,2}(S^2, Q) \\ [u] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}] \end{array} \right\} \Big/ \sim$$

$$\cong \left\{ u \in W^{3,2}(S^2, Q) \mid [u] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}] \right\} / \sim \;\; = Z. \qquad \square$$

3.2 Trivial Isotropy

In this section, we show that all (not necessarily pseudoholomorphic) maps in the space Z defined in (8) have trivial isotropy due to their specific homology class. We start by showing that maps of nontrivial finite energy have finite isotropy groups for any compact domain and target.

Lemma 3.2 *Let (Q, ω) be any symplectic manifold, (Σ, j) a compact connected Riemann surface, and let $u : \Sigma \to Q$ be a C^1-map with positive symplectic area $\int_\Sigma u^*\omega > 0$. Then it has a finite isotropy group*

$$G_u := \{\psi : (\Sigma, j) \to (\Sigma, j) \text{ biholomorphic} \mid u \circ \psi = u\}.$$

Proof Having positive symplectic area implies that there exists an open ball $B \subset \Sigma$, so that u is injective on B. Indeed, since $\int u^*\omega > 0$ there must be a point $p \in \Sigma$ such that $(u^*\omega)_p$ does not vanish as a bilinear form on $\mathrm{T}_p\Sigma$, i.e. there are vectors $v, w \in \mathrm{T}_p\Sigma$ with $(u^*\omega)_p(v, w) > 0$. This is equivalent to $\omega_{u(p)}(\mathrm{d}u(p)(v), \mathrm{d}u(p)(w)) > 0$. Since ω is skew-symmetric, we know that $\mathrm{d}u(p)(v), \mathrm{d}u(p)(w)$ are linearly independent. Thus $\mathrm{d}u(p)$ has maximal rank 2, and we can find a ball B around p such that $u|_B$ is injective and $\int_B u^*\omega > 0$.

Next, we claim that the images $g(B)$ of B under the automorphisms $g \in G_u$ are all disjoint. Assume this is not the case. Then there exists a $g \in G_u \setminus \{\mathrm{id}\}$ such that $B \cap g(B) \supset U$ contains a nonempty open set U. For every $p \in U$ we have $p = g(q_p)$ for some $q_p \in B$. Since $u \circ g = u$ we have $u(p) = u(q_p)$, so that the injectivity of $u|_B$ implies that $p = q_p$. This shows $g|_U \equiv \mathrm{id}$, so that unique continuation for biholomorphisms on the connected surface Σ implies $g = \mathrm{id}$, contradicting the assumption.

Therefore, $\bigcup_{g \in G_u} g(B)$ is a disjoint union of open sets, and each restriction $u|_{g(B)}$ has the same positive energy

$$\int_{g(B)} u^*\omega = \int_B g^*u^*\omega = \int_B u^*\omega =: \delta > 0.$$

If $u^*\omega$ is everywhere non-negative, this implies that G_u cannot have more than $\delta^{-1} \int_\Sigma u^*\omega$ elements. This is the case for u being pseudoholomorphic. To prove finiteness of G_u in general, we pick metrics on Σ and Q with respect to which $\mathrm{d}u$ and ω are bounded. Then we have $\int_{g(B)} u^*\omega \leq C \operatorname{Vol}(g(B))$ for some constant

$C > 0$, and hence $\mathrm{Vol}(g(B)) \geq \frac{\delta}{C}$. Since the total volume of Σ is finite, and the sets $g(B)$ are disjoint, this implies that G_u must be finite. □

For spheres in the specific homology class $[u] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}]$ in $Q = \mathbb{CP}^1 \times T$ we can extend this argument to show that the isotropy groups are in fact trivial.

Lemma 3.3 *If $[u] \in Z$, then $G_u = \{\mathrm{id}\}$.*

Proof The Sobolev embedding $W^{3,2}(S^2, Q) \subset \mathcal{C}^1(S^2, Q)$ and Lemma 3.2 imply that elements of Z have finite isotropy groups. To prove that they are trivial, we consider $u \in W^{3,2}(S^2, Q)$ with finite but nontrivial isotropy group $G_u \neq \{\mathrm{id}\}$ and we will show that $[u] \neq [\mathbb{CP}^1 \times \{\mathrm{pt}\}]$, and thus $[u] \notin Z$.

Note that $G_u \subset \mathrm{Aut}(S^2, i) = \mathrm{PSL}(2, \mathbb{C})$ is a subgroup of the Möbius group. Since G_u is finite, it must consist of elements of finite order. Möbius transformations are classified into parabolic, elliptic and hyperbolic/loxodromic ones, corresponding to their geometric and algebraic properties.[10] The only Möbius transformations of finite order $k > 1$ are elliptic ones corresponding to a rotation by angle $\frac{2\pi}{k}$ around two different fixed points in the extended complex plane. Since G_u is assumed to be nontrivial, it must contain some such rotation $f \in \mathrm{Aut}(S^2, i)$. We can moreover choose a biholomorphism ψ of $\mathbb{C} \cup \{\infty\}$ that maps the fixed points of the rotation to 0 and ∞. Then the map $u' := u \circ \psi^{-1}$ represents the same homology class as u and its isotropy group contains $g := \psi \circ f \circ \psi^{-1}$, which is a rotation of order $k > 1$ fixing 0 and ∞. Thus, $g : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ is given by $g(z) = e^{2\pi i/k} z$, and we have $u'(re^{i\theta + m \cdot 2\pi i/k}) = u'(re^{i\theta})$ for all $m \in \mathbb{Z}$ since $g^m \in G_{u'}$. This allows us to factorize $u' = v \circ \rho_k$ with $v(re^{i\theta}) := u'(re^{i\theta/k})$ and $\rho_k(re^{i\theta}) := re^{ki\theta}$ for $r \in (0, \infty)$, $\theta \in [0, 2\pi]$. By identifying $S^2 \cong \mathbb{C} \cup \{\infty\}$, this defines continuous maps $v : S^2 \to Q$ and $\rho_k : S^2 \to S^2$ with $v(0) = u(0)$, $v(\infty) = u(\infty)$, $\rho_k(0) = 0$, and $\rho_k(\infty) = \infty$. Finally, this implies $[u] = [u'] = \deg(\rho_k) \cdot [v] \in H_2(Q)$, where $\deg(\rho_k) = k > 1$, because ρ_k is a k-fold cover of S^2. This contradicts $[u] = [\mathbb{CP}^1 \times \{\mathrm{pt}\}]$, since $\frac{1}{k}[\mathbb{CP}^1] \in H_2(\mathbb{CP}^1)$ is not representable by a map $\mathrm{pr}_{\mathbb{CP}^1} \circ v : S^2 \to \mathbb{CP}^1$. □

3.3 The Base Space

Within this section, we explain how to equip the base space $\mathcal{B}$ defined in (4) (and thus also $[0, 1] \times \mathcal{B}$—see Corollary 3.9) with a polyfold structure. Since the isotropy is trivial by Lemma 3.3, this means we give $\mathcal{B}$ an M-polyfold structure, as discussed in Sect. 1.1. In fact, due to the absence of nodal maps, we can specify this further to an atlas of local homeomorphisms to open subsets of sc-Hilbert spaces, whose transition maps are sc-smooth.

[10] See for example the lecture notes [16] for a detailed geometric description of the Möbius group, especially [16, Cor. 12.1] for the statement about elements of finite order.

Remark 3.4 Before stating this result rigorously, we need to introduce one more piece of polyfold notation (also see [3, §4.1]). Every polyfold (and thus also M-polyfold or sc-Hilbert manifold) Z contains a dense subset $Z_\infty \subset Z$ of so-called smooth points and a nested sequence of subsets

$$Z_\infty \subset \ldots \subset Z_{k+1} \subset Z_k \subset \ldots Z_0 = Z.$$

Each of these is equipped with its own metrizable topology, so that, in particular, the inclusion maps $Z_{k+1} \hookrightarrow Z_k$ are continuous.

In most applications, the smooth points $z \in Z_\infty$ are the smooth maps modulo reparameterization, whose domains may be nodal. For the Gromov-Witten polyfold Z in Eq. (8), the points $[u] \in Z_k$ are given by maps $u : S^2 \to Q$ of class $W^{3+k,2}$, and Z_k is equipped with the quotient of the $W^{3+k,2}$ -topology. Correspondingly, Z_∞ consists of the equivalence classes of smooth maps.

Theorem 3.5 *After replacing the base space $\mathcal{B}$ with an open neighborhood $\mathcal{B}' \subset \mathcal{B}$ of the smooth points $\mathcal{B}_\infty = \mathcal{B} \cap Z_\infty$, it carries the natural structure of a sc-Hilbert manifold and thus of an M-polyfold.*

Remark 3.6 In practice, we expect $\mathcal{B}' = \mathcal{B}$ by an analogue of the estimates in [10, Theorems 3.8, 3.10], which guarantee that charts constructed on neighbourhoods of smooth points $b \in \mathcal{B}_\infty := \mathcal{B} \cap Z_\infty$ cover all of $\mathcal{B}$.

We need to center charts at smooth points in both proof approaches that we will present. Since we avoid all avoidable estimates, this proves a sc-Hilbert structure on an open neighbourhood $\mathcal{B}' \subset \mathcal{B}$ of $\mathcal{B}_\infty$. That is, $\mathcal{B}'$ contains all equivalence classes of smooth maps $u : S^2 \to Q$. (In fact, $\mathcal{B}'$ contains a $W^{3,2}$-neighbourhood of each such smooth point $[u]$). However, we may have $[v] \in \mathcal{B} \setminus \mathcal{B}'$ for some $v \in W^{3,2} \setminus \mathcal{C}^\infty$. While this makes the description of the base space less explicit, it does not affect the rest of the proof.

In fact, we could even allow for a more drastic restriction of the base space $\mathcal{B}' \subset \mathcal{B}$ as follows. Proving Theorem 3.5, by applying [4, Thm. 5.10 (I)] directly, establishes an M-polyfold structure on a Z_1-open subset $\mathcal{B}' \subset \mathcal{B} \cap Z_1$ such that $\mathcal{B} \cap Z_\infty \subset \mathcal{B}'$. This would remove all maps in $W^{3,2} \setminus W^{4,2}$ from $\mathcal{B}'$ and guarantee only that $\mathcal{B}'$ contains a $W^{4,2}$-neighbourhood of smooth maps.

However, the moduli spaces $\mathcal{M}_t$ and their perturbations automatically lie in the ∞-level $\mathcal{B}_\infty$ due to the regularizing property of sc-Fredholm sections [9, Def.3.8]. So neither a shift to Z_1-topology nor restricting to a neighbourhood of $\mathcal{B}_\infty$ affects how we can use the M-polyfold regularization scheme (see Theorem 1.7) in the proof of Theorem 1.1.

There are several ways to prove Theorem 3.5. We will first explain how the natural[11] polyfold structure for $Z \subset Z^{\mathrm{HWZ}}$ constructed in [10] induces a polyfold

[11] The choices made in the general construction for Gromov-Witten spaces, of a "gluing profile" and sequence $2\pi > \ldots \delta_{m+1} > \delta_m > \ldots > 0$ of exponential weights, turn out to be irrelevant in our special case due to the absence of nodes.

structure on its subset $\mathcal{B} = \mathrm{ev}^{-1}(\{p_0\}) \subset Z$. This proof applies an implicit function theorem by Filippenko [4] to the evaluation map $\mathrm{ev} : Z \to Q$.

Proof of Theorem 3.5 by Implicit Function Theorem In [10], the space Z^{HWZ} and thus also its (open and closed) component Z is given a polyfold structure. Following the proof in [10], one sees that for the component Z considered here (i.e. genus 0, homology class $[\mathbb{CP}^1 \times \{\mathrm{pt}\}]$), the model spaces are sc-Hilbert spaces (i.e. all retractions are identity maps). Their scale structure corresponds to the dense subsets $Z_m = \{[u] \in Z \mid u \in W^{3+m,2}(S^2, Q)\} \subset Z$. Moreover, by Lemma 3.3 we have trivial isotropy. So [10] actually constructs Z as a sc-Hilbert manifold.

Now we will use the description of $\mathcal{B} = \mathrm{ev}^{-1}(\{p_0\})$ in Eq. (9) as a preimage of $p_0 \in Q$ under the evaluation map $\mathrm{ev} : Z \to Q, [u] \mapsto u(z_0)$. Note that the evaluation map is classically smooth on each level $Z_m \subset Z$, so it is sc-smooth[12] by Hofer et al. [9, Cor. 1.1]. Furthermore, [4, §5.1] explains why the evaluation map is transverse (in the sense of [4, Def. 5.9]) to every submanifold of Q, so in particular transverse to $\{p_0\} \subset Q$. Now we can apply [4, Thm. 5.10 (I)] to deduce that $\mathcal{B} = \mathrm{ev}^{-1}(\{p_0\})$ is an M-polyfold. Though, strictly speaking, [4] constructs an M-polyfold structure on an open neighbourhood of the ∞-level $\mathcal{B}_\infty := \mathcal{B} \cap Z_\infty$ in the 1-level $\mathcal{B} \cap Z_1$. The proof of nonsqueezing could work with such a neighbourhood (see Remark 3.6), but we will argue that the proof of [4, Thm. 5.10(I)] in our specific setting does not actually require a shift in the topology—just a restriction to a neighbourhood $\mathcal{B}' \subset \mathcal{B}$ of $\mathcal{B}_\infty$.

By Filippenko [4, Rmk.1.4 (ii)], the linear model for $\mathcal{B}$ near $[u] \in \mathcal{B}$ is given by the kernel of the differential $\mathrm{D\,ev}([u]) : T_{[u]}Z \to T_{p_0}Q$, which we need to equip with a sc-structure. This yields two reasons for the restrictions in [4]: First, the tangent space $T_{[u]}Z$ of an M-polyfold carries a sc-structure only at $[u] \in Z_\infty$. This is why [4] builds charts centered at smooth points only, and so do we. Second, the differential of a general sc-smooth map $Z \to Q$ is defined only at points in the 1-level Z_1. In our case, the evaluation map ev is classically smooth on Z_0, so there is no need to restrict to $[u] \in Z_1$ when using the differential $\mathrm{D\,ev}([u])$.

Moreover, the construction of local charts for $\mathcal{B}$ uses a local submersion normal form, which [4, Lemma 2.1] guarantees only if the map is $\mathcal{C}^1$ (see [4, Rmk.1.4 (iii)]). Since [4] considers general sc-smooth maps $Z \to Q$, this requires a restriction to the $\mathcal{C}^1$-map $Z_1 \to Q$. In our case there is no need for this restriction, since the evaluation map $\mathrm{ev} : Z_0 \to Q$ is $\mathcal{C}^1$ without a shift.

So the charts for $\mathcal{B}$ constructed by Filippenko [4] are homeomorphisms between open neighbourhoods of smooth points $[u] \in \mathcal{B}_\infty$ and open subsets of the kernel of the differential $\mathrm{D\,ev}([u])$. (No retractions appear here due to the absence of nodes in $\mathcal{B}$.) Since these kernels are sc-Hilbert spaces, this induces a sc-Hilbert manifold structure on the subset $\mathcal{B}' \subset \mathcal{B}$ that is covered by the charts. □

[12] The manifold Q is finite-dimensional and so carries the constant scale structure where $Q_m = Q$ for all m.

Another way to construct the sc-Hilbert manifold structure on $\mathcal{B}$ is to directly incorporate the condition $u(z_0) = p_0$ in the construction of charts from [10]. Going through this proof should also serve to illuminate the general approach of [10] in this simplified setting.

Proof of Theorem 3.5 by Construction of Charts To begin, one needs to check that $\mathcal{B}$ is a metrizable space. In general, this is proven in [10, Thm.3.27] using the Urysohn criteria (Hausdorff, second countable, and completely regular) that imply metrizability. These criteria are easily checked in our setting: First note that the topology on $W^{3,2}(S^2, Q)$ can be obtained by viewing it as a subset of the Hilbert space $W^{3,2}(S^2, \mathbb{R}^N)$ via some choice of embedding $M \subset \mathbb{R}^N$. Now $\mathcal{B} = \widehat{\mathcal{B}}/\mathrm{Aut}(S^2, i, z_0)$ is the quotient of a subset $\widehat{\mathcal{B}} \subset W^{3,2}(S^2, Q)$ (given by specifying the homology class and value $u(z_0) = p_0$), modulo reparameterization by the biholomorphisms in $\mathrm{Aut}(S^2, i, z_0)$ that fix $z_0 \in S^2$. The relative topology on the subset $\widehat{\mathcal{B}}$ is automatically metric (thus Hausdorff) and second countable, however not all these properties are inherited by the quotient. The metric induces a pseudometric on the quotient, which implies that the quotient topology is completely regular. To show that the quotient is Hausdorff and second countable, it suffices to prove that the quotient map $\pi : \widehat{\mathcal{B}} \to \mathcal{B}$ is open. To check this we consider any open subset $\widehat{U} \subset \widehat{\mathcal{B}}$ and show that $\pi(\widehat{U})$ is open by checking that any $u \in \pi^{-1}(\pi(\widehat{U}))$ has an open neighbourhood contained in $\pi^{-1}(\pi(\widehat{U}))$. Note that $u \circ \psi = \widehat{u} \in \widehat{U}$ for some $\psi \in \mathrm{Aut}(S^2, i, z_0)$ and $\{w \in \widehat{\mathcal{B}} \,|\, \|w - \widehat{u}\|_{W^{3,2}} < \varepsilon\} \subset \widehat{U}$ for some $\varepsilon > 0$ since $\widehat{U}$ is open. Now let $C > 0$ be a constant bounding all derivatives of ψ up to third order. Then we claim that the $\frac{\varepsilon}{C}$ ball around u provides the required open neighbourhood, i.e. is contained in $\pi^{-1}(\pi(\widehat{U}))$. Indeed, $\|v - u\|_{W^{3,2}} < \frac{\varepsilon}{C}$ implies $\|v \circ \psi - \widehat{u}\|_{W^{3,2}} < \varepsilon$, thus $v \circ \psi \in \widehat{U}$ and $v \in \pi^{-1}(\pi(\widehat{U}))$. This proves that the quotient map is open and thus finishes the proof of metrizability.

Next, the main technical work is the construction of a chart for a neighbourhood of a given point $\alpha \in \mathcal{B}$. For that purpose we pick a representative $u : S^2 \to Q$ of $\alpha = [u]$. Since $\mathcal{B}$ is a quotient by the reparameterization action, the charts are constructed as local slices to this action, which involves choices of additional marked points and transverse constraints. More precisely, we need to choose **good data centered at** u in the sense of [10, Def. 3.6]. Such good data exists by Hofer et al. [10, Prop. 3.7]. Recall here that the isotropy group $G = G_u = \{\mathrm{id}\}$ is trivial in our case. Then good data in our setting consists of the following objects with the following properties[13]

1. **Marked points stabilizing the surface:** These exist by Hofer et al. [10, Lemma 3.2]. In our case a stabilization of $(S^2, i, \{z_0\}, \emptyset)$ for the map u consists of two points $\Sigma = \{z_1, z_2\} \subset S^2$ that satisfy the following conditions:
   - $z_0, z_1, z_2 \in S^2$ are pairwise different.

[13] We essentially use the numbering of [10, Def. 3.6], but merged (5) into 8.), merged (7) into 2.) and 4.), merged (8) into 7.), and left out (9), (10) which are trivially satisfied in our case.

- Denote $p_1 := u(z_1)$, $p_2 = u(z_2)$. Then p_0, p_1, p_2 are pairwise different.
- For $i = 1, 2$, the map $\mathrm{d}u(z_i)$ is injective, the bilinear form $u^*\omega(z_i)$ is non-degenerate, and it determines the correct orientation on $\mathrm{T}_{z_i} S^2$.

Now the Riemann surface with the additional marked points $(S^2, i, \{z_0, z_1, z_2\})$ is stable; in fact its isotropy group is trivial in our case. [10] also requires a choice of good uniformizing family parameterizing variations of the surface and marked points, but in our case, since the Deligne-Mumford space of three marked points on a sphere is trivial, this family is constant. It remains to choose **small disk structures**, that is holomorphic embeddings of the closed disk $D^2 \simeq D_{z_i} \subset S^2$ with center $0 \simeq z_i$ that are disjoint for $i = 1, 2$.

2. **Charts for the target space:** Open neighbourhoods $U(p_i) \subset Q$ of p_i for $i = 1, 2$ and diffeomorphisms $\psi_i : (U(p_i), p_i) \to (\mathbb{R}^{2n}, 0)$ are chosen so that

   - $U(p_1)$ and $U(p_2)$ are disjoint;
   - $u|_{D_{z_i}}$ is an embedding for $i = 1, 2$, with image contained in $U_1(p_i)$.

   Here and in the following we denote by $U_\rho(p_i) := \psi_i^{-1}(\{x \in \mathbb{R}^{2n} \,|\, |x| < \rho\})$ the preimages of balls of any radius $\rho > 0$.
3. **A Riemannian metric** on Q is chosen such that it agrees on the open sets $U_4(p_i)$ with the pullback of the standard metric on $\mathbb{R}^{2n}$ by ψ_i.
   Moreover, we choose an **open neighbourhood of the zero-section** $\widetilde{\mathcal{O}} \subset \mathrm{T}Q$, which is fiberwise convex and such that for every $q \in Q$, the exponential map induced by the chosen metric, $\exp : \widetilde{\mathcal{O}}_q := \widetilde{\mathcal{O}} \cap \mathrm{T}_q Q \to Q$ is an embedding.
4. **Transverse hypersurfaces:** We choose submanifolds $M_{p_i} \subset U(p_i) \subset Q$ of codimension 2 for $i = 1, 2$ such that $p_i \in M_{p_i}$ and

   - $\psi_i(M_{p_i}) \subset \mathbb{R}^{2n}$ is a linear subspace;
   - $\mathrm{T}_{p_i} Q = \operatorname{im} \mathrm{d}u(z_i) \oplus H_{p_i}$ for $H_{p_i} := \mathrm{T}_{p_i} M_{p_i}$;
   - $\{z_i\} = D_{z_i} \cap u^{-1}(M_{p_i})$ is the only point in D_{z_i} that u maps to M_{p_i}.

   This in particular implies that $u|_{D_{z_i}}$ is transverse to M_{p_i} for $i = 1, 2$.
6. **Concentric subdisks:** We choose $SD_{z_i} \subset D_{z_i}$ for $i = 1, 2$ to be the image of smaller disks under the holomorphic embeddings $D^2 \simeq D_{z_i}$.
7. **An open neighbourhood** $U \subset W^{3,2}(S^2, u^*\mathrm{T}Q)$ of 0 is chosen such that

   - every section $\eta \in U$ takes values in $u^*\widetilde{\mathcal{O}}$;
   - for every $\eta \in U$ and $i = 1, 2$, the map $u' := \exp_u(\eta) : S^2 \to Q$ satisfies $u'(D_{z_i}) \subset U_2(p_i)$, and $u'|_{D_{z_i}}$ is an embedding transverse to M_{p_i} that intersects M_{p_i} at a single point $p_i' = u'(z_i')$, the preimage of some $z_i' \in SD_{z_i}$.

From now on we will assume that we chose good data for a smooth map $u \in \mathcal{C}^\infty$, since it suffices to construct charts centered at smooth points $\alpha \in \mathcal{B}_\infty$.[14] Now the

[14] Hofer et al.[10, Thms.3.8, 3.10] imply that for every $[u'] \in Z$ there is a smooth $u : S^2 \to Q$ and good data centered at u such that $u' = \exp_u(\eta)$ for some $\eta \in U$. This means that the charts coming from good data centered at the smooth points of Z will cover all of Z. This can be extended to include the conditions $u'(z_0) = p_0$ and $\eta(z_0) = 0$, and thus prove the same for $\mathcal{B}$.

only place where our constructions for $\mathcal{B} = \{[u'] \,|\, u'(z_0) = p_0\} \subset Z$ differ from the constructions in [10] for Z is the definition of a linear subspace of the sc-space $\left(W^{3+k,2}(S^2, u^*\mathrm{T}Q)\right)_{k\in\mathbb{N}_0}$. Our choice

$$E_u := \{\eta \in W^{3,2}(S^2, u^*\mathrm{T}Q) \mid \eta(z_0) = 0, \eta(z_i) \in H_{p_i} \text{ for } i = 1, 2\} \tag{10}$$

adds the condition $\eta(z_0) = 0$ that linearizes the condition $u'(z_0) = p_0 = u(z_0)$. We then consider its open subset $\mathcal{O} := E_u \cap U$ and claim that the map

$$\mathcal{O} \to \mathcal{B}, \quad \eta \mapsto [\exp_u \eta] \tag{11}$$

is a homeomorphism onto a neighbourhood of $[u]$. While this is not explicitly stated in [10, §3.1], we can check the properties of this map in our case:

Continuity Equation (11) is the composition of two continuous maps: the pointwise exponential map and a quotient projection. The latter is continuous by definition of the quotient topology on $\mathcal{B}$. The first map is the composition $\eta \mapsto E \circ \eta$ with the smooth exponential map $E : u^*\mathrm{T}Q \to Q$. Checking that this is continuous between $W^{3,2}$-topologies requires local estimates, which hold since $3 \cdot 2 > \dim S^2$ (see e.g. [19, Lemma B.8]).

Injectivity Consider $\eta, \widetilde{\eta} \in \mathcal{O}$ with $[\exp_u \eta] = [\exp_u \widetilde{\eta}] \in \mathcal{B}$. This means that there is a biholomorphism $\psi : S^2 \to S^2$ with $\psi(z_0) = z_0$ and $\exp_u \eta = \exp_u \widetilde{\eta} \circ \psi$. By property 4.), the maps $\exp_u \eta$ and $\exp_u \widetilde{\eta}$ intersect the submanifolds M_{p_i}, $i = 1, 2$ in unique points $p_i', \widetilde{p}_i \in M_{p_i}$ with unique preimages $z_i', \widetilde{z}_i \in SD_{z_i}$. It follows that $\psi(z_i') = \widetilde{z}_i$ for $i = 1, 2$. On the other hand, $\eta(z_i), \widetilde{\eta}(z_i) \in H_{p_i} = \mathrm{T}_{p_i} M_{p_i}$ and the fact that $M_{p_i} \subset U(p_i)$ is totally geodesic imply $\exp_u \eta \in M_{p_i}$ and $\exp_u \widetilde{\eta} \in M_i$, so the uniqueness of intersection points in property 7.) implies $z_i' = z_i$ and $\widetilde{z}_i = z_i$ for $i = 1, 2$. Thus we have $\psi(z_i) = z_i$ for $i = 0, 1, 2$, which implies $\psi = \mathrm{id}_{S^2}$. From this we deduce $\exp_u \eta = \exp_u \widetilde{\eta}$ and thus $\eta = \widetilde{\eta}$.

Continuity of the Inverse To show that the map in (11) is a homeomorphism, it remains to check that it maps open subsets of E_u to open subsets of $\mathcal{B}$. So we fix some $\eta \in \mathcal{O}$ and need to show that any $\alpha \in \mathcal{B}$ sufficiently close to $[\exp_u \eta]$ can be written as $\alpha = [\exp_u \eta']$ for some $\eta' \in \mathcal{O}$. First note that $v := \exp_u \eta$ satisfies $v(z_0) = p_0$ along with the slicing conditions $v(z_i) \in M_{p_i}$ for $i = 1, 2$. Next, α being close to $[v]$ in the quotient topology of $\mathcal{B}$ means that $\alpha = [v']$ for some representative v' that is $W^{3,2}$-close to v. By construction of $\mathcal{B}$, this map satisfies $v'(z_0) = p_0$. Moreover, since v is locally transverse to the slicing conditions (as specified in 4.), we will have $v'(z_i') \in M_{p_i}$ for some $z_i' \approx z_i$. Now we can compose v' with a small Möbius transformation that fixes z_0 and maps z_i to z_i' to obtain a new representative $\alpha = [w]$ that satisfies $w(z_0) = p_0$ and $w(z_i) \in M_{p_i}$. This adjustment in slicing conditions guarantees that $\eta'(z) := \exp_{u(z)}^{-1}(w(z))$ defines a section $\eta' \in E_u$. Moreover, the construction is done such that w is still $W^{3,2}$-close

to v, which guarantees $\eta' \approx \eta$ and thus $\eta' \in \mathcal{O}$ for α sufficiently close to $[v]$. This proves openness and thus continuity of the inverse.

Thus we have constructed $\mathcal{C}^0$-charts for $\mathcal{B}$ centered at any point $[u]$. It remains to equip the local models with sc-structures and show that the transition maps between these charts are sc-smooth. This is where we (just as [10, §3.2]) have to restrict to centering our charts at points $[u] \in \mathcal{B}_\infty$ represented by smooth maps $u : S^2 \to Q$. This regularity is required to give the pullback bundle $u^*\mathrm{T}Q$ a classically smooth structure, so that we can define the Sobolev spaces $W^{3+k,2}(S^2, u^*\mathrm{T}Q)$ of sections by closure of the smooth sections. Then we obtain a sc-structure on the Hilbert space E_u in (11) by the subspaces $E_u \cap W^{3+k,2}(S^2, u^*\mathrm{T}Q)$ for $k \in \mathbb{N}_0$.

Finally, compatibility of the charts follows directly from [10], since our transition maps are the same maps as theirs, just restricted to subsets of their domains. More precisely, transition maps between different charts for Z can be obtained from its ep-groupoid description by composing local inverses of the source map with the target map. Both of these structure maps are local sc-diffeomorphisms by the étale property in [9, Def.7.3] (for the Gromov-Witten case this is established in [10, Prop.3.19]). So sc-smoothness of transition maps follows since they are compositions of local sc-diffeomorphisms. □

Remark 3.7 (sc-Smoothness) The key point why sc-smoothness appears here is the following: The choice of stabilization points z_1, z_2 above depended on u. For different u' we might have other z_1', z_2', and so the transition map between the charts centered at u and u' needs to reparameterize all the vector fields η. But the reparameterization biholomorphism is not fixed for one transition map, but depends also on the vector field. So we get a map of the form

$$\mathcal{O} \ni \eta \longmapsto \eta \circ \mathrm{d}\psi_\eta \in \mathcal{O}',$$

which is not classically differentiable w.r.t. any of the usual Sobolev or Hölder norms. It is however scale-smooth; see [3, §2.1,2.2] for further discussion.

Remark 3.8 (Good Data Centered at u_0) For the holomorphic map $u_0(z) = (z, m_0)$ from Sect. 2.2, it is easy to find good data: In the first factor, u_0 is the identity, and in the second factor it is the constant map to $m_0 \in T$. For this map, any choice of two different points $z_1, z_2 \in S^2 \setminus \{z_0\}$ yields a stabilization as required. Then $M_{p_i} := \{z_i\} \times T$ are transverse hypersurfaces since their tangent spaces $H_{p_i} = \{0\} \times \mathrm{T}_{m_0}T$ satisfy

$$\operatorname{im} \mathrm{d}u_0(z_i) \oplus H_{p_i} = \mathrm{T}_{z_i}\mathbb{CP}^1 \oplus \mathrm{T}_{m_0}T = \mathrm{T}_{u_0(z_i)}Q.$$

This choice of charts will be used in Sect. 3.7. In particular, we will use

$$E_{u_0} := \{\eta \in W^{3,2}(S^2, u^*\mathrm{T}Q) \mid \eta(z_0) = 0, \eta(z_i) \in \{0\} \times \mathrm{T}_{m_0}T \text{ for } i = 1, 2\}. \tag{12}$$

Corollary 3.9 *The product* $[0,1] \times \mathcal{B}$ *is a sc-Hilbert manifold with boundary* $\partial([0,1] \times \mathcal{B}) = \{0,1\} \times \mathcal{B}$.

Proof As a finite dimensional manifold, the interval $[0,1]$ is trivially a sc-Hilbert manifold and the notion of boundary is the same as the classical notion (see also Remark 1.6). This means that the product $[0,1] \times \mathcal{B}$ also is a sc-Hilbert manifold. In fact, for every pair $(t_0, [u]) \in [0,1] \times \mathcal{B}$ we can choose an open interval $I_{t_0} \subseteq [0,1]$ containing t_0 and a chart $\mathcal{O} \to \mathcal{B}$ centered at $[u]$ as in (11), then a sc-Hilbert manifold chart for a neighbourhood of $(t_0, [u])$ is given by

$$I_{t_0} \times \mathcal{O} \longrightarrow [0,1] \times \mathcal{B}, \qquad (t, \eta) \longmapsto (t, [\exp_u \eta]).$$

Sc-smoothness of transition maps between these charts follows directly from sc-smooth compatibility of the charts for $\mathcal{B}$. Finally, $\mathcal{O}$ has no boundary, and boundary of I_{t_0} arises only from $\partial[0,1] = \{0,1\}$. Since boundary and corner stratifications are determined in local charts, this proves the claim. □

3.4 The Bundle

The purpose of this section is to give the projection $\mathcal{E} \to [0,1] \times \mathcal{B}$ defined in (6) the structure of a tame strong polyfold bundle. To achieve this, we first describe the bundle $W \to Z^{\mathrm{HWZ}}$ from [10] restricted to $Z \subset Z^{\mathrm{HWZ}}$.

The fibers of W are defined in [10, §1.2] with respect to a fixed almost complex structure J on the target manifold. Using the previous simplifications for $Z \subset Z^{\mathrm{HWZ}}$ (i.e. genus 0, $S = S^2$, homology class $[\mathbb{CP}^1 \times \{\mathrm{pt}\}]$) we see that the fiber W_α over $\alpha \in Z$ is the quotient space

$$W_\alpha = \left\{ (u, \eta) \;\middle|\; \begin{array}{c} [u] = \alpha \\ \eta \in \Lambda^{0,1}_J\big(S^2, u^*\mathrm{T}(\mathbb{CP}^1 \times T)\big) \text{ of class } W^{2,2} \end{array} \right\} \Big/ \sim$$

of complex antilinear 1-forms of class $W^{2,2}$ with values in the pullback bundle along a representative u. Here complex antilinearity of η is required with respect to a fixed almost complex structure J on $\mathbb{CP}^1 \times T$ and the standard complex structure on S^2. The equivalence relation is given by

$$(u, \eta) \sim (v, \mu) \;:\Longleftrightarrow\; \begin{array}{c} \exists \psi : S^2 \to S^2 \text{ holomorphic} \\ \text{with } u = v \circ \psi \text{ and } \eta = \mu \circ \mathrm{d}\psi. \end{array}$$

For $\alpha \in \mathcal{B} \subset Z$ we can choose a representative $[u] = \alpha$ with $u(z_0) = p_0$, and restricting to such representatives reduces the equivalence relation to biholomorphisms ψ with $\psi(z_0) = z_0$. So if we choose $J = J_t$ for some $t \in [0,1]$, then this identifies our fibers $\mathcal{E}_{(t,\alpha)} = W_\alpha$ with the fibers of a tame strong polyfold bundle constructed in [10]. However, we wish to simultaneously extend and restrict

the base space: We extend by allowing the almost complex structure to vary, and we restrict to curves through the fixed point $p_0 \in Q$.

Theorem 3.10 *Let $\mathcal{B}' \subset \mathcal{B}$ be as in Theorem 3.5,*[15] *then $\mathcal{E}|_{[0,1]\times\mathcal{B}'} \to [0,1] \times \mathcal{B}'$ is a tame strong M-polyfold bundle.*

There are again several ways to prove this. Filippenko [4] has a result about the restriction of bundles to sub-polyfolds which we explain in Remark 3.13. Unfortunately, this would require an existing polyfold description for Gromov-Witten moduli spaces with varying J, which we discuss in Remark 3.11. In any case, the proof of transversality of the section σ at $t = 0$ requires a fairly explicit bundle chart, so our actual proof is an adaptation of [10]—to our simplified setting, but extending the constructions to varying J.

Remark 3.11 The analysis in [10] is formulated for a fixed almost complex structure J on the target manifold Q. This might be deemed as sufficient due to the following graph trick:

Given a smooth family $(J_t)_{t\in\mathbb{R}^k}$ of almost complex structures, parameterized by a finite dimensional space $\mathbb{R}^k$, we can identify the moduli space of J_t-holomorphic curves in Q for some $t \in \mathbb{R}^k$ with a moduli space of certain pseudoholomorphic curves in the product manifold $\widetilde{Q} := \mathbb{C}^k \times Q$ as follows. We define an almost complex structure $\widetilde{J}$ on $\widetilde{Q}$ by $\widetilde{J}(t+is, q) := J_{\mathbb{C}^k} \times J_t$ at the point $(t+is, q) \in \widetilde{Q} = \mathbb{C}^k \times Q$. Then $\widetilde{J}$-holomorphic maps $\widetilde{u} : S \to \widetilde{Q} = \mathbb{C}^k \times Q$ in a class $\widetilde{A} := [\{\text{pt}\}] \times A$, $A \in H_2(Q)$ are constant in $\mathbb{C}^k$, so that a constraint $\widetilde{u}(z_0) \in \mathbb{R}^k \times Q$ at a marked point $z_0 \in S$ picks out the maps $\widetilde{u}$ that are of the form $z \mapsto (t, u(z))$ for some $t \in \mathbb{R}^k$ and a J_t-holomorphic map $u : S \to Q$. Thus pairs (t, u) of J_t-holomorphic maps in class A for some $t \in \mathbb{R}^k$ can be identified with $\widetilde{J}$-holomorphic maps in class $\widetilde{A}$ satisfying the point constraint.

Our variation of almost complex structures $(J_t)_{t\in[0,1]}$ can be formulated in this way by choosing J_t to be constant near $t = 0$ and $t = 1$, so that its constant extension to $t \in \mathbb{R}$ is smooth. However, the polyfold bundle $\widetilde{W} \to \widetilde{Z}$ that [10] constructs for the almost complex manifold $(\widetilde{Q}, \widetilde{J})$ contains already in its base many more (reparameterization classes of) maps $\widetilde{u} : S^2 \to \mathbb{C} \times Q$ than those of the form $\widetilde{u}(z) = (t, u(z))$. The point constraint $\widetilde{u}(z_0) \in \mathbb{R} \times Q$ does not force the $\mathbb{C}$-component to be constant. Similarly, the fibers of $\widetilde{W}$, consisting of anti-linear 1-forms with values in $\mathbb{C} \times Q$, contain an extra factor $\Lambda^{0,1}(S, \mathbb{C})$ compared with the fibers W_α above that are used in the polyfold description of a moduli space for fixed J. These infinite dimensional extra factors would have to be split off near maps of the form $(0, u(z))$ or $(1, u(z))$ in order to relate the moduli spaces for J_0 and J_1 with parts of the moduli space for $\widetilde{J}$ in this setup.

Alternatively, extending such a splitting along $(t, u(z))$ for all $t \in \mathbb{R}$ would yield a smaller polyfold description in which the base space consists only of maps of the

[15] Recall we expect $\mathcal{B}' = \mathcal{B}$ by Remark 3.6.

form $(t, u(z))$. Rather than attempting such a splitting construction, we will directly construct the resulting polyfold bundle.

Proof of Theorem 3.10 We construct bundle charts as in [10, §6.3].[16] Fix a pair $(t_0, [u]) \in [0,1] \times \mathcal{B}_\infty$, a representative u of $[u]$, an open interval $I_{t_0} \subset [0,1]$ around t_0 (whose 'sufficiently small' choice will be specified below), and good data centered at u. These choices determine a chart for the base space as in Sect. 3.3, by

$$I_{t_0} \times \mathcal{O} \to [0,1] \times \mathcal{B}, \qquad (t, \eta) \mapsto (t, [\exp_u(\eta)]).$$

(See Sect. 3.3 for the definition of $\mathcal{O} = E_u \cap U \subset W^{3,2}(S^2, u^*\mathrm{T}Q)$.) To build charts for the bundle, recall that we abbreviate $Q := \mathbb{CP}^1 \times T$. Then, using the chosen representative u, the bundle fiber $\mathcal{E}_{t_0,[u]}$ can be identified with the Hilbert space

$$F := \quad W^{2,2}\text{-closure of} \quad \Lambda^{0,1}_{J_{t_0}}(S^2, u^*\mathrm{T}Q). \tag{13}$$

Here $\Lambda^{0,1}_J(\ldots)$ denotes smooth complex-antilinear 1-forms, and to take their $W^{2,2}$-closure we view them as functions on S^2 with values in a bundle whose fiber over $z \in S^2$ are the linear maps $\mathrm{T}_z S^2 \to \mathrm{T}_{u(z)} Q$. This Hilbert space can be equipped with a scale structure whose k-th level consists of forms of regularity $W^{2+k,2}$. So we can define a trivial strong bundle[17]

$$I_{t_0} \times \mathcal{O} \lhd F \to I_{t_0} \times \mathcal{O} \tag{14}$$

which we will use as local model for $\mathcal{E} \to [0,1] \times \mathcal{B}$ near the pair $(t_0, [u])$.

To trivialize the bundle, we need to map J_{t_0}-antilinear one-forms with values in $u^*\mathrm{T}Q$ to J_t-antilinear one-forms with values in $v^*\mathrm{T}Q$ for pairs (t, v) near (t_0, u). We will do this in two steps, first changing the almost complex structure and then the pullback bundle. For every $t \in I_{t_0}$ we can define a linear map from the space of antilinear 1-forms with respect to J_0 to the space of antilinear 1-forms with respect to J_t (both with values in $u^*\mathrm{T}Q$),

$$\begin{aligned} K_t \;:\; \Lambda^{0,1}_{J_{t_0}}\left(S^2, u^*\mathrm{T}Q\right) &\longrightarrow \Lambda^{0,1}_{J_t}\left(S^2, u^*\mathrm{T}Q\right) \\ \xi &\longmapsto \tfrac{1}{2}\left(\xi + J_t(u) \circ \xi \circ i\right). \end{aligned} \tag{15}$$

[16] Notation here is the same as in [10], with lots of simplifications because we work with a constant Riemann surface and trivial isotropy. On the other hand, we allow the almost complex structure to vary in a 1-parameter family, whereas [10] fixes it.

[17] As a set, $\mathcal{O} \lhd F$ is the Cartesian product $\mathcal{O} \times F$. The symbol $\lhd$ means that we consider it as a strong bundle, which is a condition on the scale structure: For $\eta \in \mathcal{O}$ on level m, i.e. of Sobolev class $W^{3+m,2}$, it makes sense to talk about sections along $u' := \exp_u \eta$ of Sobolev class up to $W^{3+m,2} = W^{2+k,2}$, i.e. up to level $k = m + 1$.

Note that K_{t_0} is the identity map. Moreover, the explicit form of K_t allows us to check that this family of operators is smooth with respect to the $W^{2+k,2}$-topology on $u^*\mathrm{T}Q$-valued 1-forms for any $k \in \mathbb{N}_0$.[18] In particular, it extends continuously to F, which is the $W^{2,2}$-closure of $\Lambda^{0,1}_{J_{t_0}}\left(S^2, u^*\mathrm{T}Q\right)$, taking values in the $W^{2,2}$-closures of $\Lambda^{0,1}_{J_t}\left(S^2, u^*\mathrm{T}Q\right)$. So by choosing I_{t_0} as sufficiently small neighbourhood of t_0 we can guarantee that K_t, after extension to the $W^{2,2}$-closures, is a linear sc-isomorphism for all $t \in I_{t_0}$.

The rest of the bundle chart construction—i.e. the step from the pullback bundle $u^*\mathrm{T}Q$ to the pullback bundle $v^*\mathrm{T}Q$—proceeds as in [10]; we just need to ensure sc-smooth dependence on the extra parameter. We construct a family of connections $\widetilde{\nabla}^t$ on $\mathrm{T}Q$ for $t \in [0,1]$ as follows:[19] Let ∇^t denote the Levi–Civita connection of the metric $g_t := \omega(\cdot, J_t\cdot)$ on Q and define a new connection by $\widetilde{\nabla}^t_X Y := \frac{1}{2}(\nabla^t_X Y - J_t\nabla^t_X(J_t Y))$. Then the family is smooth in t, and $\widetilde{\nabla}^t$ is a complex connection on the almost complex vector bundle $(\mathrm{T}Q, J_t)$, that is, it satisfies $\widetilde{\nabla}^t_X(J_t Y) = J_t(\widetilde{\nabla}^t_X Y)$. Moreover, recall that the good data used to construct the above chart for $\mathcal{B}$ included the choice of an open neighbourhood $\widetilde{\mathcal{O}}$ of the zero section of $\mathrm{T}Q$, such that $\widetilde{\mathcal{O}}$ is fiberwise convex and the exponential map (for a fixed metric on Q that does not vary with t) restricts to an embedding on each fiber $\widetilde{\mathcal{O}}_q$, $q \in Q$. Then for a tangent vector $\eta_q \in \widetilde{\mathcal{O}}_q$, consider the geodesic path $[0,1] \to Q, s \mapsto \exp_q(s\eta_q)$ from q to $p := \exp_q(\eta_q)$. Parallel transport with respect to $\widetilde{\nabla}^t$ along this path defines a J_t-complex linear map[20]

$$\Gamma_t(\eta_q) \;:\; (\mathrm{T}_q Q, J_t(q)) \longrightarrow (\mathrm{T}_p Q, J_t(p)). \tag{16}$$

This in fact is an isomorphism for each $t \in [0,1]$ and $\eta_q \in \mathrm{T}_q Q$, and these isomorphisms vary smoothly with $t \in [0,1]$, $q \in Q$ and $\eta \in TQ$.

The resulting bundle chart covering the chart (11) of the base space given by $I_{t_0} \times \mathcal{O} \to [0,1) \times \mathcal{B}, (t,\eta) \mapsto (t, [\exp_u \eta])$ is

$$\begin{aligned} I_{t_0} \times \mathcal{O} \lhd F \;&\longrightarrow\; \mathcal{E} \\ (t,\eta,\xi) \;&\longmapsto\; \Big(t, \Big[\big(\exp_u \eta\,,\; \Xi(\eta,\xi) := \Gamma_t(\eta) \circ K_t(\xi)\big)\Big]\Big). \end{aligned} \tag{17}$$

Here the complex antilinear 1-form $\Xi(\eta,\xi) \in \Lambda^{0,1}_{J_t}(S^2, v^*\mathrm{T}Q)$ with values in the pullback bundle by $v := \exp_u \eta$ is given at each $z \in S^2$ by the complex antilinear

[18] For fixed k, the continuity of K_t requires $u \in \mathcal{C}^{2+k}$, which is why we center these charts at smooth points $[u] \in \mathcal{B}_\infty$.

[19] Any smooth family of complex connections suffices for the present construction. In Sect. 3.6, however, we will need $\widetilde{\nabla}^0$ to split along the factors of $Q = \mathbb{CP}^1 \times T$, see Remark 3.12.

[20] Note here that parallel transport is defined along any path, so there is no issue with the fact that the path is induced by an exponential map for a different metric than the family of metrics used in the construction of the connection.

map

$$\Gamma_t(\eta(z)) \circ K_t(\xi)(z) \,:\, \big(\mathrm{T}_z S^2, i\big) \to \big(\mathrm{T}_{u(z)} Q, J_t(u(z))\big) \to \big(\mathrm{T}_{v(z)} Q, J_t(v(z))\big).$$

It coincides with the construction in [10, (3.9)] in case $K_t = \mathrm{id}$ due to J being fixed. We can now follow the arguments of [10, §3.6] to construct $\mathcal{E}$ as tame strong M-polyfold bundle. For that purpose first note that each map (17) covers an M-polyfold chart for $[0,1] \times \mathcal{B}' \subset [0,1] \times \mathcal{B}$, so the images of these maps cover only $\mathcal{E}|_{[0,1]\times\mathcal{B}'}$. Second, these are bundle charts in the sense that they are linear bijections on each fiber. They are local homeomorphisms because the topology on the total space $\mathcal{E}|_{[0,1]\times\mathcal{B}'}$ is defined in this manner as in [10, Thm.1.9].

Strong sc-smooth compatibility of these bundle charts is proven for fixed J in [10, Prop.3.39], using the language of [9, Prop.3.39]. The strong bundle isomorphism μ that is considered here, and proven to be a local sc-diffeomorphism, in fact encodes all transition maps between different bundle charts. The proof of [10, Prop.3.39] directly extends to the case of varying J thanks to its explicit nature (15) of K_t as family of linear 0-th order operators. Here we have to again require $u \in \mathcal{C}^\infty$ to ensure that K_t is a bounded operator on each of the $W^{2+k,2}$-scales. In fact, these operators vary smoothly with $t \in [0,1]$ on each scale. Thus the charts (17) equip $\mathcal{E}|_{[0,1]\times\mathcal{B}'}$ with the structure of a strong M-polyfold bundle. Finally, tameness is a condition on the underlying M-polyfold that is automatically satisfied in our case since all retractions are trivial; see [9, Def.2.17]. □

Remark 3.12 For $t = 0$, remember that $J_0 = i \oplus J_T$ splits along the factors of $Q = \mathbb{CP}^1 \times T$. This means that also the metric splits and so does its Levi-Civita connection ∇. Then also the connection $\widetilde{\nabla}^0$ splits, and thus the parallel transport map Γ_0 in (16) preserves the factors of $\mathrm{T}_{(z,p)} Q = \mathrm{T}_z \mathbb{CP}^1 \times \mathrm{T}_p T$.

Remark 3.13 An alternative proof of Theorem 3.10 is to construct the bundle $\mathcal{E} \to [0,1] \times \mathcal{B}$ in (6) from an implicit function theorem in [4].

This proof requires as starting point a polyfold description of the Gromov-Witten moduli space $\mathcal{M}$ for a family $(J_t)_{t\in[0,1]}$ of almost complex structures. Such a description is obtained by performing the constructions of Theorem 3.10 over $[0,1] \times Z$ to obtain a tame strong M-polyfold bundle $\widetilde{p} : \widetilde{W} \to [0,1] \times Z$ which restricts on every slice $\{t\} \times Z$ to the bundle $W \to Z$ from [10] for the almost complex structure J_t.

Now the projection $\mathcal{E} \to [0,1] \times \mathcal{B}$ in (6) is obtained from $\widetilde{W} \to [0,1] \times Z$ by restriction to $\mathcal{B} = \mathrm{ev}^{-1}(\{p_0\}) \subset Z$ as in the first proof of Theorem 3.5. Here the map $[0,1] \times Z \to Q$, $(t,[u]) \mapsto \mathrm{ev}[u] = u(z_0)$ is transverse to $\{p_0\} \in Q$ since the evaluation map—without the $[0,1]$-factor—is already transverse by Filippenko [4, §1.5]. Thus [4, Theorem 5.10 (II)] provides an open neighbourhood $\widetilde{\mathcal{B}} \subset [0,1] \times (\mathcal{B} \cap Z_1)$ of $[0,1] \times \mathcal{B}_\infty$ such that $\widetilde{p}^{-1}(\widetilde{\mathcal{B}}) = \mathcal{E}|_{\widetilde{\mathcal{B}}} \to \widetilde{\mathcal{B}}$ inherits the structure of a tame strong M-polyfold bundle. In our special case, the shift to Z_1-topology is actually not needed for the reasons already stated in Theorem 3.5, and since the retracts are all trivial. Furthermore, $\widetilde{\mathcal{B}}$ can be replaced by $[0,1] \times \mathcal{B}''$ for an open

neighbourhood $\mathcal{B}'' \subset \mathcal{B}$ of $\mathcal{B}_\infty$ that may just be somewhat smaller than $\mathcal{B}'$ from Theorem 3.10.

Indeed, $\widetilde{\mathcal{B}}$ is given by a union of charts for $[0,1] \times \mathcal{B}$ centered at smooth points, which lift to strong bundle charts. Each of these charts is constructed in product form, and since $[0,1]$ is compact we can find for every $b \in \mathcal{B}_\infty$ finitely many such product charts that cover $[0,1] \times \{b\}$. This implies $[0,1] \times \mathcal{U}_b \subset \widetilde{\mathcal{B}}$ for some open neighbourhood $\mathcal{U}_b \subset \mathcal{B}$ of b. This proves the claim with $\mathcal{B}' = \bigcup_{b \in \mathcal{B}_\infty} \mathcal{U}_b$.

3.5 The Section

This section finalizes the polyfold description of the moduli space $\mathcal{M} = \sigma^{-1}(0)$ by establishing the relevant properties of the section $\sigma : [0,1] \times \mathcal{B} \to \mathcal{E}$ introduced in (7). Up to quotienting by reparameterization in the base, its principal part (given by its values in the fibers) is $(t,u) \mapsto \overline{\partial}_{J_t} u$.

Theorem 3.14 *Let $\mathcal{B}' \subset \mathcal{B}$ be as in Theorem 3.5,*[21] *then $\sigma : [0,1] \times \mathcal{B}' \to \mathcal{E}$ is a sc-Fredholm section of Fredholm index* 1.

As for the bundle structure in Remark 3.13, the Fredholm property of the section could be proven by the restriction results of Filippenko [4]—if the Fredholm property of the Cauchy-Riemann operator with varying J was firmly established. We explain this approach in Remark 3.15, after giving a direct proof of the Fredholm property based on the explicit bundle charts in Theorem 3.10.

Proof of Theorem 3.14 We work in local coordinates centered at a pair $(t_0, [u]) \in [0,1] \times \mathcal{B}_\infty$ that were defined in (17). The principal part of the section σ is given in these coordinates by

$$f : I_{t_0} \times \mathcal{O} \longrightarrow F, \quad (t,\eta) \longmapsto K_t^{-1}\big(\Gamma_t(\eta)^{-1}(\overline{\partial}_{J_t} \exp_u \eta)\big),$$

with the family of sc-isomorphisms $K_t : F \to \Lambda^{0,1}_{J_t}\left(S^2, u^*\mathrm{T}Q\right)$ given in (15), and parallel transport $\Gamma_t(\eta)$ as in (16). Both of these are linear and explicitly given in terms of point-wise operations. Thus routine computations show that f is not just sc-smooth, but for any $k \in \mathbb{N}_0$ restricts to a classically smooth map with respect to the $W^{k+3,2}$-norm on $\mathcal{O}$ and the $W^{k,2}$-norm on F. (For the present proof, it suffices to check continuous differentiability.) We can moreover see that the section σ has classical Fredholm linearizations at any zero $(t_0, [u]) \in \sigma^{-1}(0)$. Indeed, in the coordinates centered at a point with $\overline{\partial}_{J_{t_0}} u = 0$ we have $\mathrm{d}f(t_0,0)(T,\zeta) : \mathbb{R} \times E_u \to F$ given by

$$\mathrm{d}f(t_0,0)(T,\zeta) \;=\; -\tfrac{1}{2} T \cdot J_{t_0}(\partial_t J_t)|_{t=t_0} \partial_{J_{t_0}} u \;+\; (\mathrm{D}_u \overline{\partial}_{J_{t_0}})\zeta, \tag{18}$$

[21] Recall we expect $\mathcal{B}' = \mathcal{B}$ by Remark 3.6.

with E_u given in (10). The first part, $T \mapsto -\frac{1}{2}T \cdot J_{t_0}(\partial_t J_t)|_{t=t_0}\partial_{J_{t_0}} u$, is a bounded linear operator $\mathbb{R} \to F$ with respect to any $W^{\ell,2}$-norm on F, hence it is compact with respect to the $W^{2,2}$-norm on F. The second part is a restriction of the classical Cauchy-Riemann operator $\mathrm{D}_u\overline{\partial}_{J_{t_0}} : X := W^{3,2}(S^2, u^*\mathrm{T}Q) \to F$. This classical operator is known to be Fredholm, see e.g. [15, Thm.C.1.10], and restriction to the finite codimension subspace $E_u \subset X$ preserves the Fredholm property. This shows that $\mathrm{d}f(t_0, 0)$ is classically Fredholm with index given by the index of the restriction $\mathrm{D}_u\overline{\partial}_{J_{t_0}}|_{E_u}$ plus the dimension of the domain $\mathbb{R}$ of the compact factor. The index of $\mathrm{D}_u\overline{\partial}_{J_{t_0}}$ is $2n+2c_1([\mathbb{CP}^1 \times \{\mathrm{pt}\}])$ by the Riemann-Roch Theorem. The Chern number can be computed with respect to any compatible almost complex structure on Q, and for $J_0 = J_{\mathbb{CP}^1} \oplus J_T$ we have

$$\begin{aligned} c_1([\mathbb{CP}^1 \times \{\mathrm{pt}\}]) &= \int_{\mathbb{CP}^1 \times \{\mathrm{pt}\}} c_1(\mathrm{T}Q, J_0) \\ &= \int_{\mathbb{CP}^1} c_1(\mathrm{T}\mathbb{CP}^1, J_{\mathbb{CP}^1}) + \int_{\{\mathrm{pt}\}} c_1(\mathrm{T}T, J_T) \;=\; 2 + 0. \end{aligned}$$

Finally, restricting an operator to a subspace reduces the Fredholm index (via a mix of its effects on kernel and image) by the codimension of the subspace. In this case, $E_u \subset W^{3,2}(S^2, u^*\mathrm{T}Q)$ has codimension $2n+4$ since it is given by the codimension $2n$ condition $\eta(z_0) = 0$ and the two codimension 2 conditions $\eta(z_i) \in H_{p_i}$. Thus we obtain the claimed Fredholm index

$$\mathrm{index}(\mathrm{d}f(t_0, 0)) \;=\; 1 + 2n + 2c_1([\mathbb{CP}^1 \times \{\mathrm{pt}\}]) - 2n - 4 \;=\; 1.$$

This also establishes the classical Fredholm property of the section σ in local coordinates. The nonlinear sc-Fredholm property of polyfold theory, however, demands more than just the linearizations being sc-Fredholm.[22] The sc-Fredholm property of a section as defined in [9, Def. 3.8] requires three conditions. The first, sc-smoothness, follows from the classical smoothness in local coordinates. The second condition, σ being regularizing, means that if $(t, v) \in [0, 1] \times W^{3+m,2}$ with $\overline{\partial}_{J_t} v \in W^{2+m+1,2}$, then $(t, v) \in [0, 1] \times W^{3+m+1,2}$. This follows from the corresponding property of ∂_{J_t} for fixed t—which is known from elliptic regularity. The third condition seems less transparent but is very important for the implicit function theorem in scale calculus: At every smooth point $(t_0, [u]) \in [0, 1] \times \mathcal{B}_\infty$, the section needs to have the sc-Fredholm germ property [9, Def. 3.7], that is, after subtraction of a local sc^+-section, (a filled version of)[23] the germ of σ needs to be conjugate to a basic germ as defined in [9, Def. 3.6]. The latter means that

[22] The linear sc-Fredholm property [9, Def. 1.8] is a direct analogue of the classical linear Fredholm property—kernel and image need to have complements that respect the sc-structure, these complements need to be sc-isomorphic, and kernel and cokernel should be finite dimensional.

[23] In our setting, there is no need for a filling since all retractions are trivial.

after splitting off a finite dimensional factor from the domain and projecting to the complement of a finite dimensional factor in the image, the germ is the identity plus a contraction mapping. It is this third property that implies that linearizations of a sc-Fredholm section are sc-Fredholm operators, see [9, Prop. 3.10]. The index is then defined as the Fredholm index of the linearization at the lowest level of the scale structure, which we computed above to be 1.

To establish the equivalence to a contraction germ normal form, we proceed similar to [10, Prop. 4.26], using the fact that the section is classically differentiable in all but finitely many directions. In fact, the local representative f above is continuously differentiable in all directions, and thus satisfies the conditions of being sc-Fredholm with respect to the trivial splitting $E_u \cong \{0\} \times E_u$, as defined in [18, Def. 4.1]. Indeed, we already established the regularizing property (i). The differentiablity conditions (ii) in the trivial splitting follow from classical continuous differentiability. Besides, in this setting the linearized sc-Fredholm property (iii) is only required of $\mathrm{D}_0 f$—though for any (not necessarily holomorphic) base point $(t_0, [u])$. We can make up for the latter complication by subtracting from f the sc^+-section

$$f_0 : I_{t_0} \times \mathcal{O} \longrightarrow F, \quad (t, \eta) \longmapsto K_t^{-1}\big(\overline{\partial}_{J_t} u\big),$$

which takes the same value at $(t_0, 0)$ as f. Thus the linearization of $f - f_0$ at $(t_0, 0)$ is well defined, and it can be computed as follows:

$$\mathrm{d}(f - f_0)(t_0, 0)(T, \zeta) = T \cdot \partial_t(\Gamma_t(\eta)^{-1})(\overline{\partial}_{J_{t_0}} u) + \mathrm{D}_0 \Gamma_t^{-1}(\zeta)(\overline{\partial}_{J_{t_0}} u) + \mathrm{D}_u \overline{\partial}_{J_{t_0}} \zeta.$$

To check that this is a linear sc-Fredholm operator we use the conditions of [18, Def. 3.1]. (1) It is bounded on each level of the scale structure. (2) It is regularizing by elliptic regularity for the linearized Cauchy-Riemann operator. Lastly, we have to check in (3) the classical Fredholm property on the lowest level of the scale structure. Note that the first two summands actually are bounded with respect to the $W^{3,2}$-norm on F, and thus induce compact operators to F with the $W^{2,2}$-topology. Thus the Fredholm property again follows from the corresponding property of the linearized Cauchy-Riemann operator. All in all, we have shown that the section $f - f_0$ in local coordinates satisfies all conditions of [18, Thm. 4.5], which implies its contraction germ normal form.[24] This finishes the proof. □

Remark 3.15 An alternative proof of Theorem 3.14 is to combine an implicit function theorem from [4] with a polyfold description of the Gromov-Witten moduli space $\mathcal{M}$ for a family $(J_t)_{t\in[0,1]}$ of almost complex structures. Such a description is obtained as follows. First, one follows Remark 3.13 to construct an M-polyfold bundle $\widetilde{p} : \widetilde{W} \to [0, 1] \times Z$ which restricts on every slice $\{t\} \times Z$ to the

[24] Strictly speaking, the statement of this theorem shifts the scale structure. We can avoid this by observing that all the differentiability and linearized Fredholm properties of $f : \mathbb{R} \times E_u \to F$ persist w.r.t. the $W^{2,2}$-topology on E_u and the $W^{1,2}$-topology on F.

bundle $W \to Z$ from [10] for the almost complex structure J_t. Then, the section $\overline{\partial} : [0,1] \times Z \to \widetilde{W}, (t,[u]) \mapsto \big(t, \big[(u, \overline{\partial}_{J_t} u)\big]\big)$ is shown to be sc-Fredholm by following the arguments of [10, Thm.4.6][25] or our proof of Theorem 3.14. Given such a description, we claim that the sc-Fredholm property is preserved when we restrict from $[0,1] \times Z$ to the preimage $[0,1] \times \mathcal{B} = \widetilde{\mathrm{ev}}^{-1}(\{p_0\})$ of the submanifold $\{p_0\} \subset Q$ under the evaluation map $\widetilde{\mathrm{ev}} : [0,1] \times Z \to Q, (t,[u]) \mapsto \mathrm{ev}([u])$. For that purpose we can again quote the results by Filippenko: [4, §5.1] explains why the Cauchy-Riemann section, the evaluation map and the submanifold $\{p_0\} \subset Q$ satisfy all compatibility conditions of [4, Theorem 5.10 (III)]. A direct application of that result asserts that $\sigma|_{\widetilde{\mathcal{B}}} : \widetilde{\mathcal{B}} \to W|_{\widetilde{\mathcal{B}}}$ is sc-Fredholm for some open subset $\widetilde{\mathcal{B}} \subset [0,1] \times \mathcal{B}_1$ containing $[0,1] \times \mathcal{B}_\infty$. However, the shift in topology is again not needed since the evaluation map is classically smooth on all levels. Furthermore, as in Remark 3.13, the open subset $\widetilde{\mathcal{B}} \subset [0,1] \times \mathcal{B}$ can be replaced by $[0,1] \times \mathcal{B}'''$ for an open neighbourhood $\mathcal{B}''' \subset \mathcal{B}$ of $\mathcal{B}_\infty$ that may just be smaller than $\mathcal{B}'$ from Theorem 3.10.

3.6 *Linearization*

The goal of Sect. 3.7 will be to prove transversality of the unperturbed section σ at $t = 0$. For that purpose we will need to consider its linearization at the unique solution $[u_0] \in \mathcal{B}$ for $t = 0$ (see Sect. 2.2). In fact, it will be sufficient to consider the linearization of $\sigma_0 := \sigma(0,\cdot) : \mathcal{B} \to \mathcal{E}|_{\{0\}\times\mathcal{B}}$ at $[u_0] \in \mathcal{B}$, which is what we will compute now. We do the computation in a local chart for the restricted bundle $\mathcal{E}|_{\{0\}\times\mathcal{B}}$ centered at $[u_0]$, given by

$$\mathcal{O} \times F \to \mathcal{E}|_{\{0\}\times\mathcal{B}}, \qquad (\eta,\xi) \mapsto \big(0, [\exp_{u_0}\eta, \Gamma_0(\eta)\circ\xi]\big).$$

This chart is obtained from the chart (17) for the bundle $\mathcal{E} \to [0,1] \times \mathcal{B}$ centered at $(0,[u_0])$ as in Sect. 3.4, by restriction to $\mathcal{O} \times F \cong \{0\} \times \mathcal{O} \lhd F \subset [0,1] \times \mathcal{O} \lhd F$. Here $\mathcal{O} \subset E_{u_0}$ is an open subset of the vector space E_{u_0} given in (12), F is defined in (13), the exponential map is induced by a fixed metric chosen as part of the good data in the second proof of Theorem 3.5, and Γ_t is defined in (16). The map K_t from (15) does not show up here because we have $t = t_0 = 0$ and so K_t is the identity map. In this chart, the restricted section σ_0 from (7) is given by

$$\overline{\sigma}_0 : \mathcal{O} \to \mathcal{O} \times F, \qquad \eta \mapsto \big(\eta, \Gamma_0(\eta)^{-1} \circ \overline{\partial}_{J_0}\big(\exp_{u_0}\eta\big)\big).$$

Since u_0 is the center of the chart, it corresponds to $\eta = 0$. So the linearized operator $\mathrm{D}_{u_0}\sigma_0$ in the coordinates of this chart is represented by $\mathrm{d}\overline{\sigma}_0(0) : E_u \to F$, which

[25] To generalize the Fredholm analysis in [10] to allow for a finite dimensional family of almost complex structures, note that this introduces an extra factor in the domain—in our case $[0,1]$—along which the section is classically smooth. Since it is finite dimensional, it can also be split off when constructing the contraction germ normal form.

we can compute for $\hat{\eta} \in E_u$ as follows:

$$\begin{aligned} \mathrm{d}\overline{\sigma}_0(0)(\hat{\eta}) &= \frac{d}{d\theta}\Big|_{\theta=0} \xi(\theta\hat{\eta}) = \frac{d}{d\theta}\Big|_{\theta=0} \underbrace{\Gamma_0\big(\theta\hat{\eta}\big)^{-1}}_{=:A_\theta} \circ \underbrace{\overline{\partial}_{J_0}\big(\exp_{u_0}\theta\hat{\eta}\big)}_{=:\mu_\theta} \\ &= \underbrace{A_0}_{=\mathrm{id}} \circ \frac{d}{d\theta}\Big|_{\theta=0} \mu_\theta + \frac{d}{d\theta}\Big|_{\theta=0} A_\theta \circ \underbrace{\mu_0}_{=0} = \tfrac{1}{2}\left(\nabla\hat{\eta} + J_0(u_0)\circ\nabla\hat{\eta}\circ i\right). \end{aligned}$$

Here ∇ denotes the Levi-Civita connection on Q corresponding to the metric g_0 that is compatible with J_0. It is $A_0 = \Gamma_0(0)^{-1} = \mathrm{id}_{u_0^*\mathrm{T}Q}$, as for every $z \in S^2$ it is (the inverse of) parallel transport from $u_0(z)$ to itself via the constant path, and $\mu_0 = \overline{\partial}_{J_0}\left(\exp_{u_0} 0\right) = \overline{\partial}_{J_0} u_0 = 0$ since u_0 is J_0-holomorphic. In the last step we used a formula from [15, Prop. 3.1.1] and again the fact that u_0 is J_0-holomorphic. Thus we have shown that the polyfold-theoretic linearization $\mathrm{D}_{u_0}\sigma_0$ in an appropriate bundle chart is given by a restriction of the classical Cauchy-Riemann operator of the complex bundle $(u_0^*\mathrm{T}Q, J_0(u_0))$,

$$\mathrm{d}\overline{\sigma}_0(0) : E_u \to F, \qquad \hat{\eta} \mapsto \tfrac{1}{2}\left(\nabla\hat{\eta} + J_0(u_0)\circ\nabla\hat{\eta}\circ i\right) = \mathrm{D}_{J_0(u_0)}\hat{\eta}.$$

Indeed, this operator differs from the classical Cauchy-Riemann operator $\mathrm{D}_{J_0(u_0)} : X \to F$ as in [15, Rmk.C.1.2] only by the domain E_{u_0} being a subspace of $X := W^{3,2}(S^2, u_0^*\mathrm{T}Q)$.

3.7 Transversality at the Boundary

The last missing ingredient for the proof of Gromov's nonsqueezing Theorem 1.2 in Sect. 2.5 is to show that the unique J_0-holomorphic curve is cut out transversely.

Remark 3.16 The following proof is a first instance of the general principle "classical transversality implies polyfold transversality". The core difference between the two notions is that, classically, one usually proves surjectivity of a linearized Cauchy-Riemann operator $\mathrm{D} : X \to F$ on a tangent space $X = \mathrm{T}_{u_0}\mathcal{X}$ to a space $\mathcal{X}$ of all maps (in a given homology class, etc.). In our case (ignoring the homology class and point constraint $u(z_0) = p_0$), this total space would be $\mathcal{X} = W^{3,2}(S^2, Q)$. This space still carries the action of a group of reparameterizations. In our case a group Aut of automorphisms of S^2 acts on $\mathcal{X}$, and preserves the Cauchy-Riemann operator, so that the tangent space of its orbit lies in the kernel, $\mathrm{T}_{u_0}\{u_0 \circ \varphi \mid \varphi \in \mathrm{Aut}\} \subset \ker \mathrm{D}$.

In contrast, the polyfold setup works with the quotient space $\mathcal{B} = \mathcal{X}/\mathrm{Aut}$ whose tangent space $\mathrm{T}_{[u_0]}\mathcal{B}$ is represented by a subspace $E_{u_0} \subset X = W^{3,2}(S^2, u_0^*\mathrm{T}Q)$. So, the main challenge in deducing surjectivity of the linearized polyfold section $\mathrm{D}_{[u_0]}\sigma = \mathrm{D}|_{E_{u_0}}$ from surjectivity of the classical Cauchy-Riemann operator D is

in showing that this quotient construction results in a splitting of the total space $X = E_{u_0} + A$ with a complement $A \subset \ker \mathrm{D}$ that represents the infinitesimal action of reparameterizations.

Theorem 3.17 *σ is transverse to the zero section at $t = 0$.*

Proof Since u_0 is the only solution for $t = 0$ (see Lemma 2.2), for checking transversality of σ at the boundary $t = 0$ it suffices to consider the linearization of σ at $(0, u_0)$ and show that it is surjective. For this it is sufficient to show that the linearization of $\sigma_0 := \sigma(0, \cdot)$ is surjective. In Sect. 3.6 we computed this linearization $\mathrm{D}_{u_0}\sigma_0$ in a local chart centered at $u_0 \in \mathcal{B}$ to be the restriction $\mathrm{d}\overline{\sigma}_0(0) = \mathrm{D}_{J_0(u_0)}|_{E_{u_0}} : E_{u_0} \to F$ of the classical Cauchy-Riemann operator $\mathrm{D}_{J_0(u_0)} : X \to F$ to the subspace $E_{u_0} \subset X = W^{3,2}(S^2, u_0^*\mathrm{T}Q)$ given by (12). Its codomain F is defined in (13). We will first show that this classical operator is surjective; see also [20, Lemma 5.5].

Recall from Sect. 2.2 that $u_0 : S^2 \to Q = \mathbb{CP}^1 \times T, z \mapsto (z, m_0)$ is the product of the identifying map $S^2 \cong \mathbb{CP}^1$ and a constant map to the torus. Thus the complex structure along u_0 splits $J_0(u_0) = J_{\mathbb{CP}^1} \oplus J_T(m_0)$ into the standard complex structure $J_{\mathbb{CP}^1}$ on $\mathrm{T}\mathbb{CP}^1 \cong \mathrm{id}^*\mathrm{T}\mathbb{CP}^1$ and the constant complex structure $J_T(m_0) = J_{\mathrm{st}}$ on $\mathrm{T}_{m_0}T = \mathbb{R}^{2n-2}$. Thus we have a natural splitting of complex vector bundles

$$u_0^*(\mathrm{T}Q, J_0) = (\mathrm{T}\mathbb{CP}^1, J_{\mathbb{CP}^1}) \oplus E_0^{(n-1)},$$

where $E_0^{(n-1)}$ is the trivial complex bundle of rank $n-1$ over S^2 with fibers $(\mathbb{R}^{2n-2}, J_{\mathrm{st}})$. The domain of the Cauchy-Riemann operator thus splits into

$$X = W^{3,2}\left(S^2, u_0^*\mathrm{T}\left(\mathbb{CP}^1 \times T\right)\right) = \underbrace{W^{3,2}(S^2, \mathrm{T}\mathbb{CP}^1)}_{=:X_{\mathbb{CP}^1}} \times \underbrace{W^{3,2}(S^2, E_0^{(n-1)})}_{=:X_T},$$

and, analogously, the codomain splits $F = F_{\mathbb{CP}^1} \times F_T$ into the $W^{2,2}$-closures of smooth complex antilinear 1-forms on S^2 with values in $\mathrm{T}\mathbb{CP}^1$ resp. $E_0^{(n-1)}$. This shows that the classical Cauchy-Riemann operator splits

$$\mathrm{D}_{J_0(u_0)} = \mathrm{D}_{J_{\mathbb{CP}^1}} \oplus \mathrm{D}_{J_{\mathrm{st}}} : X = X_{\mathbb{CP}^1} \times X_T \to F = F_{\mathbb{CP}^1} \times F_T$$

into the classical Cauchy-Riemann operator $\mathrm{D}_{J_{\mathbb{CP}^1}} : X_{\mathbb{CP}^1} \to F_{\mathbb{CP}^1}$ of the complex line bundle $(\mathrm{T}\mathbb{CP}^1, J_{\mathbb{CP}^1})$ over $\mathbb{CP}^1 \cong S^2$ and the Cauchy-Riemann operator $\mathrm{D}_{J_{\mathrm{st}}} : X_T \to F_T$ of the trivial bundle $E_0^{(n-1)}$ over S^2. The latter splits further $\mathrm{D}_{J_{\mathrm{st}}} = \mathrm{D}_i \oplus \ldots \oplus \mathrm{D}_i$ into $n-1$ copies of the Cauchy-Riemann operator D_i of the trivial complex line bundle $E_0^{(1)} = \mathbb{C} \times S^2 \to S^2$. These complex line bundles satisfy $c_1(\mathrm{T}\mathbb{CP}^1) + 2\chi(S^2) = 2 + 2 \cdot 2 > 0$ resp. $c_1(E_0^{(1)}) + 2\chi(S^2) = 0 + 2 \cdot 2 > 0$, so the Riemann-Roch Theorem (e.g. [15, Thm.C.1.10.(iii)]) guarantees that the

corresponding Cauchy-Riemann operators are surjective. This proves surjectivity of the product $\mathrm{D}_{J_0(u_0)}$ of these surjective operators arising from complex line bundles.

Towards proving surjectivity of the restriction $\mathrm{d}\overline{\sigma}_0(0) = \mathrm{D}_{J_0(u_0)}|_{E_{u_0}} : E_{u_0} \to F$, note the above surjectivity means $\mathrm{D}_{J_0(u_0)}(X) = F$. If we can now show that $X = E_{u_0} + V$ can be written as the sum of E_{u_0} and a subspace of the kernel $V \subset \ker \mathrm{D}_{J_0(u_0)}$, then it follows that $\mathrm{D}_{J_0(u_0)}(E_{u_0}) = \mathrm{D}_{J_0(u_0)}(E_{u_0} + V) = F$, so the restriction $\mathrm{d}\overline{\sigma}_0(0) = \mathrm{D}_{J_0(u_0)}|_{E_{u_0}}$ is surjective as well. To find this subspace V note that our choice of transverse hypersurfaces in Remark 3.8 was made such that the vector space E_{u_0} defined in (12) splits as $E_{u_0} = E_{\mathbb{CP}^1} \times E_T$, where

$$E_{\mathbb{CP}^1} = \{\eta \in X_{\mathbb{CP}^1} \mid \eta(z_i) = 0 \text{ for } i = 0, 1, 2\}, \qquad E_T = \{\eta \in X_T \mid \eta(z_0) = 0\}.$$

So we can construct $V \subset X$ as product $V := V_{\mathbb{CP}^1} \times V_T$ of subspaces satisfying

(a) $X_{\mathbb{CP}^1} = E_{\mathbb{CP}^1} + V_{\mathbb{CP}^1}$ and $\mathrm{D}_{J_{\mathbb{CP}^1}}(V_{\mathbb{CP}^1}) = 0$,
(b) $X_T = E_T + V_T$ and $\mathrm{D}_{J_{\mathrm{st}}}(V_T) = 0$.

To meet these requirements, we choose the subspaces of holomorphic sections

$$V_{\mathbb{CP}^1} := \mathrm{T}_{\mathrm{id}}\mathrm{Aut}(S^2) \quad \text{and} \quad V_T := \{\eta_T : S^2 \to \mathrm{T}_{m_0}T \ \text{ constant}\}.$$

With this choice it is easy to verify (b): Every $\eta \in X_T$ is the sum of the constant vector field taking the value $\eta(z_0)$ and the vector field $\eta(\cdot) - \eta(z_0) \in E_T$. Moreover, $\mathrm{D}_{J_{\mathrm{st}}}(V_T) = 0$ holds since all the elements of V_T are constant.

To verify (a), first recall that $\mathrm{Aut}(S^2)$ is the group of biholomorphisms $S^2 \to \mathbb{CP}^1$, i.e. solutions $\psi : S^2 \to \mathbb{CP}^1$ of the nonlinear Cauchy-Riemann equation $\overline{\partial}_{J_{\mathbb{CP}^1}}\psi = 0$. So the tangent space $\mathrm{T}_{\mathrm{id}}\mathrm{Aut}(S^2)$ to this finite dimensional family of holomorphic maps at the identity map is a subspace of the kernel of the linearized Cauchy-Riemann operator $\ker \mathrm{D}_{J_{\mathbb{CP}^1}} \subset X_{\mathbb{CP}^1}$. In particular, $V_{\mathbb{CP}^1} = \mathrm{T}_{\mathrm{id}}\mathrm{Aut}(S^2)$ is a subspace of the space of smooth sections $\mathcal{C}^\infty(S^2, \mathrm{T}\mathbb{CP}^1) \subset X_{\mathbb{CP}^1} = W^{3,2}(S^2, \mathrm{T}\mathbb{CP}^1)$ since these tangent vectors are obtained as derivatives of paths of maps in $\mathrm{Aut}(S^2)$,

$$\mathrm{T}_{\mathrm{id}}\mathrm{Aut}(S^2) = \left\{ \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}\gamma(t) \;\middle|\; \gamma\colon (-\varepsilon, \varepsilon) \to \mathrm{Aut}(S^2) \text{ with } \gamma(0) = \mathrm{id} \right\}.$$

These derivatives take values $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}\gamma(t)(z) \in \mathrm{T}_{\gamma(0)(z)}\mathbb{CP}^1 = \mathrm{T}_z\mathbb{CP}^1$ at $z \in S^2$. So, strictly speaking, they are sections of the pullback bundle $\mathrm{id}^*\mathrm{T}\mathbb{CP}^1$ under the identification $\mathrm{id} : S^2 \xrightarrow{\cong} \mathbb{CP}^1$. More concretely, the automorphisms of $S^2 = \mathbb{C} \cup \{\infty\}$ are of the form $\infty \neq z \mapsto \frac{az+b}{cz+d}$ and $\infty \mapsto \frac{a}{c}$ with $a, b, c, d \in \mathbb{C}$ such that $ad - bc \neq 0$. Thus, each $\gamma\colon (-\varepsilon, \varepsilon) \to \mathrm{Aut}(S^2)$ is given by differentiable functions $a, b, c, d : (-\varepsilon, \varepsilon) \to \mathbb{C}$ that satisfy $a(0) = 1 = d(0)$, $b(0) = 0 = c(0)$ and $a(t)d(t) - b(t)c(t) \neq 0$ for all t. We can compute their derivative at any $z \in S^2 \backslash \{\infty\}$,

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\gamma(t)(z) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\frac{a(t)z+b(t)}{c(t)z+d(t)} = \frac{\dot a(0)z+\dot b(0)}{0z+1} - \frac{(1z+0)(\dot c(0)z+\dot d(0))}{(0z+1)^2}$$
$$= z\dot a(0) + \dot b(0) - z^2\dot c(0) - z\dot d(0).$$

Conversely, any choice of numbers $A, B, C, D \in \mathbb{C}$ induces a tuple of functions $(-\varepsilon,\varepsilon) \to \mathbb{C}$ as above, given by $a(t) = 1 + tA$, $b(t) = tB$, $c(t) = tC$, $d(t) = 1 + tD$. These define a section[26] $\eta_V = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}\gamma(t)\colon S^2 \to \mathrm{T}S^2$ in $X_{\mathbb{CP}^1}$, which on $S^2 \setminus \{\infty\} \simeq \mathbb{C}$ is of the form[27]

$$\eta_V(z) \;=\; z(A-D) + B - z^2C \;\in\; \mathbb{C} \cong \mathrm{T}_z\mathbb{CP}^1. \tag{19}$$

Now, to prove (a), given any $\eta \in X_{\mathbb{CP}^1}$, we need to find $\eta_E \in E_{\mathbb{CP}^1}$ and $\eta_V \in V_{\mathbb{CP}^1}$ with $\eta = \eta_E + \eta_V$. Since η_E needs to satisfy $\eta_E(z_i) = 0$ for $i = 0, 1, 2$, we need η_V to agree with η on all marked points. Without loss of generality, we can assume that $z_0 = 0$, and by Remark 3.8 we are free to choose the points $z_1, z_2 \in S^2$ as we like. We choose $z_1 := 1$ and $z_2 := -1$. Then, given $\eta \in X_{\mathbb{CP}^1}$ the requirements $\eta_V(z_i) = \eta(z_i)$ translate by (19) into

$$\begin{aligned} B \;&=\; \eta_V(0) \;=\; \eta(0),\\ A + B - C - D \;&=\; \eta_V(1) \;=\; \eta(1),\\ -A + B - C + D \;&=\; \eta_V(-1) \;=\; \eta(-1). \end{aligned}$$

This system of 3 equations for 4 variables can be solved by choosing $D = 0$,

$$A \;=\; \tfrac{\eta(1)-\eta(-1)}{2}, \qquad B \;=\; \eta(0), \qquad C \;=\; \eta(0) - \tfrac{\eta(1)+\eta(-1)}{2}, \qquad D = 0.$$

As explained above, this choice defines a vector field $\eta_V \in \mathrm{T}_{\mathrm{id}}\mathrm{Aut}(S^2) = V_{\mathbb{CP}^1}$ on S^2, and using (19) we ensured that it has the desired values $\eta_V(z_i) = \eta(z_i)$. Now $\eta_E := \eta - \eta_V$ satisfies $\eta_E(z_i) = 0$ for $i = 0, 1, 2$, and hence we have $\eta = \eta_E + \eta_V$ with $\eta_E \in E_{\mathbb{CP}^1}$. This finishes the proof of (a) and thus proves this theorem. □

Remark 3.18 If $\mathcal{M}_1 = \emptyset$, then σ is transverse to the zero section at $t = 1$. The statement that the linearization is surjective for all $u \in \mathcal{M}_1$ is then vacuously true.

[26] Since $a(t)d(t) - b(t)c(t) \neq 1$ is an open condition and satisfied at $t = 0$, by continuity it is satisfied for all small enough t.

[27] This computation uses the chart $\mathbb{CP}^1 \setminus \{\infty\} \cong \mathbb{C}$. To compute the value of η_V at $\infty \in \mathbb{CP}^1$, we would have to consider a chart near ∞. This is not needed here if we just avoid putting a marked point at ∞.

Appendix 1: The Monotonicity Lemma for Pseudoholomorphic Maps

The purpose of this appendix is to give a detailed proof of the monotonicity lemma for J_{st}-holomorphic maps to $\mathbb{R}^{2n}$ that was used in Sect. 2.3, which avoids the use of special properties of holomorphic maps such as their local representation. Other proofs can be found in the literature; e.g. [20, Thm.5.2.1]. We also use the opportunity to establish the result in maximal generality—for maps of regularity C^1, and with $(\mathbb{R}^{2n}, J_{\mathrm{st}})$ generalized to a complex Hilbert space as follows.

Definition A.1 A **complex Hilbert space** (V, J) consists of a Hilbert space V with inner product $\langle \cdot, \cdot \rangle$ and a compatible complex structure J, i.e. an endomorphism $J : V \to V$ with $J^2 = -\mathrm{id}_V$ that preserves the inner product. The associated symplectic structure $\omega : V \times V \to \mathbb{R}$ is $\omega(v_1, v_2) = \langle J v_1, v_2 \rangle$.

A **pseudoholomorphic map** $v : (S, j) \to (V, J)$ consists of a compact Riemann surface (S, j) with (possibly empty) boundary ∂S and a C^1-map[28] $v : S \to V$ satisfying the Cauchy-Riemann equation $\mathrm{d}v \circ j = J \circ \mathrm{d}v$.

Lemma A.2 *Consider a nonconstant*[29] *(j, J)-holomorphic map $v : S \to V$ and an open ball $\mathring{B}_R(p) := \{q \in V \mid \|q - p\| < R\}$ centered at a point $p \in v(S)$ in the image, of radius $R > 0$ such that $\|v(z) - p\| \geq R$ for all $z \in \partial S$. Then the symplectic area of v within the ball is at least the area of the flat disk of radius R, that is*

$$\int_{v^{-1}(\mathring{B}_R(p))} v^*\omega \geq \pi R^2.$$

Proof We begin by rewriting the 2-form $v^*\omega$ on S in local holomorphic coordinates $s + it \in \mathbb{C}$ for $(S, j) \simeq (\mathbb{C}, i)$ as

$$\begin{aligned} v^*\omega &= \omega(\partial_s v, \partial_t v)\, \mathrm{d}s \wedge \mathrm{d}t \\ &= \tfrac{1}{2}\big(\|\partial_s v\|^2 + \|\partial_t v\|^2\big)\, \mathrm{d}s \wedge \mathrm{d}t \\ &= \tfrac{1}{2}\big\langle \big(\partial_s v\, \mathrm{d}s + \partial_t v\, \mathrm{d}t\big) \wedge \big(\partial_s v\, \mathrm{d}t - \partial_t v\, \mathrm{d}s\big)\big\rangle \;=\; \tfrac{1}{2}\,\langle \mathrm{d}v \wedge *\mathrm{d}v \rangle\,. \end{aligned} \tag{20}$$

Here the Cauchy-Riemann equation $\partial_s v + J\partial_t v = 0$ in local coordinates together with compatibility of ω and J with the inner product implies

$$\omega(\partial_s v, \partial_t v) = \langle J\partial_s v, \partial_t v \rangle = \langle -J^2 \partial_t v, \partial_t v \rangle = \|\partial_t v\|^2 = \|\partial_s v\|^2.$$

[28] If V is finite dimensional, then v is automatically smooth by elliptic regularity.

[29] More precisely, we assume that v is not constant on any connected component of the domain. This excludes the pathological case of $v \equiv p$ on one connected component and on the other components covering just a small symplectic area in $V \setminus B_R(p)$.

The final result $\frac{1}{2}\langle \mathrm{d}v \wedge *\mathrm{d}v\rangle$ is a well defined global expression (i.e. independent of coordinates) that uses the Hodge operator $* : TS \to TS$ induced by the complex structure j. In local holomorphic coordinates the Hodge operator is given by $*\mathrm{d}s = \mathrm{d}t$ and $*\mathrm{d}t = -\mathrm{d}s$. Here and below the notation $\langle \alpha \wedge \beta\rangle$ for differential forms with values in V denotes the wedge product $\wedge$ on the level of differential forms, with two values in V being multiplied via the inner product $\langle \cdot, \cdot\rangle$.

The second expression for $v^*\omega$ in (20) shows that the area $A(r) := \int_{v^{-1}(\mathring{B}_r(p))} v^*\omega$ is non-negative and grows monotone with r, since it integrates a non-negative multiple of the area form on S over domains $v^{-1}(\mathring{B}_r(p))$ that grow with r. Classical monotonicity proofs now argue that the ratio $a(r) := r^{-2}A(r)$ as function[30] of $r \in (0, R]$ satisfies $\frac{\mathrm{d}}{\mathrm{d}r}a \geq 0$ and $\lim_{r\to 0} a(r) \geq \pi$, which implies the claim $a(R) \geq \pi$. We will follow the same line of argument but avoid differentiability concerns by establishing a uniform difference estimate

$$A(r+\varepsilon) - A(r) \;\geq\; \tfrac{2\varepsilon}{r+\varepsilon} A(r) \qquad \text{for all } 0 < r < R,\ 0 < \varepsilon < R - r. \tag{21}$$

Proof of Difference Estimate To prove (21), we simplify notation by assuming without loss of generality that $p = 0$. This can be achieved by applying a global shift which does not affect the area. We can estimate the area of v in $\mathring{B}_r(0)$ by

$$A(r) = \textstyle\int_{v^{-1}(\mathring{B}_r(0))} v^*\omega \;\leq\; \frac{1}{2}\int_S f_\varepsilon(\|v\|)\langle \mathrm{d}v \wedge *\mathrm{d}v\rangle, \tag{22}$$

where $f_\varepsilon : [0,\infty) \to [0,1]$ is the continuous, piecewise linear cutoff function with $f_\varepsilon|_{[0,r]} \equiv 1$, $f_\varepsilon|_{[r+\varepsilon,\infty]} \equiv 0$, and $f_\varepsilon'|_{(r,r-\varepsilon)} = -\varepsilon^{-1}$. On the other hand, integration by parts yields

$$\begin{aligned}
\textstyle\int_S f_\varepsilon(\|v\|)\langle \mathrm{d}v \wedge *\mathrm{d}v\rangle &= \textstyle\int_{\partial S} f_\varepsilon(\|v\|)\langle v, *\mathrm{d}v\rangle - \int_S f_\varepsilon'(\|v\|)\,\mathrm{d}\|v\| \wedge \langle v, *\mathrm{d}v\rangle \\
&\leq -\textstyle\int_S f_\varepsilon'(\|v\|)\frac{1}{2}\|v\|\langle \mathrm{d}v \wedge *\mathrm{d}v\rangle \\
&\leq \tfrac{r+\varepsilon}{\varepsilon}\textstyle\int_{v^{-1}(\mathring{B}_{r+\varepsilon}(0))\setminus v^{-1}(\mathring{B}_r(0))} \frac{1}{2}\langle \mathrm{d}v \wedge *\mathrm{d}v\rangle \\
&= \tfrac{r+\varepsilon}{\varepsilon}\big(A(r+\varepsilon) - A(r)\big).
\end{aligned} \tag{23}$$

Here the first step uses a weak version of the Laplace equation $\mathrm{d} * \mathrm{d}v = 0$ which follows from the Cauchy-Riemann equation $\overline{\partial}_J v = 0$. If v is twice differentiable then this can be checked in local coordinates,

$$*\mathrm{d} * \mathrm{d}v = -\partial_s^2 v - \partial_t^2 v = (-\partial_s + J\partial_t)(\partial_s v + J\partial_t v) = (\overline{\partial}_J)^*\overline{\partial}_J v = 0.$$

[30] Monotonicity of $A(r)$ implies that this function is differentiable almost everywhere and has at most countably many jump discontinuities. Then the same holds for the ratio function $a(r)$ since $r \mapsto r^{-2}$ is smooth on the domain $(0, R]$.

Otherwise we first consider smooth functions $w : S \to V$ with $w|_{\partial S} = 0$ and calculate using integration by parts

$$\int_S \langle \mathrm{d}w \wedge * \mathrm{d}v \rangle = \int_S \langle \mathrm{d} * \mathrm{d}w, v \rangle = \int_S \langle (\overline{\partial}_J)^* \overline{\partial}_J w, v \rangle \, \mathrm{dvol} = \int_S \langle \overline{\partial}_J w, \overline{\partial}_J v \rangle \, \mathrm{dvol} = 0.$$

Then we note that this identity extends by continuity to $\mathcal{C}^1$ functions such as $w = f_\varepsilon(\|v\|)v$, which vanishes on ∂S since $\|v(\partial S)\|$ takes values in $[R, \infty)$, where f_ε vanishes since $r + \varepsilon < R$. This yields the first step in (23), with the integral over ∂S vanishing.

The second step in (23) can be checked in local holomorphic coordinates and using (20) again,

$$\begin{aligned} \mathrm{d}\|v\| \wedge \langle v, * \mathrm{d}v \rangle &= \tfrac{1}{2\|v\|} \big(\langle v, \partial_s v \rangle \, \mathrm{d}s + \langle v, \partial_t v \rangle \, \mathrm{d}t \big) \wedge \big(\langle v, \partial_s v \rangle \, \mathrm{d}t - \langle v, \partial_t v \rangle \, \mathrm{d}s \big) \\ &= \tfrac{1}{2\|v\|} \big(\langle v, \partial_s v \rangle^2 + \langle v, \partial_t v \rangle^2 \big) \mathrm{d}s \wedge \mathrm{d}t \\ &\geq \tfrac{1}{2\|v\|} \big(\|v\|^2 \|\partial_s v\|^2 + \|v\|^2 \|\partial_t v\|^2 \big) \mathrm{d}s \wedge \mathrm{d}t = \tfrac{\|v\|}{2} \langle \mathrm{d}v \wedge * \mathrm{d}v \rangle. \end{aligned}$$

The third step in (23) follows from $f_\varepsilon'(\|v\|) \equiv 0$ unless $r + \varepsilon \geq \|v\| \geq r$, and $f_\varepsilon' = \varepsilon^{-1}$ where it doesn't vanish. Now combining (22) and (23) proves (21),

$$\tfrac{r+\varepsilon}{\varepsilon} \big(A(r + \varepsilon) - A(r) \big) \;\geq\; \textstyle\int_S f_\varepsilon(\|v\|) \langle \mathrm{d}v \wedge * \mathrm{d}v \rangle \;\geq\; 2A(r).$$

Monotone Growth of Ratio Function In terms of the ratio function $a(r) = r^{-2} A(r)$, the difference estimate (21) implies

$$\begin{aligned} (r + \varepsilon)^2 \big(a(r + \varepsilon) - a(r) \big) &= A(r + \varepsilon) - \tfrac{(r+\varepsilon)^2}{r^2} A(r) \\ &\geq A(r) + \tfrac{2\varepsilon}{r+\varepsilon} A(r) - \tfrac{(r+\varepsilon)^2}{r^2} A(r) \\ &= \tfrac{r^3 + \varepsilon r^2 + 2\varepsilon r^2 - r^3 - 3\varepsilon r^2 - 3\varepsilon^2 r - \varepsilon^3}{(r+\varepsilon) r^2} A(r) \;=\; -\varepsilon^2 \tfrac{3r+\varepsilon}{r+\varepsilon} a(r). \end{aligned}$$

Now for fixed $0 < r_0 < r_1 < R$ we have $a(r+\varepsilon) - a(r) \geq -C\varepsilon^2$ for all $r \in [r_0, r_1]$ and $0 < \varepsilon < R - r_1$, with a non-negative constant

$$C \;:=\; \max_{r \in [r_0, r_1], \varepsilon \in (0, R - r_1)} \frac{3r + \varepsilon}{(r + \varepsilon)^3} a(r) \;\leq\; \frac{3r_1 + (R - r_1)}{{r_0}^5} \int_S u^* \omega.$$

Then summation with $\varepsilon = \frac{r_1 - r_0}{N}$ (with N sufficiently large for $\varepsilon < R - r_1$) yields

$$a(r_1) - a(r_0) = \textstyle\sum_{n=0}^{N-1} a(r_0 + n\varepsilon + \varepsilon) - a(r_0 + n\varepsilon) \geq -NC \big(\frac{r_1 - r_0}{N} \big)^2.$$

Taking $N \to \infty$ this implies $a(r_1) \geq a(r_0)$ for any $0 < r_0 < r_1 < R$. Next, taking the limit $r_1 \to R$ for fixed $r_0 > 0$ yields

$$A(R) \;\geq\; \lim_{r_1 \to R} A(r_1) \;\geq\; \lim_{r_1 \to R} r_1^2 a(r_1) \;\geq\; R^2 a(r_0). \tag{24}$$

So to prove the claim $A(R) \geq \pi R^2$ it remains to establish $\lim_{r_0 \to 0} a(r_0) \geq \pi$.

Centering the Ball at a Regular Point Before studying the $r \to 0$ limit, we claim that it suffices to prove the area bound $A(R) \geq \pi R^2$ after replacing $p = v(z)$ with a sequence $p_n = v(z_n) \to p$ of images of regular points $S \setminus \partial S \ni z_n \to z$, regular just meaning that they are not critical points of v. Indeed, given such a sequence and assuming the area bound holds on all balls $\mathring{B}_R(p_n)$, we have

$$A(R) \;=\; \int_{v^{-1}(\mathring{B}_R(p))} v^*\omega \;\geq\; \int_{v^{-1}(\mathring{B}_{R-\|p_n-p\|}(p_n))} v^*\omega \;\geq\; \pi(R - \|p_n - p\|)^2$$

since $\mathring{B}_{R-\|p_n-p\|}(p_n) \subset \mathring{B}_R(p)$. This proves $A(R) \geq \pi R^2$ in the limit $\|p_n - p\| \to 0$. Moreover, such a sequence of regular points exists since the critical points of v are isolated in S. For $\dim V < \infty$ this is proven in [15, Lemma 2.4.1]. If $\dim V = \infty$ first note that by the Cauchy-Riemann equation in local coordinates, $J\partial_s v = \partial_t v$, any $z \in S$ is either regular (i.e. $\mathrm{d}_z v$ is injective) or critical (i.e. $\mathrm{d}_z v = 0$). Next, choose a complex splitting $(V, J) \simeq (\mathbb{C}, i) \oplus (V', J')$ in which the first component $\mathrm{pr}_{\mathbb{C}} \circ v$ is nonconstant (e.g. by splitting off $\operatorname{im} \mathrm{d}_z v \cong \mathbb{C}$ at a regular point). Then classical complex analysis asserts that the critical points of the nonconstant holomorphic map $\mathrm{pr}_{\mathbb{C}} \circ v : S \to \mathbb{C}$ are isolated, so in particular the critical points of v are isolated.

Lower Bound for Ratio as $\mathbf{r \to 0}$ Once $p = v(z_0)$ is the image of a regular point $z_0 \in S \setminus \partial S$ of v, we choose a neighbourhood $U_0 \subset S$ of z_0 with local holomorphic coordinates $U_0 \cong D_\delta := \{(s,t) \in \mathbb{R}^2 \,|\, s^2 + t^2 < \delta^2\}$ for some $\delta > 0$, so that $z_0 \cong (0,0)$ and $v|_{U_0}$ is given by $v_0(s,t) = p + sX_0 + tJX_0 + h(s,t)$ with a vector $X_0 = \partial_s v_0(0,0) \in V$ of length $\|X_0\| = 1$ and an error term h with $h(0,0) = 0$ and $\mathrm{d}h(0,0) = 0$. Then we can bound the area $A(r) \geq \int_{D_\delta} \chi\big(\frac{\|v_0 - p\|}{r}\big) v_0^*\omega$ by an integral of the characteristic function χ with $\chi|_{[0,1]} \equiv 1$, $\chi|_{(1,\infty)} \equiv 0$. Moreover, the limit $\lim_{r \to 0} a(r)$ exists, since we already proved monotonicity of a and it is bounded below by 0. So to prove $\lim_{r \to 0} a(r) \geq \pi$ it suffices to find a sequence $r_n \to 0$ with $\lim_{n \to \infty} r_n^{-2} \int_{D_{r_n}} \chi\big(\frac{\|v_0 - p\|}{r_n}\big) v_0^*\omega = \pi$. To construct this sequence of radii we use the continuous differentiability of h to choose $0 < r_n \leq \delta$ for each $n \in \mathbb{N}$ so that $\|\mathrm{d}h(s,t)\| \leq \frac{1}{n}$ and $\|h(s,t)\| \leq \frac{1}{n}|(s,t)|$ for $|(s,t)| \leq r_n$. Then on D_{r_n} we can estimate

$$\begin{aligned} \big|\omega(\partial_s v, \partial_t v) - 1\big| \;&=\; \big|\omega\big(X_0 + \partial_s h\,,\; JX_0 + \partial_t h\big) - \omega(X_0, JX_0)\big| \\ &=\; \big|\omega(X_0, \partial_t h) + \omega(\partial_s h, X_0) + \omega(\partial_s h, \partial_t h)\big| \\ &\leq\; \tfrac{1}{n} + \tfrac{1}{n} + (\tfrac{1}{n})^2 \;\leq\; \tfrac{3}{n}. \end{aligned}$$

To control the distance $\|v(s,t) - p\|$ we use $2xy \leq \frac{1}{n}x^2 + ny^2$ to obtain

$$\begin{aligned}\big|\|v(s,t)-p\|^2-|(s,t)|^2\big| \;&=\; \big\|sX_0+tJX_0+h(s,t)\big\|^2-(s^2+t^2)\\ &\le 2|s||\langle X_0,h(s,t)\rangle|+2|t||\langle JX_0,h(s,t)\rangle|+\|h(s,t)\|^2\\ &\le \tfrac{1}{n}(s^2+t^2)+n\big(|\langle X_0,h(s,t)\rangle|^2+|\langle JX_0,h(s,t)\rangle|^2\big)+\|h(s,t)\|^2\le \tfrac{4}{n}|(s,t)|^2.\end{aligned}$$

This holds for $(s,t)\in D_{r_n}$ and implies $\chi\big(\frac{\|v(s,t)-p\|}{r_n}\big)=1$ for $|(s,t)|\le \frac{r_n}{\sqrt{1+4n^{-1}}}=:\rho_n$. Now writing $\pi=r_n^{-2}\int_{D_{r_n}}\mathrm{d}s\,\mathrm{d}t$, from the above we obtain

$$\begin{aligned}&\Big|r_n^{-2}\textstyle\int_{D_{r_n}}\chi\big(\frac{\|v-p\|}{r_n}\big)v^*\omega\;-\;\pi\Big|\\ &=r_n^{-2}\Big|\textstyle\int_{D_{r_n}}\big(\chi\big(\frac{\|v-p\|}{r_n}\big)\omega(\partial_s v,\partial_t v)\;-\;1\big)\,\mathrm{d}s\,\mathrm{d}t\Big|\\ &\le r_n^{-2}\Big(\textstyle\int_{D_{\rho_n}}\big|\omega(\partial_s v,\partial_t v)-1\big|\;+\;\int_{D_{r_n}\setminus D_{\rho_n}}\big|\chi\big(\frac{\|v-p\|}{r_n}\big)\omega(\partial_s v,\partial_t v)-1\big|\Big)\\ &\le r_n^{-2}\Big(\textstyle\int_{D_{\rho_n}}\frac{4}{n}\,\mathrm{d}s\,\mathrm{d}t\;+\;\int_{D_{r_n}\setminus D_{\rho_n}}1\,\mathrm{d}s\,\mathrm{d}t\Big)\\ &=r_n^{-2}\Big(\tfrac{4}{n}\pi\rho_n^2+\pi r_n^2-\pi\rho_n^2\Big)\;=\;\tfrac{4\pi n^{-1}}{1-4n^{-1}}+\pi-\tfrac{\pi}{1-4n^{-1}}\quad\underset{n\to\infty}{\longrightarrow}\quad 0.\end{aligned}$$

Together with (24) this finishes the proof,

$$A(R)\;\ge\;R^2\,\lim_{n\to\infty}a(r_n)\;\ge\;R^2\,r_n^{-2}\textstyle\int_{D_{r_n}}\chi\big(\frac{\|v-p\|}{r_n}\big)v^*\omega\;\underset{n\to\infty}{\longrightarrow}\;R^2\pi.\qquad\square$$

Acknowledgments This project was initiated by the 'Women in Symplectic and Contact Geometry and Topology' workshop (WiSCon) at ICERM in 2019. We are deeply grateful to all the organizers and fellow participants for giving us the space and inspiration for this work! Further gratitude to Ben Filippenko and Wolfgang Schmaltz for helpful discussions.

This work is supported by Deutsche Forschungsgemeinschaft (DFG) under Germany?s Excellence Strategy EXC-2181/1 - 390900948 (the Heidelberg STRUCTURES Excellence Cluster), CRC/TRR 191 ?Symplectic Structures in Geometry, Algebra and Dynamics?, (Project-ID 281071066 ? TRR 191) and RTG 2229, and by the US National Science Foundation grants DMS-1708916 and DMS-1807270.

References

1. Benjamin Ackermann. Pseudo-holomorphic curves and Gromov's non-squeezing theorem. (unpublished lecture notes from 2015).
2. Tobias Holck Colding and William P. Minicozzi, II. *A course in minimal surfaces*, volume 121 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2011.
3. Oliver Fabert, Joel W. Fish, Roman Golovko, and Katrin Wehrheim. Polyfolds: A first and second look. *EMS Surv. Math. Sci.*, 3(2):131–208, 2016.
4. Benjamin Filippenko. Polyfold Regularization of Constrained Moduli Spaces. to appear in Journal of Symplectic Geometry, arXiv:1807.00386.

5. Benjamin Filippenko and Katrin Wehrheim. A Polyfold Proof of the Arnold conjecture. arXiv:1810.06180.
6. Victor Guillemin and Alan Pollack. *Differential topology*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1974.
7. M. Gromov. Pseudo holomorphic curves in symplectic manifolds. *Invent. Math.*, 82(2):307–347, 1985.
8. Christoph Hummel. *Gromov's compactness theorem for pseudo-holomorphic curves*, volume 151 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1997.
9. Helmut Hofer, Kris Wysocki, and Eduard Zehnder. Polyfold and Fredholm Theory. arXiv:1707.08941.
10. H. Hofer, K. Wysocki, and E. Zehnder. Applications of polyfold theory I: The polyfolds of Gromov-Witten theory. *Mem. Amer. Math. Soc.*, 248(1179):v+218, 2017.
11. Helmut Hofer and Eduard Zehnder. *Symplectic invariants and Hamiltonian dynamics*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, 1994.
12. Dusa McDuff. Groupoids, branched manifolds and multisections. *J. Symplectic Geom.*, 4(3):259–315, 2006.
13. Ieke Moerdijk. Orbifolds as groupoids: an introduction. In *Orbifolds in mathematics and physics (Madison, WI, 2001)*, volume 310 of *Contemp. Math.*, pages 205–222. Amer. Math. Soc., Providence, RI, 2002.
14. Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1998.
15. Dusa McDuff and Dietmar Salamon. *J-holomorphic curves and symplectic topology*, volume 52 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.
16. John Olsen. The Geometry of Möbius transformations. (unpublished lecture notes from 2010).
17. Wolfgang Schmaltz. Naturality of polyfold invariants and pulling back abstract perturbations. arXiv:1912.13370.
18. Katrin Wehrheim. Fredholm notions in scale calculus and Hamiltonian Floer theory. arXiv:1209.4040.
19. Katrin Wehrheim. *Uhlenbeck compactness*. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2004.
20. Chris Wendl. Lectures on Holomorphic Curves in Symplectic and Contact Geometry. arXiv:1011.1690.
21. Zhengyi Zhou. Quotient Theorems in Polyfold Theory and S^1-equivariant Transversality. *Proceedings of the London Mathematical Society*, 121(5):1337–1426, 2020.

Infinite Staircases for Hirzebruch Surfaces

Maria Bertozzi, Tara S. Holm, Emily Maw, Dusa McDuff, Grace T. Mwakyoma, Ana Rita Pires, and Morgan Weiler

1 Introduction

1.1 Overview of Results

Given a four-dimensional symplectic manifold (X, ω), we define its **ellipsoid embedding** function to be

M. Bertozzi
Ruhr Universität Bochum, Bochum, Germany
e-mail: maria.bertozzi@rub.de

T. S. Holm (✉)
Cornell University, Ithaca, NY, USA
e-mail: tsh@math.cornell.edu

E. Maw
University College London, London, UK

D. McDuff
Department of Mathematics, Barnard College, Columbia University, New York, NY, USA
e-mail: dusa@math.columbia.edu

G. T. Mwakyoma
Instituto Superior Técnico, Lisbon, Portugal
e-mail: grace.mwakyoma@tecnico.ulisboa.pt

A. R. Pires
University of Edinburgh, Edinburgh, UK
e-mail: apires@ed.ac.uk

M. Weiler
Rice University, Houston, TX, USA
e-mail: morgan.weiler@rice.edu

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_2

$$c_X(z) := \inf\left\{\lambda \;\middle|\; E(1,z) \overset{s}{\hookrightarrow} \lambda X\right\},$$

where $z \geq 1$ is a real variable, $\lambda X := (X, \lambda\omega)$ is the symplectic scaling of X, an ellipsoid $E(c,d) \subset \mathbb{C}^2$ is the set

$$E(c,d) = \left\{(\zeta_1,\zeta_2) \in \mathbb{C}^2 \;\middle|\; \pi\left(\frac{|\zeta_1|^2}{c} + \frac{|\zeta_2|^2}{d}\right) < 1\right\},$$

and we write $E \overset{s}{\hookrightarrow} \lambda X$ if there is a symplectic embedding of E into λX. One simple property of this function is that it is bounded below by the **volume curve**, because symplectomorphisms preserve volume:

$$c_X(z) \geq \sqrt{\frac{z}{\mathrm{vol(X)}}},$$

where vol(X) is the appropriately normalized volume of X. Similarly, the invariance of symplectomorphisms under scaling $\omega \mapsto \lambda\omega$ of the symplectic form implies the following scaling property

$$c_X(\lambda z) \leq \lambda c_X(z), \qquad z \geq 1, \lambda \geq 1; \tag{1.1.1}$$

indeed, if $E(1,z) \overset{s}{\hookrightarrow} X$ then $E(1,\lambda z) \subset \lambda(E(1,z)) \overset{s}{\hookrightarrow} \lambda X$.

The first ellipsoid embedding function to be computed was that of a ball B^4, by McDuff and Schlenk in [17]. They describe the following striking behavior: on the interval $[1, \tau^4]$ the function is piecewise linear, alternating between horizontal line segments and line segments that extend to lines passing through the origin in increasingly smaller intervals, creating what looks like an infinite staircase accumulating at τ^4, where $\tau = \left(1+\sqrt{5}\right)/2$ is the golden ratio. Similar behavior has been described for other targets X: the symmetric polydisk $B^2(1) \times B^2(1)$ in [8], certain convex toric domains in [7], and certain infinite families of polydisks $B^2(1) \times B^2(\beta)$ with irrational ratios β in [19]. For Delzant toric domains, [7] establishes that the open manifold has the same embedding function as its closure; for example the functions for the 4-ball and the projective plane $\mathbb{C}P^2$ are the same. Further, for all convex toric domains X of finite type, the capacity function exhibits packing stability: it coincides with the volume function for sufficiently large z (cf. [7, Prop. 2.1]). Moreover, the function $z \mapsto c_X(z)$ is piecewise linear when not identically equal to the volume constraint curve, and when its graph has infinitely many nonsmooth points lying above the volume curve, we say that c_X has an **infinite staircase**: see Definition 24.

For $0 \leq b < 1$, we denote by

$$H_b := \mathbb{C}P^2_1 \# \overline{\mathbb{C}P}^2_b$$

the one point blow up of $\mathbb{C}P^2$, where the line has symplectic area 1 and the exceptional divisor has symplectic area b. In light of [7, 17], we know that the ellipsoid embedding functions of $\mathbb{C}P^2$ and of $H_{1/3}$ have infinite staircases. In this note we report on the search for infinite staircases in the ellipsoid embedding functions of the Hirzebruch surfaces in H_b.

By Cristofaro-Gardiner et al. [7, Theorem 1.11], which applies because the moment polytope of H_b has rational conormals, we know that if c_{H_b} has an infinite staircase, then its accumulation point is at the point $z = \mathrm{acc}(b)$, the unique solution > 1 of the following quadratic equation involving b:

$$z^2 - \left(\frac{(3-b)^2}{1-b^2} - 2 \right) z + 1 = 0 \tag{1.1.2}$$

Furthermore, if an infinite staircase exists, then at $z = \mathrm{acc}(b)$ the ellipsoid embedding function must coincide with the volume constraint curve, that is we must have

$$c_{H_b}(\mathrm{acc}(b)) = \sqrt{\frac{\mathrm{acc}(b)}{1-b^2}} =: V_b(\mathrm{acc}(b)). \tag{1.1.3}$$

As shown in Fig. 1, the function $b \mapsto \mathrm{acc}(b)$ decreases for $b \in [0, 1/3)$, with minimum value $\mathrm{acc}(1/3) = 3+2\sqrt{2}$, and then increases. The nature of this function leads to an interesting and as yet not well understood interplay between the z-coordinates that parametrize the domain[1] and the b-coordinates parametrizing the target. The point $b = 1/3$ is the only known nonzero rational number such that H_b has a staircase, and according to [7, Conjecture 1.20] it is expected to be the only such value. In this paper, we focus on irrational values of b.

The capacity function c_{H_b} may be described either in terms of exceptional divisors $\mathbf{E}$ as recounted in Sect. 2 or in terms of ECH capacities as laid out in Sect. 4. Both approaches have added to our understanding in this project. ECH capacities gave us a way to understand the broad outlines of the problem and to discover interesting phenomena. Exceptional divisors are then well-adapted to establishing the intricate details; they also yield intriguing numerical data. The most relevant exceptional classes have "centers" at rational numbers that, in the case of a staircase, form the first coordinate of the corners of the steps. We now lay out two results that we obtained early on. For notation, see Sect. 2. We say that a staircase ascends (resp. descends) if the steps occur at increasing (resp. decreasing) values of the domain parameter z.

In the first result, we identify intervals of shape parameters b for which H_b cannot admit an infinite staircase. Exceptional divisors $\mathbf{B}$ give rise to **obstruction functions** $\mu_{b,\mathbf{B}}(z)$ that give lower bounds

[1] Our convention is that a general point in the domain is denoted by the letter z while special points (such as step corners, or break points of constraints) are denoted a.

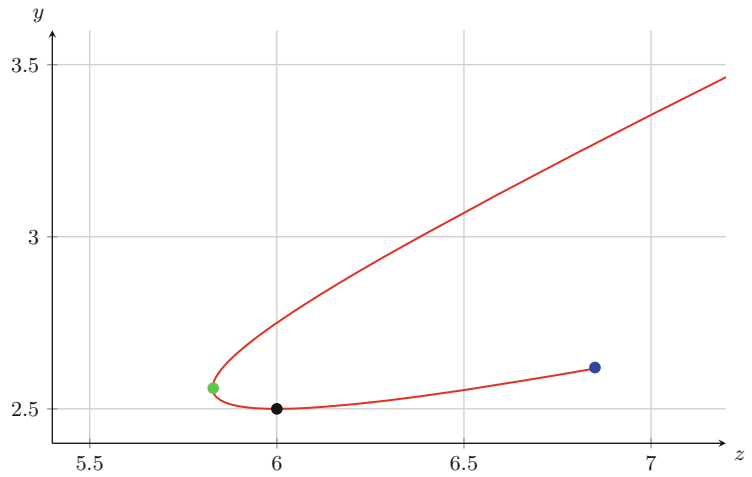


Fig. 1 This curve indicates the location of the accumulation point $(z, y) = \big(\mathrm{acc}(b), V_b(\mathrm{acc}(b))\big)$ for $0 \le b < 1$. The blue point with $b = 0$ is at (τ^4, τ^2) and is the accumulation point for the Fibonacci stairs. The green point at $(3 + 2\sqrt{2}, 1/3)$ is the accumulation point for the stairs in $H_{1/3}$, and is the minimum of the function $b \mapsto \mathrm{acc}(b)$. The black point at $(6, 5/2)$ corresponds to $b = 1/5$ and is the location where $V_b(\mathrm{acc}(b))$ takes its minimum

$$c_{H_b}(z) \ge \mu_{\mathbf{B},b}(z).$$

For a fixed class $\mathbf{B}$, then, we can let b vary and determine intervals of b for which

$$\mu_{\mathbf{B},b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b)),$$

violating (1.1.3) and thereby blocking the existence of a staircase for those b parameters.

Theorem 1 *For each $n \ge 0$, the exceptional divisor in class*

$$\mathbf{B}_n^U := (n+3)L - (n+2)E_0 - \sum_{i=1}^{2n+6} E_i$$

blocks the existence of a staircase for b in the interval

$$1/3 < \beta_{\mathbf{B}_n^U,\ell} := \frac{(2n^2+6n+3) - \sqrt{\sigma_n}}{2n^2+6n+2} < b < \frac{(n+3)\left(3n+7+\sqrt{\sigma_n}\right)}{5n^2+30n+44} =: \beta_{\mathbf{B}_n^U,u} < 1,$$

where $\sigma_n := (2n+1)(2n+5)$. Correspondingly, there is no staircase with accumulation point in the interval

$$\alpha_{\mathbf{B}_n^U,\ell} := \frac{\sigma_n + (2n+3)\sqrt{\sigma_n}}{2(2n+1)} < z < 6 + \frac{\sigma_n + (2n+3)\sqrt{\sigma_n}}{2(2n+5)} =: \alpha_{\mathbf{B}_n^U,u},$$

where $\alpha_{\mathbf{B}_n^U,\ell}$ *and* $\alpha_{\mathbf{B}_n^U,u}$ *have continued fraction expansions*

$$\alpha_{\mathbf{B}_n^U,\ell} = [\{2n+5, 2n+1\}^\infty], \qquad \alpha_{\mathbf{B}_n^U,u} = [2n+7; \{2n+5, 2n+1\}^\infty]$$

For a proof see the end of Sect. 3.2. We will say that an interval such as $\left(\beta_{\mathbf{B}_n^U,\ell}, \beta_{\mathbf{B}_n^U,u}\right)$ that is blocked by an exceptional divisor $\mathbf{B}$ is a **B-blocked** *b*-**interval**, and denote it by $J_{\mathbf{B}}$. The corresponding interval $\mathrm{acc}(J_{\mathbf{B}})$ on the z-axis is denoted $I_{\mathbf{B}} = (\alpha_{\mathbf{B},\ell}, \alpha_{\mathbf{B},u})$, and, if $J_{\mathbf{B}} \subset (1/3, 1)$ (resp. if $J_{\mathbf{B}} \subset [0, 1/3)$), consists of points that cannot be the accumulation point of any staircase for $b > 1/3$ (resp. $b < 1/3$). The class $\mathbf{B}$ itself is called a **blocking class**.

In Theorem 1, the blocked z-intervals $(\alpha_{\mathbf{B}_n^U,\ell}, \alpha_{\mathbf{B}_n^U,u})$ have 'centers' $2n$ at distance 2 apart, while their lengths increase with limit 2. Correspondingly most values of b close to 1 are blocked; see Fig. 4. Further, the above classes $\mathbf{B}_n^U$ are **center-blocking** in the sense of Definition 37, which implies by Proposition 42 that they also block the existence of a descending (resp. ascending) staircase that accumulates at the lower endpoint $\alpha_{\mathbf{B}_n^U,\ell}$ (resp. upper endpoint $\alpha_{\mathbf{B}_n^U,u}$) of $I_{\mathbf{B}}$.

In our second main result, proved in Sect. 3.3, we find a new infinite family of shape parameters for which the embedding capacity function does include an infinite staircase.

Theorem 2 *There is a decreasing sequence of parameter values* $b_n := b^E_{u,n,\infty}$ *in the interval* $(1/5, 1/3)$*, for* $n \geq 0$*, with limit* $1/5$*, so that each* H_{b_n} *admits a descending staircase* $\mathcal{S}^E_{u,n} = \left(\mathbf{E}^E_{u,n,k}\right)_{k\geq 0}$ *whose steps have exterior corners at the points* $z = a^E_{u,n,k}$ *with continued fraction expansions*

$$a^E_{u,n,k} = [5; 1, 2n+6, \{2n+5, 2n+1\}^k, 2n+4], \quad k \geq 0.$$

The accumulation point $\mathrm{acc}(b^E_{u,n,\infty})$ *is*

$$a^E_{u,n,\infty} = [5; 1, 2n+6, \{2n+5, 2n+1\}^\infty].$$

These accumulation points form an increasing sequence that converges to 6 *as* $n \to \infty$.

We found this family of staircases by trial and error, using a combination of the computer programs described in Sect. 5 and the algebro-geometric methods described in Remark 19.

Remark 3

(i) Note that, in contrast to the z-coordinates, the continued fraction expansions of the shape parameters $b^E_{u,n,\infty}$ seem to have no special properties. On the other hand, because the value σ_n that occurs in the formula for b_n is so closely related

to the iterated pair $2n+5, 2n+1$ in the continued fraction of the accumulation points $a^E_{u,n,\infty}$, one can show that the $a^E_{u,n,\infty}$ are also numbers in $\mathbb{Q} + \mathbb{Q}\sqrt{\sigma_n}$. From this point of view, the fact that the shape parameters $b^E_{u,n,\infty}$ also involve the surd $\sqrt{\sigma_n}$ is rather mysterious, but this follows because the new staircases accumulate at the end points of blocked z-intervals; see Corollary 45.

(ii) The staircases in Theorem 2 actually consist of two intertwining sequences of classes that are fully described in Theorem 58 below. In fact, all the new staircases that we find in this paper have this form. In Fig. 2, one can see the steps with centers $[5; 1, 5, 1, 4] = 199/34 \approx 5.853$ and $[5; 1, 5, 1, 5, 2] = 521/89 \approx 5.8539$, while in Fig. 3, one can see the steps with centers $[7; 5, 2] = 79/11 \approx 7.1818$ and $[7; 5, 1, 4] = 208/29 \approx 7.172$. The small step with center just > 7.190 is not part of the staircase. ♢

Rather surprisingly, the two theorems above are related both numerically (the same square roots and continued fractions occur) and geometrically, in ways that we explore in Sect. 2.4 and Sect. 3.1. In general, we make the following conjecture:

Conjecture 4 For each center-blocking class $\mathbf{B}$ with corresponding blocked b-interval $J_{\mathbf{B}} = (\beta_{\mathbf{B},\ell}, \beta_{\mathbf{B},u})$, both $H_{\beta_{\mathbf{B},\ell}}$ and $H_{\beta_{\mathbf{B},u}}$ admit staircases $\mathcal{S}_{\beta_{\mathbf{B},\ell}}, \mathcal{S}_{\beta_{\mathbf{B},u}}$ where $\mathcal{S}_{\beta_{\mathbf{B},\ell}}$ ascends and $\mathcal{S}_{\beta_{\mathbf{B},u}}$ descends if $J \subset (1/3, 1)$ and the reverse holds true if $J \subset [0, 1/3)$. Further, if b is irrational and H_b has a staircase, then b is the endpoint of some (center-)blocked b-interval $J_{\mathbf{B}}$.

Figures 2 and 3 illustrate Conjecture 4 for the center-blocking class $\mathbf{B}^U_0$ of Theorem 1.

As evidence for this conjecture, we prove the following theorem in Sect. 3.3.

Theorem 5 *For each $n \geq 1$ there is a center-blocking class $\mathbf{B}^E_n$ blocking b-values $J_{\mathbf{B}^E_n} \subset (0, 1/3)$ such that the accumulation point $a^E_{u,n,\infty}$ of the decreasing staircase $\mathcal{S}^E_{u,n}$ in Theorem 2 is the upper endpoint $\alpha_{\mathbf{B}^E_n,u}$ of the corresponding z-interval $I_{\mathbf{B}^E_n}$.*

It turns out that this second set of center-blocking classes $(\mathbf{B}^E_n)_{n\geq 1}$ is related to the first set $(\mathbf{B}^U_n)_{n\geq 0}$ by a symmetry (described in Corollary 60) that arises from the properties of the recursive sequence $1, 6, 35, 204, \ldots$. In fact, in Sect. 3.1 we define three families of center-blocking classes, the two $(\mathbf{B}^U_n)$, $(\mathbf{B}^E_n)$ that are mentioned above, together with one more called $(\mathbf{B}^L_n)$, that are all related by symmetries. We also describe the six associated families of staircases, with full proofs of our claims given in Sects. 3.2–3.4. As explained in Remark 61, these symmetries generate yet more staircases, that will be explored more fully in our next paper.

As we show in Sect. 3.1, these new staircases are variants of the Fibonacci staircase for the ball, and belong to a class of staircases (called **pre-staircases** and defined in Definition 46) made up of exceptional classes whose entries are determined by the centers of the steps together with a linear relation. As we show in Theorem 52, the coefficients of this relation are in turn determined by the parameters of the associated center-blocking class. Thus the structure of these staircases is very different from that of the staircase at $b = 1/3$ (described in Example 35 on p. 76), which is not associated to any blocking class.

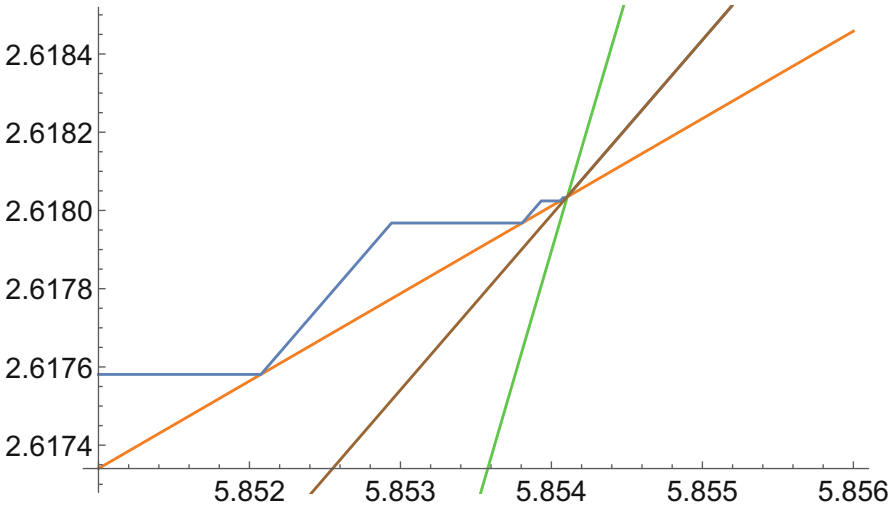


Fig. 2 In blue is the plot of a lower bound for c_{H_b}, where $b = \beta_{\mathbf{B}_0^U,\ell} = \left(3-\sqrt{5}\right)/2 > 1/3$, which contains the infinite staircase with the same numerics as the $n = 0$ case of the family $\mathcal{S}^U_{\ell,n}$ (Theorem 56) whose accumulation point has a z-value with continued fraction $[\{5,1\}^\infty]$; see Remark 57. The volume constraint V_b is in orange, and the parametrized curve $\big(\mathrm{acc}(b), V_b(\mathrm{acc}(b))\big)$, which intersects the volume constraint at the accumulation point of $\mathcal{S}^U_{\ell,0}$, is in green. The obstruction $\mu_{\mathbf{B}_0^U,b}$ (see (2.1.6)) from $\mathbf{B}_0^U$ is in brown

Our last main result concerns the point $b = 1/5$ that has associated accumulation point $\mathrm{acc}(1/5) = 6$. It is relatively straightforward to show that $H_{1/5}$ is **unobstructed**, that is its capacity and volume functions agree at the point $z = \mathrm{acc}(b)$; see Examples 22, 28 (i). As we remarked above, this is a necessary condition for H_b to admit a staircase; however, as the following result shows, it is not sufficient.

Theorem 6 *$H_{1/5}$ is unobstructed, but does not have a staircase.*

We prove this fact in Sect. 4.3, using ECH tools developed in [4].

The fundamental question that concerns us here is that of the structure of the sets

$$\mathit{Stair} := \{b \mid H_b \text{ has a staircase}\} \subset [0,1), \quad \text{and}$$

$$\mathit{Block} := \bigcup \left\{ J_{\mathbf{B}} \;\middle|\; \mathbf{B} \text{ is a blocking class} \right\} \subset [0,1).$$

Block is a union of open intervals, and as far as we know it could be dense in $[0,1)$. Indeed, Proposition 49 shows that most (if not all) of the classes in the new staircases are themselves blocking classes. Note also that if Conjecture 4 is true, then H_b is unobstructed at each end of a center-blocked b-interval (provided these are irrational), so that each such class determines a single connected component of *Block*. It is unknown whether there are other components of *Block*. Further, *Stair* is

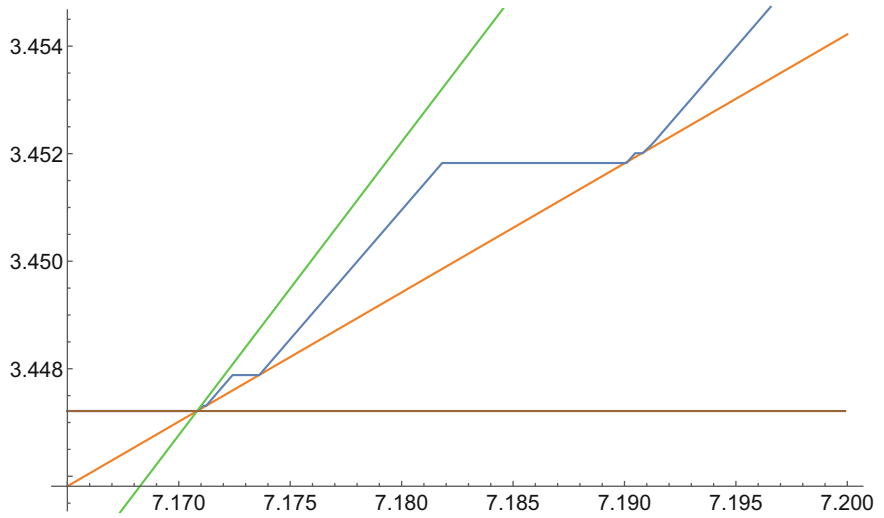


Fig. 3 In blue is the plot of a lower bound for c_{H_b} where $b = \beta_{\mathbf{B}_0^U,u}$, which contains the infinite staircase $\mathcal{S}_{u,0}^U$ (Theorem 56) whose accumulation point has a z-value with continued fraction $[7; \{5, 1\}^\infty]$. The volume constraint V_b is in orange, and the parameterized curve $\big(\mathrm{acc}(b), V_b(\mathrm{acc}(b))\big)$, which intersects the volume constraint at the accumulation point of $\mathcal{S}_{u,0}^U$, is in green. The obstruction $\mu_{\mathbf{B}_0^U,b}$ (see (2.1.6)) from the blocking class $\mathbf{B}_0^U$ is in brown

contained in the closed set $[0, 1)\setminus \mathit{Block}$, but is not equal to it by Theorem 6. This result also shows that the subset *Stair* is not closed, since Theorem 2 shows that $1/5$ is the limit point of a sequence in *Stair* (and also the limit of a sequence of endpoints of blocked b-intervals by Theorem 5), while $1/5 \notin \mathit{Stair}$ by Theorem 6.

We conjecture above that the irrational points in *Stair* are precisely the union of the endpoints of the center-blocked intervals in *Block*, while, by Cristofaro-Gardiner et al. [7, Conj. 1.20], the only rational points in *Stair* are believed to be 0 and $1/3$. However, at the moment we cannot say anything more about these sets. There are many open questions here. For example, are there any blocking classes that are not center-blocking? Do all staircases (except for the one at $b = 1/3$) have the special structure of a pre-staircase? We will explore such questions in more depth in our forthcoming manuscript.

Remark 7 All the staircases that we mention here are generated by families of perfect exceptional classes in the sense of Definition 12. Hence, the proof of [6, Prop.3.6.1] should generalize to show that they all stabilize when we multiply both domain and target by $\mathbb{C}^k$. ◊

1.2 Organization of the Paper

In Sect. 2, we first explain how to define the embedding capacity function c_{H_b} from knowledge of exceptional spheres. We then develop the general properties of exceptional classes, contrasting the notion of a **quasi-perfect** class (which is numerically defined) with that of a **perfect** class (which satisfies an addition geometric condition); see Definition 12. Proposition 21 shows how to tell when a class is **live**, i.e. contributes to the capacity function; while Propositions 26 and 30 formulate our basic staircase recognition criteria. After introducing the notion of a **blocking class** in Sect. 2.3, we show in Proposition 42 that suitable knowledge about the capacity function forces a quasi-perfect blocking class to be perfect. We also begin to analyse the relation between blocking classes and staircases. Finally Sect. 2.4 introduces the notion of a **pre-staircase** (Definition 46), which is a numerically defined set of classes of special form. The main results are Theorem 51 which gives criteria for a pre-staircase to be an actual staircase, and Theorem 52 which explains how the existence of suitable perfect pre-staircases that converge to the ends of a potential blocking class forces that class to be a perfect blocking class.

The next section describes the three staircase families in Theorems 54, 56 and 58. This can be read once the basic definitions are understood. The rest of the section develops the techniques necessary to prove that these really are staircases. There are two parts to this proof. The first part (in Sects. 3.2 and 3.3) involves showing that the staircase classes are quasi-perfect and that the estimates in Theorem 51 hold. The second part is a proof that the staircase classes are perfect, in other words that they satisfy the necessary geometric condition (reduction under Cremona moves) to be true (rather than fake) exceptional classes. Here the proof is greatly eased by the existence of the symmetries described in Corollary 60 that transform one staircase into another: see Proposition 86.

Section 4 describes a different (but equivalent) way to compute the capacity function, this time using capacities coming from embedded contact homology. These "ECH capacities" are the most efficient way to do certain calculations, as is shown in the proof in Sect. 4.3 that c_{H_b} has no infinite staircase at $b = 1/5$. In Sect. 4.2 we also prove Lemma 92, which explains the relationship between the obstructions from perfect classes and those from ECH capacities. This allowed us to identify live classes using computer models based on ECH capacities, as explained in Sect. 5.4.

Finally, Sect. 5 explains the computer programs that we used for our graphical explorations. In Sect. 5.1 we explain the combinatorics of lattice paths which allowed us to reduce computation time; our methods should generalize to any convex toric domain with two sides with rational slopes. After discussing our actual Mathematica code in Sects. 5.2 and 5.3, in Sect. 5.4 we explain our strategy for identifying values of b for which c_{H_b} has an infinite staircase. We conclude with plots of c_{H_b} which illustrate several of the major phenomena we discuss in this paper. Remark 19 describes how the computer-aided approach of Sect. 5 interacted fruitfully with the methods of Sect. 2 in our search for potential staircases.

2 Embedding Obstructions from Exceptional Spheres

In this section, we explain how to define the embedding capacity function c_{H_b} using exceptional spheres. The first subsection explains the general properties of exceptional classes, and the second discusses the properties of quasi-perfect and perfect classes, and formulates our first staircase recognition criteria (in Propositions 26 and 30). The next subsection explains the properties of blocking classes, while the last introduces the notion of a pre-staircase and, by combining their properties with those of blocking classes, gives in Theorem 51 a powerful staircase recognition criterion and in Theorem 52 a blocking class recognition criterion.

2.1 The Role of Exceptional Spheres

It was shown in [15] that the existence of a symplectic embedding of a rational ellipsoid $E(1, z)$ into either the projective plane $\mathbb{C}P^2$ or another rational ellipsoid is equivalent to a particular embedding of balls into $\mathbb{C}P^2$. This result extends trivially to embeddings of ellipsoids into any blow up of $\mathbb{C}P^2$; with some additional argument to embeddings into rational convex toric domains (see [3, Theorem 1.6] and [7, Theorem 1.2]); and into certain other toric domains (see [14]). In our framework, when z is rational there is a symplectic embedding

$$E(1,z) \overset{s}{\hookrightarrow} \lambda H_b \iff \bigsqcup_{j=1}^{n} B(w_j) \overset{s}{\hookrightarrow} \lambda H_b \iff \bigsqcup_{j=1}^{n} B(w_j) \bigsqcup B(\lambda b) \overset{s}{\hookrightarrow} B(\lambda). \tag{2.1.1}$$

Here, $\mathbf{w}(z) = (w_1, \ldots, w_n)$ is the **weight expansion** of $z \in \mathbb{Q}$ defined in (2.1.3) below. Thus the ellipsoid embedding question converts into a question about symplectic forms on an $(n+1)$-fold blowup[2] of $\mathbb{C}P^2$, a problem that was solved in the 1990s.

The natural basis for $H_2\left(\mathbb{C}P^2\#(n+1)\overline{\mathbb{C}P^2}\right)$ consists of the original generator L on $\mathbb{C}P^2$, plus the classes of the $n+1$ exceptional spheres $E_0, \ldots, E_n$. It is known that there is a ball embedding as in (2.1.1) if and only if there is a symplectic form ω on $\mathbb{C}P^2\#(n+1)\overline{\mathbb{C}P^2}$ satisfying

$$PD[\omega] = \lambda L - \lambda b E_0 - \sum_{i=1}^{n} w_i E_i$$

with canonical class K satisfying

[2]For discussion of the relationship between ball embeddings and the symplectic blow up, see for example, [18, Ch.7.1].

$$PD(K) = -3L + E_0 + \sum_{i=1}^{n} E_i.$$

We set $\mathscr{E}_{n+1}$ to be the set of **exceptional classes** in $H_2\left(\mathbb{C}P^2\#(n+1)\overline{\mathbb{C}P^2}\right)$. It consists of those classes A that have square -1, $A \cdot K = -1$, and that can be represented by a smoothly embedded, symplectic sphere. Li and Li have shown that the requirement that a sphere be symplectic is superfluous [12]; thus $\mathscr{E}_{n+1}$ is independent of ω (though it depends on K). An exceptional class

$$\mathbf{E} := (d, m; \mathbf{m}) := dL - mE_0 - \sum_{i\geq 1} m_i E_i$$

must satisfy the following Diophantine equations, which correspond to the homological conditions $A^2 = -1$ and $A \cdot K = -1$:

$$d^2 - m^2 - \sum_{i=1}^{n} m_i^2 = -1 \text{ and } 3d - m - \sum_{i=1}^{n} m_i = 1. \tag{2.1.2}$$

We call classes that satisfy these two equations (2.1.2) **Diophantine classes**. As explained in Sect. 3.4, Diophantine classes are exceptional if and only if they reduce to the class E_1 under repeated Cremona moves.

Li-Liu [13] proved that the symplectic cone of $\mathbb{C}P^2\#(n+1)\overline{\mathbb{C}P^2}$ consists of classes $[\omega]$ such that $[\omega]^2 > 0$ and $PD([\omega]) \cdot E > 0$ for all $E \in \mathscr{E}_{n+1}$. It follows that there is a symplectic embedding (2.1.1) if and only if

- the volume of $E(1, z)$ is less than that of λH_b, that is $z < \lambda^2(1 - b^2)$;
- for each exceptional class $\mathbf{E} := (d, m; \mathbf{m}) \in \mathscr{E}_{n+1}$, we have $\omega(E) > 0$, that is $mb + \sum_{i=1}^{n} m_i \frac{w_i}{\lambda} < d$; or equivalently

$$\lambda > \frac{\sum_{i=1}^{n} m_i w_i}{d - mb} = \frac{\mathbf{m} \cdot \mathbf{w}(z)}{d - mb}.$$

Thus, the exceptional classes provide potential obstructions to the desired embeddings.

For any exceptional class $\mathbf{E} = (d, m; \mathbf{m})$, we may vary z to get an obstruction function as follows. Any real number z has (possibly infinitely long) **weight expansion** $\mathbf{w}(z) = (w_1, w_2, \dots)$ defined for $z > 1$ as follows:

- $w_1 = 1$ and $w_n \geq w_{n+1} > 0$ for all n,
- if $w_i > w_{i+1} = \cdots = w_n$ (where we set $w_0 := z$), then

$$w_{n+1} = \begin{cases} w_{i+1} & \text{if } \ w_{i+1} + \cdots + w_{n+1} = (n-i+1)w_{i+1} < w_i, \\ w_i - (n-i)w_{i+1} & \text{otherwise} \end{cases} \tag{2.1.3}$$

Thus $\mathbf{w}(z)$ starts with $\lfloor z \rfloor$ copies of 1 (where $\lfloor z \rfloor$ is the largest integer $\leq z$), is finite for rational z, and satisfies the recursive relation

$$\mathbf{w}(z) = \Big(1^{\times \lfloor z \rfloor}, (z - \lfloor z \rfloor)\mathbf{w}\big(\frac{1}{z - \lfloor z \rfloor}\big)\Big),$$

where we take $\mathbf{w}\big(\frac{1}{z-\lfloor z \rfloor}\big)$ to be empty if $z = \lfloor z \rfloor$. If we think of the weight expansion as an infinitely long sequence $\mathbf{w}(z) = (w_1, w_2, \dots)$, appending infinitely many zeros if necessary, then the coordinates of the weight expansion are continuous (but not differentiable!) functions $w_i(z)$ of z.

Example 8 If $z = p/q$ (in lowest terms) and $k_0 < p/q < k_0 + 1$, then

$$q\mathbf{w}(p/q) = \Big(q^{\times k_0}, (p - k_0 q)^{\times k_1}, (q - k_1(p - k_0 q))^{\times k_2}, \dots, 1\Big) \tag{2.1.4}$$

where z has continued fraction expansion $[k_0; k_1, k_2, \dots]$, i.e. the multiplicities of the entries in $\mathbf{w}(z)$ give the continued fraction expansion of z. Notice that all the entries here are integral and that the last entry is 1. Conversely, the continued fraction expansion of z determines the (continuous!) linear functions of $z = p/q$ that give the weight expansion. For example,

$$\mathbf{w}\Big(\frac{35}{6}\Big) = \Big(1^{\times 5}, \frac{5}{6}, \frac{1}{6}^{\times 5}\Big), \qquad \frac{35}{6} = [5; 1, 5] = 5 + \frac{1}{1 + \frac{1}{5}}.$$

We call $q\mathbf{w}(p/q)$ the **integral weight expansion** of $p/q = [\ell_0; \ell_1, \dots, \ell_N]$. An easy inductive argument on the length N shows that the integers p, q found by this inductive process from a continued fraction $[\ell_0; \ell_1, \dots, \ell_N]$ are always relatively prime. For more details, see [17, §1.2], and also Example 13 below. ♢

As in [17, §1.2], it is not hard to check the following identities for the weight expansion $\mathbf{w}(z)$ at $z = p/q$:

$$\sum_i w_i = z + 1 - \frac{1}{q}, \qquad \sum_i w_i^2 = \frac{p}{q}. \tag{2.1.5}$$

To define the **obstruction function** $z \mapsto \mu_{\mathbf{E},b}(z)$ associated to $\mathbf{E}$, we first pad $(d, m; \mathbf{m})$ with zeros on the right to make it infinitely long, and then set

$$\mu_{\mathbf{E},b}(z) := \mu_{(d,m;\mathbf{m}),b}(z) := \frac{\mathbf{m} \cdot \mathbf{w}(z)}{d - mb}. \tag{2.1.6}$$

The above discussion implies that these obstruction functions, together with the volume, completely determine the embedding capacity function c_{H_b}, namely

$$c_{H_b}(z) = \sup_{\mathbf{E}\in\mathcal{E}} \big(\mu_{\mathbf{E},b}(z),\ V_b(z)\big), \tag{2.1.7}$$

where $V_b(z) := \sqrt{\frac{z}{1-b^2}}$ denotes the volume obstruction.

Remark 9 A key point in our analysis of staircases is the result in [7, Theorem 1.8] that the geometry of the target H_b determines the accumulation point $\mathrm{acc}(b)$ of any staircase via (1.1.2). This result is valid (and the proof is the same) when the target X is any blow up of $\mathbb{CP}^2_1$ by balls of sizes $b_1, \dots, b_N$, provided that we replace $3-b$ by $3-\sum_i b_i$ and $1-b^2$ by the volume $1-\sum_i b_i^2$. ◊

Definition 10 An obstruction $\mu_{\mathbf{E},b}$ is said to be **nontrivial at** z if $\mu_{\mathbf{E},b}(z) > V_b(z)$, and **live at** z if it is nontrivial and equals the capacity function c_{H_b} at z. Further, it is said to **block** b if it is nontrivial at the accumulation point $\mathrm{acc}(b)$, that is, if

$$\mu_{\mathbf{E},b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b)).$$

We say that H_b is **unobstructed at** $\mathrm{acc}(b)$ (or simply **unobstructed**) if

$$c_{H_b}(\mathrm{acc}(b)) = V_b(\mathrm{acc}(b)).$$

Otherwise, we say that H_b is **obstructed** or, equivalently, **blocked**. Further, a b-interval J is **blocked** if $c_{H_b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b))$ for all $b \in J$.

Since all functions involved are continuous, the condition that $\mu_{\mathbf{E},b}$ is nontrivial at z is continuous in both z and b. Similarly, the condition that $\mu_{\mathbf{E},b}$ blocks b is an open condition on b, so that the set

$$\mathit{Block} := \{b \in [0,1) \mid H_b \text{ is blocked}\} \tag{2.1.8}$$

is open. Thus the set of b that are unobstructed is closed and contains the set

$$\mathit{Stair} = \{b \in [0,1) \mid H_b \text{ has a staircase}\}.$$

Theorem 6 guarantees that *Stair* is a proper subset of the unobstructed values.

Observe that the condition of being live at (z_0, b_0) need not be continuous with respect to either variable z or b since it might happen that two different classes are live at (z_0, b_0), one remaining live as z or b increases and the other remaining live as z or b decreases. If b is kept fixed, there is no known example of this behavior at an outward corner[3] of the capacity function, though it does happen at other corners provided that these are not on the volume curve.

[3] I.e. a nonsmooth point where the slope decreases.

Remark 11 Hutchings [9] was the first to observe that the inequalities defining the capacity function do not require that $\mathbf{E} \in \mathcal{E}$ but only that $\mathbf{E}$ be Diophantine: that its entries satisfy (2.1.2). However, it is sometimes very important to know that a class does lie in $\mathcal{E}$ because any two such classes have nonnegative intersection. See, for example, Proposition 21. ◊

Definition 12 Let $\mathbf{E} := (d, m; \mathbf{m})$ be a Diophantine class. If there is a rational number $a = p/q$ such that $\mathbf{m} = q\mathbf{w}(a)$, then we say that $\mathbf{E}$ has **center** $a = p/q$ and (as in [19, Def.1.4]) call the class **quasi-perfect**. If in addition $\mathbf{E}$ is an exceptional class, then we say that $\mathbf{E}$ is **perfect**. A class that is quasi-perfect but not perfect, is called **fake**.

Finally, for a quasi-perfect, resp. perfect, class $\mathbf{E}$, if $|bd - m| < \sqrt{1 - b^2}$ then we say that $\mathbf{E}$ is b **-quasi-perfect**, resp. b**-perfect**.

Example 13 We now show how to compute an obstruction function on a small interval. Let $a = 35/6$ with $\mathbf{w}(35/6) = (1^{\times 5}, 5/6, (1/6)^{\times 5})$ as in Example 8. It is not hard to check that there is a corresponding exceptional class $\mathbf{E}_3 = (15, 4; 6\mathbf{w}(35/6))$. If $35/6 \le z < 6$ then $\mathbf{w}(z) = (1^{\times 5}, z - 5, (6 - z)^{\times k}, \dots)$ for some $k \ge 5$ and we have $\mathbf{m} \cdot \mathbf{w}(z) = 35$. On the other hand, if $z < 35/6$, then its weight expansion is $(1^{\times 5}, z - 5, (6 - z)^{\times 4}, (5z - 29)^{\times k}, \dots)$ provided that $4(6 - z) < z - 5 < 5(6 - z)$, which is equivalent to $29/5 < z < 35/6$. In this case

$$\mathbf{m} \cdot \mathbf{w}(z) = 30 + 5(z - 5) + 4(6 - z) + (5z - 29) = 6z.$$

Thus

$$\mu_{\mathbf{E},b}(z) = \begin{cases} \frac{6z}{15-4b} & \text{if } \frac{29}{5} < z < \frac{35}{6} \\ \frac{35}{15-4b} & \text{if } \frac{35}{6} \le z < 6. \end{cases}$$

For a more general result, see Lemma 16 below. ◊

The next task is to determine whether a particular exceptional class $\mathbf{E}$ gives us a useful embedding obstruction. We denote by $\mathcal{E} := \bigcup_n \mathcal{E}_n$ the set of exceptional classes in arbitrary blow ups of $\mathbb{C}P^2$, and will define the **length** $\ell(\mathbf{m})$ of $\mathbf{m}$ (resp. $\ell(z)$ of z) to be the number of nonzero entries in the vector $\mathbf{m}$ (resp. $\mathbf{w}(z)$). The following lemma, taken from [7, Lemma 2.21 & Prop.2.23], extends results in [17] about embeddings into a ball to the case of a general convex toric target such as H_b.

Lemma 14 *Let* $\mathbf{E} = (d, m; \mathbf{m})$ *be a Diophantine class with corresponding obstruction function* $\mu_{\mathbf{E},b}$.

(i) *The function* $\mu_{\mathbf{E},b}(z)$ *is continuous with respect to both variables* z, b, *and for fixed* b, *is piecewise linear with rational coefficients.*

(ii) *For any* b, *the obstruction* $\mu_{\mathbf{E},b}(z)$ *given by* $\mathbf{E} = (d, m; \mathbf{m})$ *is nontrivial at* z *only if* $\ell(z) \ge \ell(\mathbf{m})$.

(iii) *Any maximal interval I on which $\mu_{\mathbf{E},b}$ is nontrivial contains a unique point a, called the* ***break point****, with $\ell(a) = \ell(\mathbf{m})$. Moreover, the function*

$$I \to (0,\infty), \quad z \mapsto \mu_{\mathbf{E},b}(z) - V_b(z)$$

reaches its maximum at the break point a.

Proof Denote by $w_i(z)$ the ith weight of z considered as a function of z. Then, by (2.1.3) it is piecewise linear, and is linear on any open interval that does not contain an element a' with $\ell(a') \le i$. This implies (i).

Property (ii) is an easy consequence of the inequality

$$\sqrt{d^2-m^2}\,\sqrt{1-b^2} \le d - mb, \tag{2.1.9}$$

which, after squaring both sides simply reduces to $0 \le (db-m)^2$.

For property (iii), we note that the graph of $V_b(z) = \sqrt{\frac{z}{1-b^2}}$ is concave down, so the function $\mu_{\mathbf{E},b}(z)$ cannot be linear on any maximal interval on which it is nontrivial. On the other hand, it follows from (2.1.3) that the formula for $w_i(z)$ can change only if it, or one of the earlier weights, becomes zero. From this, it is not hard to see that $\mu_{\mathbf{E},b}$ is linear on every open interval on which $\ell(z) > \ell(\mathbf{m})$. The above remarks imply that I must contain at least one break point a, which occurs when $\ell(a) = \ell(\mathbf{m})$. Moreover this point a is unique because if $\ell(a) = \ell(a')$ for some $a < a'$, then it is shown in [17, Lem.2.1.3] that there is $z \in (a, a')$ with $\ell(z) < \ell(a) = \ell(\mathbf{m})$, contradicting property (ii).

Note that part (iii) of the following lemma shows that, if $\mathbf{E}$ is perfect with center a, then it is b-perfect in the sense of Definition 12 exactly if $\mu_{\mathbf{E},b}$ is nontrivial at a. Further, part (ii) establishes that the capacity function $c_{H_b}(z)$ is the maximum, rather than just the supremum, of the obstruction functions $\mu_{\mathbf{E},b}(z)$.

Lemma 15 *Let $\mathbf{E} = (d, m; \mathbf{m})$ be a Diophantine class.*

(i) *If $\mu_{\mathbf{E},b}$ is nontrivial at z, we must have $|bd - m| < \sqrt{1-b^2}$. Furthermore,*

$$\mu_{\mathbf{E},b}(z) \le V_b(z)\sqrt{1 + \frac{1}{d^2-m^2}}.$$

(ii) *If $c_{H_b}(z) > V_b(z)$, there is an exceptional class $\mathbf{E}$ such that $c_{H_b}(z) = \mu_{\mathbf{E},b}(z)$. Moreover, c_{H_b} is piecewise linear on any interval on which it is bounded away from the volume. Finally, if $J \subset [0,1)$ is an open subset such that $c_{H_b}(z) > V_b(z) + \varepsilon$ for all $b \in J$ and some $\varepsilon > 0$, then for all b outside of a finite subset of J, we may assume that $c_{H_{b'}}(z) = \mu_{\mathbf{E},b'}(z)$ for all b' in some neighborhood J_b of b.*

(iii) *If $\mathbf{E}$ has center $a = p/q$, then $\mu_{\mathbf{E},b}(z)$ is nontrivial at $z = a$ if and only if $|bd - m| < \sqrt{1-b^2}$. Further, when $b_0 = m/d$, we have*

$$\mu_{\mathbf{E},b_0}(a) = \frac{pd}{d^2-m^2} \le \frac{d}{q}\Big(1 + \frac{2}{pq}\Big). \tag{2.1.10}$$

Proof The obstruction $\mu_{\mathbf{E},b}$ satisfies

$$(d - mb)\,\mu_{\mathbf{E},b}(z) = \mathbf{m}\cdot\mathbf{w}(z) \le \|\mathbf{m}\|\,\|\mathbf{w}\| = \sqrt{z}\sqrt{d^2 - m^2 + 1}. \tag{2.1.11}$$

But $\sqrt{z}\frac{\sqrt{d^2-m^2+1}}{d-mb} > V_b(z)$ if

$$\frac{\sqrt{d^2 - m^2 + 1}}{d - mb} > \frac{1}{\sqrt{1 - b^2}}.$$

This happens if

$$\begin{aligned}
(d^2 - m^2 + 1)(1 - b^2) &> (d - mb)^2, \quad i.e.\\
d^2 - m^2 + 1 - b^2d^2 + b^2m^2 - b^2 &> d^2 - 2mdb + m^2b^2, \quad i.e.\\
1 - b^2 &> d^2b^2 - 2mdb + m^2, \quad i.e.\\
1 - b^2 &> (db - m)^2.
\end{aligned}$$

This proves the first claim in (i).

Combining the inequality $d - mb \ge \sqrt{d^2 - m^2}\ \sqrt{1 - b^2}$, from (2.1.9) with (2.1.11), we have

$$\mu_{\mathbf{E},b}(z) \le \frac{\sqrt{z}\sqrt{d^2 - m^2 + 1}}{\sqrt{d^2 - m^2}\,\sqrt{1 - b^2}} = V_b(z)\sqrt{1 + \frac{1}{d^2 - m^2}}.$$

This completes the proof of (i).

To prove (ii), we first show that if $c_{H_b}(z) = (1 + \varepsilon')V_b(z)$ for some $\varepsilon' > 0$ and a fixed value of b, then there are only finitely many exceptional classes $\mathbf{E}$ such that $\mu_{\mathbf{E},b}(z) > (1 + \varepsilon'/2)V_b(z)$. To this end, notice that (i) implies that $\sqrt{1 + \frac{1}{d^2-m^2}} > 1 + \varepsilon/2 > \sqrt{1 + \varepsilon}$, so that we must have $\frac{1}{d^2-m^2} > \varepsilon$. Thus there is a constant C so that $d^2 - m^2 < C$.

Since we must also have $|db - m| < 1$ by (i), we are then looking for pairs of non-negative integers (d, m) which lie in a compact region of the plane, so there is an upper bound for the degrees d of the relevant classes. Further, we must have $\ell(\mathbf{m}) \le \ell(z)$ by Lemma 14. Therefore there are only a finite number of relevant classes. Hence, for each b, there is equality for at least one such class.

Moreover, if we now consider sufficiently small neighborhoods I' of z and J' of b, then the inequality $\mu_{\mathbf{E},b'}(z') > (1 + \varepsilon/2)V_{b'}(z')$ holds for all $z' \in I'$ and $b' \in J'$, so that again there are only finitely many classes $\mathbf{E}$ such that $c_{H_{b'}}(z') = \mu_{\mathbf{E},b'}(z')$ for some $z' \in I'$ and $b' \in J'$. By Lemma 14, each function $z \mapsto \mu_{\mathbf{E},b'}(z)$ is piecewise linear and hence with a finite number of nonsmooth points; see also Lemma 16. Therefore the same is true for $z' \mapsto c_{H_{b'}}(z')$ on any interval on which it is bounded away from the volume function. Thus for fixed b', these functions $c_{H_b'}$ and $\mu_{\mathbf{E},b'}$

of z' either agree at a single point or agree on a finite union of closed z'-intervals. Moreover, as b' changes these z'-intervals change continuously. Hence if we now consider our fixed z, all but finitely many b' values lie in the interior of one of these intervals. This completes the proof of item (ii).

The first claim in (iii) follows from the first calculation in (i) since in this case, the inequality in (2.1.11) is an equality. Further, because $d^2 - m^2 = pq - 1$ and $\mathbf{m} \cdot \mathbf{w}(a) = q\mathbf{w}(a) \cdot \mathbf{w}(a) = p$, we have

$$\mu_{\mathbf{E},m/d}(a) = \frac{pd}{d^2 - m^2} = \frac{pd}{pq-1} = \frac{d}{q}\frac{1}{1-1/pq} \le \frac{d}{q}\Big(1 + \frac{2}{pq}\Big),$$

where the last inequality is a consequence of the fact that $\frac{1}{1-1/x} \le 1 + \frac{2}{x}$ for $x \ge 2$. This completes the proof.

The calculation in Example 13 generalizes as follows.

Lemma 16 *Let $\mathbf{E} = (d, m; \mathbf{m})$ be a quasi-perfect class with center $a = p/q$. Then the corresponding capacity function $\mu_{\mathbf{E},b}$ has the following form near a: there are numbers $z_1 < a < z_2$ such that*

$$\mu_{\mathbf{E},b}(z) = \begin{cases} \frac{qz}{d-mb} & \text{if } \ z_1 < z < a \\ \frac{p}{d-mb} & \text{if } \ a \le z < z_2. \end{cases} \tag{2.1.12}$$

In particular, if $\mu_{\mathbf{E},b}(a) > V_b(a)$ then this formula applies throughout the neighborhood of a on which $\mu_{\mathbf{E},b}$ is nontrivial.

Proof Consider the continued fraction expansion $a = [\ell_0; \ell_1, \dots, \ell_N]$. We will carry out the proof when N is odd, since this is the case for the exceptional classes $\mathbf{E}_{n,k}$ that give our new staircases. The proof for N even is similar.

When N is odd, by McDuff and Schlenk [17, Lem.2.2.1] there is $\varepsilon > 0$ and $h > 0$ so that $z \in (a - \varepsilon, a)$ has weight expansion

$$\mathbf{w}(z) = \big(1^{\times \ell_0}, (x_1(z))^{\times \ell_1}, \dots, (x_N(z))^{\times \ell_N}, \ (x_{N+1}(z))^{\times h}, \dots, \big), \tag{2.1.13}$$

where the functions $x_j(z) = \alpha_j + z\beta_j$ are linear functions of z, that increase w.r.t. z when j is odd and decrease when j is even. Denote the kth convergent to a by p_k/q_k so that $p_0/q_0 = \lfloor a \rfloor$ and $\partial_n/q_n = p/q = a$. Then [17, Lem.2.2.3 (ii), (iii)] shows that

$$|\alpha_j| = p_{j-1}, \qquad |\beta_j| = q_{j-1} \qquad \text{for } \ 1 \le j \le N+1.$$

Further, [17, Lem.2.2.3 (i)] shows that the weight expansion $\mathbf{w}(a)$ may be written

$$\big(1^{\times \ell_0}, (x_1(a))^{\times \ell_1}, \dots, (x_N(a) = 1/q)^{\times \ell_N}\big),$$

where the weights $x_j(a)$ are determined by the convergents to the mirror continued fraction. Because $\mathbf{m} = q\mathbf{w}(a)$, we have

$$\mathbf{m} \cdot \mathbf{w}(z) = q\mathbf{w}(a) \cdot \mathbf{w}(z) = \sum_{j=1}^{N} q\ell_j \, x_j(a) \, x_j(z) = qz, \qquad \text{if } z \in (a-\varepsilon, a),$$

where the last equality holds for odd N by McDuff and Schlenk [17, Cor.2.2.7 (ii)]. This establishes the claimed formula when $z < a$. The formula for $z > a$ now follows by McDuff and Schlenk [17, Prop.2.3.2], since the assumption that $\mathbf{m} = q\mathbf{w}$ implies that the smallest entry in $\mathbf{m}$ is 1.

This proves the first claim. The second holds because the formula for $\mu_{\mathbf{E},b}(z)$ can change only at a point a' with $\ell(a') \le \ell(\mathbf{m})$, while Lemma 14 (iii) shows that a is the unique point with this property in the maximal neighborhood of a on which $\mu_{\mathbf{E},b}$ is nontrivial.

Remark 17

(i) When N is odd, the first formula in (2.1.12) holds for all z that have a continued fraction expansion of the form (2.1.13). Hence we may take our value $z_1 = [\ell_0; \ell_1, \dots, \ell_N + 1] < a$. (Note that, somewhat counterintuitively, when you increase an odd-placed coefficient of the continued fraction the number it represents decreases; see Remark 76.) Similarly, the second formula holds for all $z < z_2$, where we have the value $z_2 = [\ell_0; \ell_1, \dots, \ell_N - 1] > a$, unless $\ell_N = 2$, in which case we have $z_2 = [\ell_0; \ell_1, \dots, \ell_{N-1} + 1]$.

(ii) In the case of a ball (i.e. H_0), the paper [17] considered exceptional classes as above but with $m = 0$. In this case, every exceptional class that has a center a is perfect, and hence (as in Lemma 15) is nontrivial at a. Moreover, the obstruction function $\mu_{\mathbf{E},0}$ is live throughout the interval on which it is nontrivial. It also turns out that $a < \tau^4 = \mathrm{acc}(0)$, and that the capacity function $c_{H_0}(z)$ for $z < \tau^4$ is completely determined by these perfect classes.

(iii) The claims in (ii) no longer hold when $b > 0$. Lemma 15 (iii) shows that every exceptional class $\mathbf{E}$ with center a is live at a for some b, but this b may be far from the most relevant value of b, namely $\mathrm{acc}^{-1}(a)$. (Here $b \mapsto \mathrm{acc}(b)$ is as in (1.1.3) and we take the appropriate branch of its inverse as in (2.3.2).) However, as we will see in Sect. 2.3, if $\mathbf{E}$ is a perfect blocking class with center a then it does share some of the characteristics of perfect classes for the ball. In particular, we show in Proposition 42 that $\mu_{\mathbf{E},b}$ is live at a for all $b \in J_{\mathbf{E}}$, and hence in particular at $\mathrm{acc}^{-1}(b)$. ◊

The following remarks about the structure of a general class $\mathbf{E}$ that gives a nontrivial $\mu_{\mathbf{E},b}$ can be disregarded on a first reading. They will be useful when establishing the existence of staircases; see Example 70. They are slight generalizations of results in [17, §2]. We always assume that $\mathbf{m} = (m_1, \dots, m_n)$ is **ordered**, i.e. $m_1 \ge m_2 \ge \cdots \ge m_n$, since, because weight decompositions are ordered, the obstruction given by an ordered $\mathbf{m}$ is at least as large as that given by the any unordered tuple. If $a = [\ell_0; \ell_1, \dots, \ell_N]$ has weight decomposition

$\mathbf{w}(a) = (1^{\times \ell_0}, w_1^{\times \ell_1}, \ldots, w_N^{\times \ell_N})$, we call each group $w_i^{\times \ell_i}$ of equal terms a **block** of $\mathbf{w}(a)$.

Lemma 18 *Let $\mathbf{E} = (d, m; \mathbf{m})$ be a Diophantine class such that the function $\mu_{\mathbf{E},b}$ is nontrivial on a maximal interval I with break point $a = [\ell_0; \ell_1, \ldots, \ell_N]$. Suppose further that $\mathbf{m} = (m_1, \ldots, m_n)$ is ordered and that $z_0 \in I$. Then:*

(i) *there is a constant $\alpha(z_0, b) > 0$ and a vector ε of length $\ell(z_0)$ such that*

$$\mathbf{m} = \alpha(z_0, b)\mathbf{w}(z_0) + \varepsilon, \qquad \|\varepsilon\|^2 = \varepsilon \cdot \varepsilon < 1, \tag{2.1.14}$$

where the constant $\alpha(z_0, b)$ is such that the obstruction from $\alpha(z_0, b)\mathbf{w}(z_0)$ is the volume constraint $V_b(z_0)$, that is,

$$\alpha(z_0, b) = \frac{d - mb}{\sqrt{z_0(1 - b^2)}}. \tag{2.1.15}$$

(ii) *The vector $\mathbf{m}$ is constant on all blocks of $\mathbf{w}(z_0)$ of length two or more, except perhaps for one block on which all the entries but one are the same. In the latter case, the entries on this block differ by* 1.

Proof Since $\mu_{\mathbf{E},b}$ is nontrivial at z_0, it follows from (2.1.14), (2.1.15) that we must have

$$\mathbf{w}(z_0) \cdot \varepsilon > 0. \tag{2.1.16}$$

Indeed, because $\mathbf{w}(z_0) \cdot \mathbf{w}(z_0) = z_0$, the choice of $\alpha(z_0, b)$ implies that

$$\begin{aligned} V_b(z_0) &< \mu_{\mathbf{E},b}(z_0) \\ &= \frac{\alpha(z_0, b)\mathbf{w}(z_0) \cdot \mathbf{w}(z_0) + \varepsilon \cdot \mathbf{w}(z_0)}{d - mb} = V_b(z_0) + \frac{\varepsilon \cdot \mathbf{w}(z_0)}{d - mb}. \end{aligned}$$

Now notice that because $\mathbf{E} \cdot \mathbf{E} = -1$ and $\mathbf{w}(z_0) \cdot \mathbf{w}(z_0) = z_0$ by (2.1.5), we have

$$\begin{aligned} d^2 - m^2 + 1 &= \mathbf{m} \cdot \mathbf{m} \\ &= (\alpha(z_0, b)\mathbf{w}(z_0) + \varepsilon) \cdot (\alpha(z_0, b)\mathbf{w}(z_0) + \varepsilon) \\ &= \alpha^2(z_0, b)\, z_0 + 2\alpha(z_0, b)\, \varepsilon \cdot \mathbf{w}(z_0) + \varepsilon \cdot \varepsilon \\ &= \frac{(d - mb)^2}{1 - b^2} + 2\alpha(z_0, b)\, \varepsilon \cdot \mathbf{w}(z_0) + \varepsilon \cdot \varepsilon \quad \text{by (2.1.15)} \\ &= d^2 - m^2 + \frac{(m - bd)^2}{1 - b^2} + 2\alpha(z_0, b)\, \varepsilon \cdot \mathbf{w}(z_0) + \varepsilon \cdot \varepsilon. \end{aligned}$$

Therefore, we must have

$$\frac{(m - bd)^2}{1 - b^2} + 2\alpha(z_0, b)\, \varepsilon \cdot \mathbf{w}(z_0) + \varepsilon \cdot \varepsilon = 1. \tag{2.1.17}$$

Now observe that $\alpha(z_0, b) > 0$ by (2.1.15), and $\varepsilon \cdot \mathbf{w}(z_0) > 0$ by (2.1.16). Hence we have $\varepsilon \cdot \varepsilon < 1$ as claimed in (i).

To prove (ii), consider a block of $\mathbf{w}(z_0)$ of length $s \geq 2$, and suppose that the corresponding entries of $\mathbf{m}$ are the integers $m_1, \ldots, m_s$. If λ is the size of the corresponding entries in $\alpha(z_0, b)\mathbf{w}(z_0)$ we have

$$\sum_{i=1}^{s}(m_i - \lambda)^2 \leq \|\varepsilon\|^2 < 1.$$

In particular, $|m_i - \lambda| < 1$ for all i, which implies that the integers m_i differ by at most one on this block. Write these entries as $m^{\times r}$, $(m-1)^{\times(s-r)}$, and set $\mu := m - \lambda$. Then

$$\sum_{i=1}^{s}(m_i - \lambda)^2 = r\mu^2 + (s-r)(1-\mu)^2 = R\mu^2 - 2S\mu + T$$

$$\text{where } R = s,\ S = (s-r),\ T = s - r$$

and hence has minimum value $T - S^2/R$ taken at $\mu = S/R$. It remains to check that

$$T - \frac{S^2}{R} = \frac{(s-r)r}{s} < 1$$

only if $r = 0, 1, s-1, s$.

Remark 19 (Relation to ECH Obstructions)

(i) As explained in Sect. 4 below, there is an alternative approach to these calculations that use the obstructions coming from embedded contact homology (ECH). Thus, the capacity function $z \mapsto c_{H_b}(z)$ may also be described as the supremum of a ratio of ECH lattice counts as in (4.0.1). If a class $\mathbf{E} := \mathbf{E}(d, m; q\mathbf{w}(p/q))$ with center p/q is live at p/q for some value of b, then $c_{H_b}(z) = \mu_{\mathbf{E},b}(z)$ for $z \approx p/q$, and it follows from Lemma 92 that $c_{H_b}(z)$ must equal a ratio of counts for a particular lattice path. Thus, there is some interval around p/q on which the function $z \mapsto \mu_{\mathbf{E},b}(z)$ agrees with the obstruction given by the kth ECH capacity, where, by (4.2.3),

$$k = \tfrac{1}{2}\big(d(d+3) - m(m+1)\big). \tag{2.1.18}$$

However, as is shown by Fig. 9, for other values of z these two obstruction functions may be quite different. Notice here that $\mathbf{E}' = (73, 20; 29\mathbf{w}(170/29))$ is an exceptional class with center at $170/29 \approx 5.682$. Further, $\big(d(d+3) - m(m+1)\big)/2 = 2564$.

(ii) If we fix b, then we can often figure out from the graph of $z \mapsto c_{H_b}(z)$ exactly which exceptional classes are giving the obstruction. For example, in Fig. 8, we see that when $b = 3/10$, the capacity function c_{H_b} has an outer corner at $a = 23/4$. There is an exceptional class $\mathbf{E}$ with center $23/4$, namely $\mathbf{E} = \big(10, 3; 4\mathbf{w}(23/4)\big)$, such that $\mu_{\mathbf{E},3/10}$ is live near $23/4$ by Proposition 21 below. Hence we see that $c_{H_b} = \mu_{\mathbf{E},3/10}$ is given near a by the 59th ECH capacity. As another example, we found that the 125th capacity has a corner at about $35/6$ and from this it was not too hard to figure out the corresponding exceptional class. Indeed, with $p/q = 35/6$ we simply need to find suitable d, m satisfying $(d-m)(d+m) = pq-1$, which we can do in this case by factoring $pq-1 = 209$. In this way, we are led to the class $\mathbf{E} = \big(15, 4; 6\mathbf{w}(35/6)\big)$. This was one of the first ways we used to identify relevant exceptional classes. Once the first few such classes were found, we found others by numerical experimentation. One method, using Mathematica, is explained in Sect. 5.4. ◊

Remark 20 The computer calculations described in Sect. 5 take into account the ECH obstructions for $k \le 25{,}000$. The obstruction $\mu_{\mathbf{E},b}$ given by a class $\mathbf{E} = (d, m; \mathbf{m})$ with break point a corresponds to the kth ECH obstruction near $z = a$, where $k = \big(d(d+3) - m(m+1)\big)/2$. If we restrict the range of b, then we can estimate the degree d of $\mathbf{E}$ as a function of k since we know from Lemma 15 that $|bd-m| < 1$. For example, if $b \le 1/3$, then all relevant classes have $m < bd+1 < d/3+1$. Combining that with the inequality $d(d+3) - m(m+1) < 50{,}000$, we find that $\frac{8}{9}d^2 < 50{,}000$, which roughly corresponds to $d \le 237$. The only proviso here is that the computer works with a fixed value for b, which is often a rational approximation to the irrational value b_∞ of interest. However, it follows from results such as Proposition 21 below that one can quantify the b-interval on which a class is visible. We will rigorously prove all our results, rather than relying on computer programs to tell us that certain classes cannot exist, since the relevant numbers grow quickly, escaping the range that is easily accessible to our computers. ◊

2.2 Characterizing Staircases

The section explains how to recognize families of obstructions that give staircases for the Hirzebruch domains $H_b, 0 < b < 1$. We begin with a useful criterion for a class to be live. It is an adaptation of [19, Lem.4.1].

Proposition 21 *Suppose that* $\mathbf{E} = (d, m; \mathbf{m})$ *is b-perfect with center a. Then:*

(i) *If b satisfies*

$$\frac{m^2-1}{dm} \le b \le \frac{1+m(d-m)}{1+d(d-m)}, \tag{2.2.1}$$

the obstruction $\mu_{\mathbf{E},b}$ *is live at a; that is,* $c_b(a) = \mu_{\mathbf{E},b}(a) > V_b(a)$.

(ii) *If $b < r/s$ satisfies*

$$\frac{m^2-1}{dm} \le b \le \frac{s+m(rd-sm)}{r+d(rd-sm)}, \tag{2.2.2}$$

then $\mu_{\mathbf{E},b}(a) \ge \mu_{\mathbf{E}',b}(a)$ for every exceptional divisor $\mathbf{E}' = (d', m', \mathbf{m}')$ with $m'/d' \le r/s$.

(iii) *If $\frac{m}{d} > \frac{r}{s}(1+\frac{1}{d^2})$ and $b > r/s$ satisfies*

$$\frac{m(sm-rd)-s}{d(sm-rd)-r} \le b \le \frac{m}{d} \tag{2.2.3}$$

then $\mu_{\mathbf{E},b}(a) \ge \mu_{\mathbf{E}',b}(a)$ for every exceptional divisor $\mathbf{E}' = (d', m', \mathbf{m}')$ with $m'/d' \ge r/s$.

Proof Suppose that b satisfies (2.2.1), and let $\mathbf{E}' = (d', m'; \mathbf{m}')$ be any exceptional class. Then because $\mathbf{E} \cdot \mathbf{E}' \ge 0$ we have

$$dd' - mm' - \mathbf{m} \cdot \mathbf{m}' \ge 0. \tag{2.2.4}$$

Therefore if $b \le m/d$ we have

$$\begin{aligned} \mu_{\mathbf{E}',b}(a) = \frac{\mathbf{m}' \cdot \mathbf{w}(a)}{d' - m'b} &\le \frac{\mathbf{m}' \cdot \mathbf{m}}{q(d' - m'\frac{m}{d})} \quad \text{since } b \le \frac{m}{d} \\ &\le \frac{d(dd' - mm')}{q(dd' - mm')} = \frac{d}{q}. \end{aligned}$$

On the other hand, because $\mathbf{w}(a) \cdot \mathbf{w}(a) = a = p/q$ by (2.1.5),

$$\mu_{\mathbf{E},b}(a) = \frac{\mathbf{m} \cdot \mathbf{w}(a)}{d - mb} = \frac{p}{d - mb} \ge \frac{d}{q} \quad \text{if} \quad pq \ge d^2 - dmb.$$

Since $pq = d^2 - m^2 + 1$ by (2.1.2), this will hold if

$$dmb \ge d^2 - pq = m^2 - 1,$$

i.e. $b \ge \left(m^2-1\right)/(dm)$. This shows that when $\left(m^2-1\right)/(dm) \le b \le m/d$ the obstruction $\mu_{\mathbf{E},b}(a)$ is at least as large as any other that is defined by an exceptional class. Finally, we have assumed that $\mathbf{E}$ is b-perfect, so $|bd - m| < \sqrt{1-b^2}$. Hence $\mu_{\mathbf{E},b}$ is live at a by Lemma 15 (iii).

Similarly, if $b > m/d$, we have

$$\mu_{\mathbf{E}',b}(a) = \frac{\mathbf{m}' \cdot \mathbf{w}(a)}{d' - m'b} \le \frac{dd' - mm'}{q(d' - m'b)} \quad \text{by (2.2.4)}$$

$$\leq \frac{d - m\frac{m'}{d'}}{q(1 - b\frac{m'}{d'})} \quad \leq \quad \frac{d-m}{q(1-b)}$$

where the last inequality uses the fact that because $m/d < b < 1$ the function $x \mapsto \left(1 - \frac{m}{d}x\right) / (1 - bx)$ increases on the interval $(0, 1]$. Hence, as above, the inequality $\mu_{\mathbf{E},b}(a) \geq \mu_{\mathbf{E}',b}(a)$ holds whenever

$$\frac{p}{d - mb} \geq \frac{d-m}{q(1-b)}.$$

Substituting $pq = d^2 - m^2 + 1$, one readily checks that this is equivalent to the upper bound in (2.2.1). Since we again have assumed $|bd - m| < \sqrt{1 - b^2}$, this completes the proof of (i).

Claim (ii) is proved in exactly the same way except that if $b > m/d$ we use the fact that $m'/d' \leq r/s$ to conclude that $\mu_{\mathbf{E}',b}(a) \leq \frac{sd-rm}{q(s-rb)}$. Hence it suffices that

$$\frac{pq}{d - mb} = \frac{d^2 - m^2 + 1}{d - mb} \geq \frac{sd - rm}{s - rb},$$

which is equivalent to the condition $b \leq \frac{s+m(rd-sm)}{r+d(rd-sm)}$.

Finally, to prove (iii) we consider the function $x \mapsto \frac{d-mx}{1-bx}$ for $r/s < b < m/d < 1$ and $r/s \leq x < 1$. This is a decreasing function. Hence, taking $x = m'/d'$ we have

$$\begin{aligned} \mu_{\mathbf{E}',b}(a) &\leq \frac{d - m\frac{m'}{d'}}{q(1 - b\frac{m'}{d'})} \quad \text{by (2.2.4)} \\ &\leq \frac{sd - rm}{q(s - rb)}. \end{aligned}$$

Hence it suffices that

$$\frac{pq}{d - mb} = \frac{d^2 - m^2 + 1}{d - mb} \geq \frac{sd - rm}{s - rb},$$

which is equivalent to the condition $b \geq \frac{sm^2-rmd-s}{smd-rd^2-r}$, since we assumed that $sm > r(d + 1/d)$.

Example 22 Consider the exceptional class $\mathbf{E} = (2, 0; 1^{\times 5})$. It has center $a = 5$ and is b-perfect for b such that $|2b| < \sqrt{1 - b^2}$, i.e. for $b \in [0, 1/\sqrt{5})$. By Lemma 16, $\mu_{\mathbf{E},b}(z) = 5/2$ for all $z \geq 5$. Notice that

$$5/2 = \mu_{\mathbf{E},b}(6) = \sqrt{\frac{6}{1 - b^2}} = V_b(6), \qquad \text{when } b = 1/5.$$

Further b satisfies (2.2.1) precisely if $b < 1/5$ so that $c_{H_b}(5) = 5/2$ for $b \leq 1/5$, and $c_{H_{1/5}}(6) \geq 5/2$, with equality unless there is another obstruction for $b = 1/5$ at $z = 6$. But this would have to have break point at 6 since $\ell(6)$ is so small, and one can easily check that no such obstruction exists. Thus $H_{1/5}$ is unobstructed. Also there is no ascending staircase at $b = 1/5$. For further discussion, see Example 34 and Sect. 4.3. ◊

Remark 23 The proof of Lemma 15 only uses the Diophantine conditions (2.1.2) satisfied by $\mathbf{E}$, and hence holds for **fake classes** $\mathbf{E} = (d, m; \mathbf{m})$ that satisfy these conditions but do not reduce correctly under Cremona moves, and so are not exceptional classes. It turns out that there are many such fake classes: see Example 28 below. By contrast, the proof of Proposition 21 is based on the fact that different classes in $\mathcal{E}$ have nonnegative intersection, which is false for fake classes. In fact, it seems likely that the obstructions given by fake classes are never live, though we do not attempt to give a proof here. In any case, it follows from (1.1.3) that if there were a live fake class for some values of a, b there would have to be a class in $\mathcal{E}$ giving the same obstruction at a, b.

By contrast, results such as Lemma 38 about blocking classes do not require the class in question to be exceptional, and hence we phrase them in terms of Diophantine classes. It turns out that many quasi-perfect blocking classes are in fact perfect: see Proposition 42. ◊

We now give a formal definition of staircase, to clarify the language used below. In Sect. 1.1, we said that H_b has an infinite staircase precisely when the graph of c_{H_b} has infinitely many nonsmooth points that lie above the volume curve. We saw in Lemma 15 (ii) that c_{H_b} has finitely many nonsmooth points on any interval on which it is bounded away from the volume. It follows that it can have only countably many nonsmooth points that lie above the volume curve. Moreover, by Cristofaro-Gardiner et al. [7, Thm. 1.11], the corresponding set of z-coordinates has a unique limit point $a = \mathrm{acc}(b)$. Hence we can always arrange these z-coordinates into a convergent sequence a_k. By Lemmas 14 (iii) and 15 (ii), each such point a_k is the break point of some obstructive class $\mathbf{E}_k$ such that $\mu_{\mathbf{E}_k, b}$ is live at a_k in the sense of Definition 10. Hence H_b has an infinite staircase precisely when there is a sequence of such obstructive classes.

Definition 24 A **staircase** $\mathcal{S}$ for H_b is a sequence of distinct Diophantine classes $\big(\mathbf{E}_k\big)_{k\geq 0}$ such that each $\mu_{\mathbf{E}_k, b}$ is live at its break point a_k for some sequence of distinct points a_k that converge to $a_\infty := \mathrm{acc}(b)$. The points a_k are called the **steps** or **exterior corners** of $\mathcal{S}$, and $\mathcal{S}$ is said to be **ascending** (respectively **descending**) if the sequence of steps a_k is increasing (resp. decreasing). We say that an ascending (resp. descending) staircase is **complete** if there is a one-sided neighborhood $(a_\infty - \varepsilon, a_\infty]$ (resp. $[a_\infty, a_\infty + \varepsilon)$) of a_∞ such that there is no class $\mathbf{E}'$ other than the $\mathbf{E}_k$ with break point a' in this neighborhood and such that $\mu_{\mathbf{E}', b}$ is live at a'. Finally we say that a staircase $\mathcal{S}$ is **perfect** if all its classes are perfect.

Remark 25

(i) It follows from the above remarks that H_b has an infinite staircase if and only if there is a sequence of classes $\mathbf{E}_k$ with the above properties. However notice that even if $\left(\mathbf{E}_k\right)$ is complete, with ascending steps a_k, then the capacity function c_{H_b} may not be determined on any interval $(a_\infty - \varepsilon, a_\infty]$ by the classes $\mathbf{E}_k$, since there might be 'big' obstructions as in Example 32.

(ii) The new staircases that we describe in Sect. 3.1 consist of two interwoven sets of classes. According to our definition, each of these sets forms a staircase, as does their union. It is likely (though we do not prove that here) that their union is complete in the sense defined above. ◊

We are now in a position to establish our first **staircase recognition** criterion. We will show in the proof that condition (ii) below cannot be satisfied if $b_\infty = 0$, though we do not assume that from the beginning.

Proposition 26 *Suppose that* $\left(\mathbf{E}_k = \left(d_k, m_k; q_k \mathbf{w}(p_k/q_k)\right)\right)_{k\geq 1}$ *is a sequence of perfect classes with* $m_k \neq 0$ *such that*

(i) $m_k/d_k \to b_\infty \in [0, 1)$ *and* $p_k/q_k \to a_\infty$*;*

(ii) $\frac{m_k^2-1}{d_k m_k} < b_\infty < \frac{1+m_k(d_k-m_k)}{1+d_k(d_k-m_k)}$ *for all* $k \geq k_0$.

Then the classes $\left(\mathbf{E}_k\right)_{k\geq k_0}$ *provide a staircase for* H_{b_∞}*. Moreover,* $a_\infty = \mathrm{acc}(b_\infty)$.

Proof Condition (ii) implies that for sufficiently large k there is a constant c depending on $1 - b_\infty > 0$ such that

$$\frac{m_k}{d_k} - \frac{1}{m_k d_k} < b_\infty < \frac{m_k}{d_k + \frac{1}{d_k - m_k}} + \frac{1}{d_k(d_k - m_k)} < \frac{m_k}{d_k} + \frac{1}{c d_k^2}.$$

Rearranging, this gives us $1/m_k > m_k - b_\infty d_k > -1/(cd_k)$. Now we must have $d_k \to \infty$, since there are only finitely many exceptional classes of bounded degree. Further, since $m_k \neq 0$, we also have $m_k \to \infty$, since otherwise $b_\infty = 0$ and we have $m_k < 1/m_k$, which is impossible since $m_k \geq 0$ is an integer. This implies $|m_k - b_\infty d_k| \to 0$. We obtain in particular the bound $|m_k - d_k b_\infty| < \sqrt{1 - b_\infty^2}$ for large k, so that the classes $\mathbf{E}_k$ are b_∞-perfect for large k. Proposition 21 (i) now implies that E_k is live at b_∞ for sufficiently large k. Thus, we have a sequence of live obstructions $\mu_{\mathbf{E}_k, b_\infty}$ whose centers converge to a_∞; in other words, a staircase. Finally, we may use [7, Thm.1.11] to deduce that $a_\infty = \mathrm{acc}(b_\infty)$, as desired.

Since condition (ii) above is rather strong, we now establish a weaker version. We begin by giving conditions that guarantee that H_{b_∞} is unobstructed.

Lemma 27 *Suppose that* $\left(\mathbf{E}_k = \left(d_k, m_k; q_k \mathbf{w}(p_k/q_k)\right)\right)_{k\geq 1}$ *is a sequence of distinct quasi-perfect classes such that* $m_k/d_k \to b_\infty \in (0, 1)$*, that* $|d_k b_\infty - m_k| < 1$

for all k, and that $a_k := p_k/q_k \to a_\infty$. Then

(i) $a_\infty = \mathrm{acc}(b_\infty)$.
(ii) *If the $\mathbf{E}_k$ are perfect, then H_{b_∞} is unobstructed.*

Proof By Lemma 15 (iii) the obstructions $\mu_{\mathbf{E}_k, b_\infty}(a_k)$ are nontrivial for all k. The proof of [7, Thm.1.11] then shows that $a_\infty = \mathrm{acc}(b_\infty)$. Indeed, a careful reading of the [7] argument shows that it never uses the hypothesis that the obstructions $\mu_{\mathbf{E}_k, b_\infty}$ are live at a_k, instead using extensions of the arguments in Lemma 18 above. This proves (i).

If we now assume that the $\mathbf{E}_k$ are perfect and set $b_k := m_k/d_k$, then Proposition 21 implies that $\mu_{\mathbf{E}_k, b_k}$ is live at a_k. Thus, by Lemma 15 (i), we have

$$V_{b_k}(a_k) < c_{H_{b_k}}(a_k) = \mu_{\mathbf{E}_k, b_k}(a_k) < V_{b_k}(a_k)\sqrt{1 + 1/(d_k^2 - m_k^2)}.$$

But $d_k^2 - m_k^2 = p_k q_k - 1 \to \infty$. Hence by the continuity of the function $(z, b) \mapsto c_{H_b}(z)$, we must have $c_{H_{b_\infty}}(a_\infty) = V_{b_\infty}(a_\infty)$. In other words, H_{b_∞} is unobstructed.

Example 28 The following examples illustrate the need for some of these conditions.

(i) For each $k \geq 1$ there is an exceptional class $\mathbf{E}_k = (5k, k+1; \mathbf{m}_k)$ with center $a_k = 6 - 1/(2k) = (12k-1)\,/\,(2k)$, where

$$\mathbf{m}_k = \big((2k)^{\times 5}, 2k-1, 1^{\times(2k-1)}\big) = 2k\mathbf{w}\Big(\frac{12k-1}{2k}\Big).$$

These are exceptional classes because they evidently satisfy (2.1.2) and it is not hard to prove by induction that they reduce correctly. Since $m_k/b_k \to 1/5$ and $a_k \to 6 = \mathrm{acc}(1/5)$, the first two conditions in Lemma 27 hold. However, condition (ii) in Proposition 26 does not hold: indeed $\mathbf{E}_k$ is not even $1/5$-perfect. Nevertheless, the existence of these classes implies that $H_{1/5}$ is unobstructed. We prove that it does not admit a staircase in Sect. 4.3.

(ii) There is an exceptional class $\mathbf{E}' = \big(73, 20; 29\mathbf{w}(170/29)\big)$ with center $a = 170/29$. Here we have $29\mathbf{w}(170/29) = (29^{\times 5}, 25, 4^{\times 6}, 1^{\times 4})$, and it not hard to check that $\mathbf{E}'$ does reduce correctly under Cremona moves. By Proposition 21 this class is live for $b \approx 20/73 = 0.27397\ldots$ while[4] $\mathrm{acc}_L^{-1}(170/29) \approx .275425\ldots$ is almost the same. Therefore one can expect that this class is relevant to the study of staircases. Indeed, it turns out that this class is the entry $\mathbf{E}^E_{0,0}$ in the staircase $\mathcal{S}^E_0$ in Theorem 2.

(iii) There is another related[5] class $\mathbf{E}'' = \big(48, 14; 19\mathbf{w}(111/19)\big)$ with center $111/19$, but one can check that this one is fake, i.e. quasi-perfect but not

[4] The function acc_L^{-1} is a branch of the inverse to $b \mapsto \mathrm{acc}(b)$; see (2.3.2).

[5] In fact there is a family of such classes, whose centers $a_i = 6 - \frac{i}{(i+1)(i+2)-1}$, $i \geq 3$, converge to 6; for $i = 2n+4$ these are the exceptional classes $\mathbf{E}_{n,0}$ in $\mathcal{S}^E_n$, but they seem to be fake for odd i.

perfect: see Remark 23. As is shown by Fig. 10, when b is chosen so that $\mathrm{acc}(b) = 111/19$ (i.e. $b \approx 0.296654$) the corresponding function $z \mapsto \mu_{\mathbf{E}'',b}(z)$ agrees with the kth ECH capacity where k is given by (2.1.18), and lies between the volume function and the capacity function c_{H_b}. ◊

Although some of the staircases that we identify do satisfy the criteria in Proposition 26, some of them only satisfy a modification of condition (ii), such as that in (2.2.2). Thus we have a sequence of classes $\mathbf{E}_k$ that are nontrivial, though perhaps not live, at b_∞. In order for such a sequence to form a staircase, we have to rule out the existence of an overshadowing class $\mathbf{E}'$, defined as follows.

Definition 29 Let $I = (y_0, y_1)$ be a maximal open interval on which the obstruction $\mu_{\mathbf{E}',b}$ is nontrivial. Then we say that

- $\mathbf{E}'$ is a **right-overshadowing class at** y_0 if
  - $\mathrm{acc}(b) = y_0$, and
  - $\mu_{\mathbf{E}',b}$ is live on some nonempty subset $(y_0, y_0 + \varepsilon) \subset I$;
- $\mathbf{E}'$ is a **left-overshadowing class at** y_1 if
  - $\mathrm{acc}(b) = y_1$, and
  - $\mu_{\mathbf{E}',b}$ is live on some nonempty subset $(y_1 - \varepsilon, y_1) \subset I$.

If $\mathcal{S}$ is a sequence of quasi-perfect classes whose centers a_k converge to a_∞, we say that $\mathbf{E}'$ **overshadows** $\mathcal{S}$ (or **is an overshadowing class**, if $\mathcal{S}$ is understood from the context) if either

- the a_k descend and $\mathbf{E}'$ is right-overshadowing (i.e. overshadows to the right of a_∞); or
- the a_k ascend and $\mathbf{E}'$ is left-overshadowing at a_∞ (i.e. overshadows to the left of a_∞).

Here is a refined version of Proposition 26. Notice that setting $r/s = 1$ in (iii) below gives the same as condition (ii) in Proposition 26.

Proposition 30 *Let* $\left(\mathbf{E}_k = \left(d_k, m_k; q_k \mathbf{w}(p_k/q_k)\right)\right)_{k\ge 0}$ *be a sequence of perfect classes that satisfies the following conditions for some* $r/s \in (0, 1]$*:*

(i) $m_k/d_k \to b_\infty \in [0, r/s)$ *and the centers* p_k/q_k *are a monotonic sequence with limit* a_∞*;*

(ii) $\frac{m_k^2-1}{d_k m_k} \le b_\infty \le \frac{s+m_k(rd_k-sm_k)}{r+d_k(rd_k-sm_k)}$ *for all* $k \ge k_0$.

(iii) *There is no overshadowing class* $\mathbf{E}'$ *at* a_∞ *of degree* $d' < s/(r - sb_\infty)$ *and with* $m'/d' > r/s$

Then there is k_0 *such that the classes* $\left(\mathbf{E}_k\right)_{k\ge k_0}$ *are live at* b_∞ *and hence form a staircase for* b_∞.

Proof We know from Lemma 27 that H_{b_∞} is unobstructed and $\mathrm{acc}(b_\infty) = a_\infty$. Further, because condition (ii) holds, it follows from Proposition 21 (ii) that

$\mu_{\mathbf{E}_k,b_\infty}(p_k/q_k) > \mu_{\mathbf{E}',b_\infty}(p_k/q_k)$ for all exceptional classes $\mathbf{E}' = (d', m', \mathbf{m}')$ with $m'/d' \le r/s$. Hence if $\mathbf{E}_k$ is not live at b_∞ there is a class $\mathbf{E}' = (d', m'; \mathbf{m}')$ with $m'/d' > r/s$ such that $\mu_{\mathbf{E}',b_\infty}(a_\infty) > \mu_{\mathbf{E}_k,b_\infty}(a_\infty)$. But then Lemma 15 implies that $|b_\infty d' - m'| < 1$ and hence $|m'/d' - b_\infty| < 1/d'$. Further, $b_\infty < r/s$ by (i), so that $m'/d' - b_\infty > r/s - b_\infty$. Therefore we must have $1/d' > r/s - b_\infty > 0$, or, equivalently, $d' < s/(r - sb_\infty)$. In particular, there are a finite number of such classes $\mathbf{E}'$. It follows that either all but finitely many $\mathbf{E}_k$ are live at b_∞ or there is a single class $\mathbf{E}'$ with $d' < s/(r - sb_\infty)$ whose obstruction for $b = b_\infty$ is live at all but finitely many of the points a_k. But in the first case the $(\mathbf{E}_k)_{k\ge k_0}$ form a staircase, while in the second, since H_{b_∞} is unobstructed, $\mathbf{E}'$ satisfies the conditions to be an overshadowing class. Since the latter is ruled out by condition (iv), the classes $(\mathbf{E}_k)_{k\ge k_0}$ must form a (possibly incomplete) staircase, as claimed.

Remark 31

(i) There is an analogous result for classes $\mathbf{E}_k$ such that the ratios $b_k = m_k/d_k$ decrease with limit $b_\infty > r/s$ in which the inequalities in (iii) above are replaced by (2.2.3) from Proposition 21 (iii):

$$\frac{m_k(sm_k - rd_k) - s}{d_k(sm_k - rd_k) - r} \le b \le \frac{m_k}{d_k}, \qquad d' < \frac{s}{sb_\infty - r}.$$

Notice that if $m_k/d_k > r/s$ for all k then, since the $d_k \to \infty$, the inequality $m_k/d_k > r(1 + 1/d_k^2)/s$ required by Proposition 21 (iii) does hold for large k. Further, as above, any potentially overshadowing class would have parameters d', m' where $m'/d' < r/s$, and could only be live at b_∞ if $b_\infty - r/s < b_\infty - m'/d' < 1/d'$, which gives the bound $d' < s/(sb_\infty - r)$.

(ii) Although in principle there can be overshadowing classes for both ascending and descending sequences of obstructions, in practice we only seem to have to worry about them in the descending case. See Lemmas 73 and 74, for example.

◊

Example 32 The exceptional class $\mathbf{E}' := \left(3, 1; 2, 1^{\times 5}\right)$ is a potentially awkward class that is a rearranged version of the large nontrivial blocking class $\mathbf{E}_0 := \left(3, 2; 1^{\times 6}\right)$ in Theorem 1. For many values of z it corresponds to the 8th ECH capacity, and, as explained in Sect. 5.5, is visible on many of the figures. Though $\mathbf{E}'$ has no center and so is not perfect, it has break point at $a = 6$, and gives the following obstruction

$$\mu_{\mathbf{E}',b}(z) = \begin{cases} \dfrac{1+z}{3-b}, & 5 < z < 6, \\ \dfrac{7}{3-b}, & z \ge 6. \end{cases}$$

Note that $\mathbf{E}'$ is nontrivial at its break point $z = 6$ precisely when

$$\mu_{\mathbf{E}',b}(6) = \frac{7}{3-b} \;>\; V_b(6) = \sqrt{\frac{6}{1-b^2}} \qquad \text{i.e. if} \quad b \in \left(\frac{1}{5}, \frac{5}{11}\right).$$

Further, one can check from (1.1.2) that the range of b in which $\mathrm{acc}(b) < 6$ is also the interval $1/5 < b < 5/11$. Therefore, for b in this range we have

$$\mu_{\mathbf{E}',b}\big(\mathrm{acc}(b)\big) = \frac{1+\mathrm{acc}(b)}{3-b},$$

and this turns out to be precisely the same as the volume obstruction. Indeed, if we define $c(b) = \frac{(3-b)^2}{1-b^2} - 2$, we have

$$\begin{aligned} &\frac{1+\mathrm{acc}(b)}{3-b} = \sqrt{\frac{\mathrm{acc}(b)}{1-b^2}} \qquad (2.2.5)\\ &\quad\Longleftrightarrow\quad \big(1+\mathrm{acc}(b)\big)^2 = \mathrm{acc}(b)\frac{(3-b)^2}{1-b^2} = \mathrm{acc}(b)\big(c(b)+2\big),\\ &\quad\Longleftrightarrow\quad \mathrm{acc}(b)^2 - c(b)\mathrm{acc}(b) + 1 = 0 \end{aligned}$$

which holds by the definition of $\mathrm{acc}(b)$ in (1.1.2). Therefore, for $b \in (1/5, 5/11)$, this class satisfies at least some of the conditions to be overshadowing, and so potentially obstructs the existence of a descending staircase.

However, it does not overshadow the staircases $\mathcal{S}^E_{u,n}, n \geq 0$, defined in Theorem 58. These are descending with limit points at $(a_{n,\infty},\ b_{n,\infty}) \in (5,6) \times (1/5, 1/3)$; and we show in Example 77 and Lemma 78 that they satisfy condition (ii) in Proposition 30 (by way of Lemma 67 (ii)) with $r/s = 1/3$. Therefore, because $\mathbf{E}'$ has $m/d = 1/3$ (rather than $> 1/3$), it is not overshadowing. One can also see the initial peaks of the steps of $\mathcal{S}^E_{u,0}$ poking above the line $z \mapsto \mu_{\mathbf{E}',b}(z)$ in Fig. 7.

Remark 33

(i) The observation in (2.2.5) about the function $b \mapsto V_b(\mathrm{acc}(b))$ turns out to be an essential ingredient of our proof of the relation between staircases and blocking classes: see Lemma 44.

(ii) All of the new ascending (resp. descending) staircases that we have found accumulate at the initial (resp. final) point $\big(a_\infty, V(a_\infty)\big)$ of a "big" obstruction, namely, an obstruction that goes through $\big(a_\infty, V(a_\infty)\big)$ and for the given value of b obstructs above (resp. below) a_∞. Moreover, in many cases (though not in the case of the Fibonacci staircase at $b = 0$) this big obstruction is given by a perfect blocking class as described in Sect. 2.3. Note that this phenomenon is quite different from that of a (potentially) overshadowing class since the latter type of obstruction overhangs the staircase, and so is active on the staircase side of the accumulation point. ◊

Example 34 The degree 5 class $\mathbf{E}_5 := (5, 1; 2^{\times 6}, 1)$ also affects the capacity function for some ranges of b, z, and for $b \approx 1/5$ the obstruction $\mu_{\mathbf{E}_5,b}$ equals the obstruction from the 19th ECH capacity for $z \approx 7$, a behavior that is discussed in Sect. 4.2. One can check that the corresponding obstruction is

$$\mu_{\mathbf{E}_5,b}(z) = \frac{6+z}{5-b}, \quad 11/2 < z < 7$$

which goes through the accumulation point $(6, 5/2)$ when $b = 1/5$, and is tangent to the graph of $z \mapsto V_{1/5}(z)$ at that point. We show in Sect. 4.3 that $c_{H_{1/5}}(z) = \mu_{\mathbf{E}_5,b}(z)$ on some interval $[1/5, 1/5 + \varepsilon)$.

We claim that $b = 1/5$ is the unique value where the obstruction given by $\mathbf{E}_5$ might affect the existence of staircases since we have

$$\mu_{\mathbf{E}_5,b}(\mathrm{acc}(b)) < V_b(\mathrm{acc}(b)), \quad b \neq 1/5. \tag{2.2.6}$$

To justify this, recall that $V_b(\mathrm{acc}(b)) = \frac{1+\mathrm{acc}(b)}{3-b}$ by (2.2.5). Since $\mu_{\mathbf{E}_5,b}(z)$ is constant for $z \geq 7$ and the increasing function $z \mapsto \frac{1+z}{3-b}$ is greater than $\mu_{\mathbf{E}_5,b}(z)$ at $z = 7$, the class $\mathbf{E}_5$ can only affect staircases with $\mathrm{acc}(b) \leq 7$. Moreover, $\mathrm{acc}(b) \geq 3 + 2\sqrt{2} > 11/2$, and in this range we have

$$\mu_{\mathbf{E}_5,b}(z) = \frac{6+z}{5-b} =: f_1(z),$$

whose graph is a line of smaller slope than that of $\frac{1+z}{3-b} =: f_2(z)$. Thus (2.2.6) will hold if we show that for each $b \neq 1/5$, the two lines $y = f_1(z)$ and $y = f_2(z)$ intersect at a point z_b with $f_2(z_b) < V_b(\mathrm{acc}(b))$. But $z_b = (13 - 5b)/2$, and $f_2(z_b) = 5/2 < V_b(\mathrm{acc}(b))$ for all $b \neq 1/5$, since the function $b \mapsto V_b(\mathrm{acc}(b))$ has a minimum value of $5/2$, taken at $b = 1/5$; see Fig. 1. Thus (2.2.6) holds. ◊

Example 35 (The Staircase at $b = 1/3$) This staircase is somewhat different from the new staircases that we find in the current paper. For one thing, it consists of three interwoven sequences, while the new ones that we discuss here consist of two such sequences.[6] For another, it does not satisfy either the condition (2.2.1) or its variant (2.2.2). Indeed, both these conditions imply that $|b_\infty d_k - m_k| = O(d_k^{-1})$, while the sequences at $b = 1/3$ satisfy $|b_\infty d_k - m_k| = O(1)$. Further, the ratios m_k/d_k lie on both sides of $1/3$, while in the new staircases these ratios are monotonic. Finally, as in the case of the Fibonacci staircase in (3.0.1), the numerators and denominators of the center points $a_k := p_k/q_k$ of each of the three substaircases fit into a single sequence $g_0, g_1, g_2 \ldots$ where $a_k = g_k/g_{k-1}$. Thus there are three interwoven sequences

[6] See [7] for a discussion of such symmetries.

$$\mathbf{E}_{k,i} = \big(d_{k,i}, m_{k,i}; g_{k-1,i}\mathbf{w}(\frac{g_{k,i}}{g_{k-1,i}})\big), \quad i = 0, 1, 2$$

where for each i the numbers $g_{k,i}$ satisfy the recursion formula

$$g_{k+1,i} = 6g_{k,i} - g_{k-1,i}$$

with seeds (i.e. initial values)

$$g_{0,0} = 1, \quad g_{1,0} = 2, \quad g_{0,1} = 1, \quad g_{1,1} = 4, \quad g_{0,2} = 1, \quad g_{1,2} = 5.$$

Thus in all cases, the sequence of centers $g_{k.i}/g_{k-1.i}$ increases,[7] and can be represented by the continued fractions $[5; \{1, 4\}^{k+\varepsilon}, \mathrm{end}_i]$, with $\mathrm{end}_0 = 2$, $\mathrm{end}_1 = (1, 3)$ and $\mathrm{end}_2 = \emptyset$, and $\varepsilon = 0, \pm 1$ as appropriate. (Compare this with the staircase descriptions in Theorems 54, 56 and 58.) The numbers $d_{k,i}, m_{k,i}$ satisfy modified versions of this recursion, and are determined for each i and k by the requirement $3d_{k,i} - m_{k,i} = g_{k,i} + g_{k-1,i}$ and the following additional relation:

$$d_{k,0} = 3m_{k,0} - (-1)^k, \quad d_{k,1} = 3m_{k,1} + (-1)^k, \quad d_{k,2} = 3m_{k,2} - (-2)^k.$$

For example, the first sequence starts with the classes:

$$\mathbf{E}_{1,0} = \big(1, 0; \mathbf{w}(2)\big), \quad \mathbf{E}_{2,0} = \big(5, 2; 2\mathbf{w}(\frac{11}{2})\big), \quad \mathbf{E}_{3,0} = \big(28, 9; 11\mathbf{w}(\frac{64}{11})\big).$$

Thus the linear relations in this staircase are inhomogeneous, and hence different in nature from those in the new staircases, which, as we will see in Theorem 52, are homogeneous and arise because the latter are associated to blocking classes.

The methods developed in [2] and [7] involve identifying staircases by constructing tight packings at the inner corners using almost toric fibrations (ATFs). It might be rather difficult to use ATF methods for the staircases identified in this manuscript, since the capacity function $c_{H_{b_\infty}}$ may well not have well defined inner corners because of the presence of an obstruction that nearly overshadows the staircase.[8] This possibility is illustrated in Fig. 7 in the case of the descending staircase $\mathcal{S}^E_{u,0}$ in Theorems 2 and 58, which is almost overshadowed by the obstruction from $(3, 1; 2, 1^{\times 5})$. Thus, in this case one cannot find tight packings of the fixed target H_{b_∞} by a sequence of ellipsoids $E(1, z_k)$ where z_k increases to a_∞.

Nevertheless, one might be able to use tight packings to show that there is a decreasing sequence of unobstructed points $a'_k = \mathrm{acc}(b'_k)$ with $b'_k \to b_\infty$. This

[7] In fact, one can check that $g_{k+1,i}g_{k-1,i} = g^2_{k,i} + c_i$, where $c_i = 7$ for $i = 0, 1$ and $c_2 = 4$.

[8] By contrast, Usher conjectures in [19, Conj.4.23] that in the cases he considers, the capacity function is defined on (a_∞, a_1) by the union of the relevant staircase classes, which would imply that the capacity function has well delineated inner corners.

would also be enough to show that $\mathcal{S}^E_{u,0}$ is a descending staircase for H_{b_∞}. Further, as is shown in Figs. 2 and 3, the staircases in the other families $\mathcal{S}^U_{\bullet,\bullet}$, $\mathcal{S}^L_{\bullet,\bullet}$ discussed in Sect. 3.1 below may not be so nearly obscured, since the problematic class $\mathbf{E} = (3, 1; 2, 1^{\times 5})$ in Example 32 is not obstructive in the relevant region. ◊

2.3 Blocking Classes

Let $\mathbf{E}$ be a Diophantine class in the sense of Definition 12, and recall from Definition 10 that an obstruction $\mu_{\mathbf{E},b}$ is said to block b if

$$\mu_{\mathbf{E},b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b)).$$

Further, we call $\mathbf{E} = \big(d, m; \mathbf{m}\big)$ a **blocking class** if there is some b_0 such that the corresponding obstruction $\mu_{\mathbf{E},b_0}$ blocks b_0. We will show in Remark 39 that b_0 need not be m/d.

As preparation for discussing the properties of these classes, recall from Sect. 1 that the point $\mathrm{acc}(b)$ is the unique solution > 1 of the equation

$$z^2 - \left(\frac{(3-b)^2}{1-b^2} - 2\right) z + 1 = 0. \tag{2.3.1}$$

As illustrated in Fig. 1, this function is at most two-to-one, with minimum value $\mathrm{acc}(1/3) = 3 + 2\sqrt{2}$. Hence for any $z \in [1, \infty)$, it has a well defined local inverse that may be calculated as follows. If $z, 1/z$ are the two solutions to (2.3.1), write $\ell := 2 + z + 1/z$, so that $\frac{(3-b)^2}{1-b^2} = \ell$. Then we have

$$b = \frac{3 \pm \sqrt{\ell^2 - 8\ell}}{\ell + 1}, \quad \ell := z + \frac{1}{z} + 2.$$

With $\ell = \ell(z)$ as above, we will consider the following functions:

$$\mathrm{acc}_L^{-1} : \left[3 + 2\sqrt{2}, \frac{7 + 3\sqrt{5}}{2}\right] \to \left[0, \frac{1}{3}\right], \quad z \mapsto \frac{3 - \sqrt{\ell^2 - 8\ell}}{\ell + 1} \tag{2.3.2}$$

$$\mathrm{acc}_U^{-1} : \left[3 + 2\sqrt{2}, \infty\right) \to \left[\frac{1}{3}, 1\right), \quad z \mapsto \frac{3 + \sqrt{\ell^2 - 8\ell}}{\ell + 1},$$

writing acc^{-1} to denote one or other of these branches if there is no need to specify the branch any further.

The following lemma will simplify some calculations below.

Lemma 36 *Let* $\mathbf{E} = \big(d, m; q\mathbf{w}(p/q)\big)$ *be a quasi-perfect class, and suppose that* $\mathrm{acc}(b) = p/q$. *Then:*

(i) *b is given by the formula*

$$b = \frac{3pq \pm (p+q)\sqrt{\sigma}}{p^2+q^2+3pq}, \quad \text{where } \sigma := p^2+q^2-6pq. \tag{2.3.3}$$

and we use the sign +, resp. −, if $b > 1/3$, *resp.* $b < 1/3$.

(ii) $\mu_{\mathbf{E},b}(p/q) = \mu_{\mathbf{E},b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b))$ *if and only if*

$$\begin{aligned} &pq\big(3(p+q)^2 \mp (p+q)\sqrt{\sigma}\big) \\ &\quad > (p+q)\Big(d(p^2+q^2+3pq) - m(3pq \pm (p+q)\sqrt{\sigma})\Big) \end{aligned} \tag{2.3.4}$$

where we use the top, resp. bottom, signs if $b > 1/3$, *resp.* $b < 1/3$.

Proof Since

$$\ell - 8 = \frac{p}{q} + \frac{q}{p} - 6 = \frac{p^2+q^2-6pq}{pq} =: \frac{\sigma}{pq},$$

we can obtain (2.3.3) by substituting for ℓ in (2.3.2).

To prove (ii), recall from (2.2.5) that $V_b(\mathrm{acc}(b)) = \frac{1+\mathrm{acc}(b)}{3-b}$. Therefore, we have $\mu_{\mathbf{E},b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b))$ if and only if $\frac{p}{d-mb} > \frac{1+\mathrm{acc}(b)}{3-b}$. Now substitute for b using (2.3.3).

The general properties of a blocking class are rather subtle, and will be addressed elsewhere. In this paper we are only concerned with classes that are center-blocking in the following sense.

Definition 37 Let $\mathbf{E} = \big(d, m; q\mathbf{w}(p/q)\big)$ be a quasi-perfect class with center $a = p/q > 3 + 2\sqrt{2}$ and with $m/d \neq 1/3$. Then we say that $\mathbf{E}$ is **center-blocking** if $\mu_{\mathbf{E},b_0}(a) > V_{b_0}(a)$, where b_0 is the solution to $\mathrm{acc}(b_0) = a$ that lies in the same component of $[0,1)\smallsetminus\{1/3\}$ as does m/d.

Lemma 38 *Let* $\mathbf{E}$ *be a center-blocking class with center a, and let* b_0 *be the solution to* $\mathrm{acc}(b_0) = a$ *that lies in the same component of* $[0,1)\smallsetminus\{1/3\}$ *as does* m/d. *Then:*

(i) *There is an open interval* $J_{\mathbf{E}}$ *containing* b_0 *such that*

- $\mu_{\mathbf{E},b}$ *blocks b for all* $b \in J_{\mathbf{E}}$;
- $\mu_{\mathbf{E},b}(z) = V_b(z)$ *for* $z = \mathrm{acc}(b)$, $b \in \partial J_{\mathbf{E}}$.

(ii) *For all* $b \in J_{\mathbf{E}}$, *the z-interval* I_b *on which* $\mu_{\mathbf{E},b}(z)$ *is obstructive contains a.*

Proof Since all functions involved are continuous, the set of b such that $\mu_{\mathbf{E},b}$ blocks b is open. Let $J_{\mathbf{E}}$ be the connected component of this set that contains b_0. Then, for

the appropriate branch of the inverse acc^{-1},

$$\mu_{\mathbf{E},b}(\mathrm{acc}^{-1}(b)) > V_b(\mathrm{acc}^{-1}(b)), \qquad \forall\, b \in J_{\mathbf{E}}.$$

Therefore by continuity, we must have $\mu_{\mathbf{E},b}(z) = V_b(\mathrm{acc}^{-1}(b))$ when $b \in \partial J_{\mathbf{E}}$. This proves (i).

Next, observe that for all $b \in J_{\mathbf{E}}$, the interval I_b on which $\mu_{\mathbf{E},b}$ is obstructive contains a break point by Lemma 14, and because this point is the unique point of shortest length in I_b and equals a when $b = b_0$, this point must be independent of b and hence equal to a. Hence $a \in I_b$ for all $b \in J_{\mathbf{E}}$. This proves (ii).

Remark 39 One might expect that a class $\mathbf{E} = (d, m; \mathbf{m})$ would block some b only if it blocked $b = m/d$. However, this is not true even if $\mathbf{E}$ is perfect. For example, by Theorem 1, the class $\mathbf{B}_0^U = (3, 2; 1^{\times 6}) = (3, 2; \mathbf{w}(6))$ blocks the b-interval $\left(\frac{3-\sqrt{5}}{2}, \frac{3(7+\sqrt{5})}{44}\right) \approx (0.382, 0.629)$; and this interval does not contain $m/d = 2/3$, though it does contain $5/11 = \mathrm{acc}_U^{-1}(6)$. For further discussion, see Sect. 5.5. ◊

Example 40 Here are some examples of obstructive exceptional classes that are not (center)-blocking classes.

- The class $\mathbf{E} = (2, 0; 1^{\times 5})$ is perfect with center of length 5. Since its center lies outside the range of $b \mapsto \mathrm{acc}(b)$ this class is certainly not center-blocking. Moreover, because $m = 0$ the obstruction $\mu_{\mathbf{E},b}(z) = 5/2$ is constant for $z > 5$, and, as shown in Fig 1, is always $\le V_b(\mathrm{acc}(b))$. Hence $\mathbf{E}$ is not a blocking class.
- Similarly, the classes in the (ascending) staircase at $b = 1/3$ (see Example 35) are not center-blocking since their centers lie outside the range of $b \mapsto \mathrm{acc}(b)$. It is unlikely that they are blocking classes; however we do not pursue this question further here.
- The class $\mathbf{E}' := (5, 1; 2^{\times 6}, 1)$ with break point 7 gives an obstruction $\mu_{\mathbf{E}',1/5}$ at $b = 1/5$ that goes through the accumulation point $(6, 5/2)$. It is not a blocking class because, as we show in Example 34, $b = 1/5$ is the unique value of b for which $\mu_{\mathbf{E}',b}(\mathrm{acc}(b)) \ge V_b(\mathrm{acc}(b))$. ◊

Lemma 41 *All the classes (except for the first) in the Fibonacci staircase of* (3.0.1) *are center-blocking.*

Proof These classes are $\big(g_k, 0; g_{k-1}\mathbf{w}(g_{k+1}/g_{k-1})\big)$, where $g_0 = 1, g_1 = 1, g_2 = 2, g_3 = 5, \ldots$. Therefore, by well-known identities for Fibonacci numbers (see [17, §3.1]), we have $d = (p+q)/3, m = 0, d^2 = pq - 1$ and $p^2 + q^2 = 7pq - 9$. By (2.3.4) we need

$$3pq(3(p+q)^2 + (p+q)\sqrt{\sigma}) > (p+q)^2(p^2+q^2+3pq) = (p+q)^2(10pq-9).$$

Since $\sigma = pq - 9$, this holds exactly if

$$3pq\sqrt{\sigma} > (p+q)(pq-9) = 3d\sigma.$$

Thus we need $pq > d\sqrt{pq-9}$. But this holds because $d^2 = pq - 1$.

Proposition 42 *Suppose that* $\mathbf{E} = \big(d, m; q\mathbf{w}(p/q)\big)$ *is a center-blocking class with center* $a = p/q$ *that blocks the b-interval* $J_{\mathbf{E}} = (\beta_\ell, \beta_u) \subset [0, 1/3) \cup (1/3, 1)$.

(i) *if* $J_{\mathbf{E}} = (\beta_\ell, \beta_u) \subset (1/3, 1)$ *then* H_{β_u} *admits no ascending staircase and* H_{β_ℓ} *admits no descending staircase. If in addition* H_b *is unobstructed for* $b \in \partial J_{\mathbf{E}}$, *then* $\mu_{\mathbf{E},b_u}$ *is live on* $[a, \alpha_u)$ *and* $\mu_{\mathbf{E},b_\ell}$ *is live on* $(\alpha_\ell, a]$.

(ii) *if* $J_{\mathbf{E}} = (\beta_\ell, \beta_u) \subset (0, 1/3)$, *then* H_{β_ℓ} *admits no ascending staircase and* H_{β_u} *admits no descending staircase. If in addition* H_b *is unobstructed for* $b \in \partial J_{\mathbf{E}}$, *then* $\mu_{\mathbf{E},b_\ell}$ *is live on* $[a, \alpha_u)$ *and* $\mu_{\mathbf{E},b_u}$ *is live on* $(\alpha_\ell, a]$.

(iii) *If* H_b *is unobstructed for* $b \in \partial J_{\mathbf{E}}$, *then* $\mu_{\mathbf{E},b}$ *is live at* a *for all* $b \in J_{\mathbf{E}}$ *and* $\mathbf{E}$ *is perfect. Moreover,* $J_{\mathbf{E}}$ *is a connected component of Block.*

Proof In case (i) the map $b \mapsto \mathrm{acc}(b)$ preserves orientation. Consider the upper endpoint β_u of $J_{\mathbf{E}}$. By Lemma 16, the function $\mu_{\mathbf{E},\beta_u}(z)$ is constant for $z \in [a, \alpha_u := \mathrm{acc}(\beta_u))$. Also we have $\mu_{\mathbf{E},\beta_u}(\alpha_u) = V_{\beta_u}(\alpha_u)$ by Lemma 38. If H_{β_u} did have an ascending staircase, we saw in (1.1.3) that we would have to have $c_{H_{\beta_u}} = V_{\beta_u}(\alpha_u)$, i.e. H_{β_u} would be unobstructed. But then, because $c_{H_{\beta_u}}$ is nondecreasing, $c_{H_{\beta_u}}$ would be constant and equal to $\mu_{\mathbf{E},\beta_u}$ on the interval $[a, \alpha_u]$. Thus there can be no ascending staircase. (Indeed in this case $\mathbf{E}$ is left-overshadowing in the sense of Definition 29.) This proves the claims in (i) that pertain to the end point β_u.

Now consider the lower endpoint β_ℓ. By Lemma 16 the graph of $\mu_{\mathbf{E},\beta_\ell}(z) = \frac{qz}{d-m\beta_\ell}$ for $z \in [\alpha_\ell, a]$ is a line through the origin that passes through the point $\big(\alpha_\ell, V_{b_\ell}(\alpha_\ell)\big)$. If H_{β_ℓ} did have a staircase, then again we would have $V_{b_\ell}(\alpha_\ell) = c_{H_{b_\ell}}(\alpha_\ell)$, i.e. H_{β_ℓ} would be unobstructed. In this case the the scaling property (1.1.1) of capacity functions implies that the graph of $c_{H_{\beta_\ell}}$, which as we just saw goes through $(\mathrm{acc}(\beta_\ell), V_{\beta_\ell})$, cannot lie above the line $z \mapsto \mu_{\mathbf{E},\beta_\ell}(z) = \frac{qz}{d-m\beta_\ell}$ for $z > \mathrm{acc}(\beta_\ell)$. Thus we again conclude that $\mu_{\mathbf{E},\beta_\ell}$ is live for $z \in (\alpha_\ell, a]$ (and so is right-overshadowing), and that there cannot be a decreasing staircase for H_{β_ℓ}. This completes the proof of (i).

The proof of (ii) is similar, and is left to the reader.

Now consider (iii). For clarity we will again suppose that $J_{\mathbf{E}} \subset (1/3, 1)$, leaving the other case to the reader. First note that if H_b is unobstructed then $b \notin \mathit{Block}$, so that $J_{\mathbf{E}}$ is a connected component of *Block* as claimed.

We next show that $\mu_{\mathbf{E},b}$ is live at a for all $b \in J_{\mathbf{E}}$. To see this, consider the set of $b \in [\beta_\ell, \beta_u]$ such that $\mu_{\mathbf{E},b'}$ is live at a for all $b' \in [\beta_\ell, b)$. This is a closed, connected subset of $[\beta_\ell, \beta_u]$. If it is empty, define $b_{\max} := \beta_\ell$, and if it is nonempty but proper, let $b_{\max} < \beta_u$ be its maximal element. Then, by Lemma 15 (ii), there must be an exceptional class $\mathbf{E}' = (d', m'_0; \mathbf{m}')$ and $\varepsilon > 0$ such that

$$\mu_{\mathbf{E}',b_{\max}}(a) = \frac{\mathbf{m}' \cdot \mathbf{w}(a)}{d' - m'_0 b_{\max}} = \frac{p}{d - m b_{\max}} = \mu_{\mathbf{E},b_{\max}}(a)$$

and $\mu_{\mathbf{E}_{\max},b}(a) > \mu_{\mathbf{E},b}(a)$ for $b \in (b_{max}, b_{max} + \varepsilon)$. In particular,

$$\begin{aligned}\frac{\partial}{\partial b}\Big|_{b=b_{\max}}\mu_{\mathbf{E}',b}(a) &= \mu_{\mathbf{E}',b_{\max}}(a)\ \frac{m'}{d'-m'_0 b_{\max}}.\\ &\geq \frac{\partial}{\partial b}\Big|_{b=b_{\max}}\mu_{\mathbf{E},b}(a)\\ &= \mu_{\mathbf{E},b_{\max}}(a)\ \frac{m}{d-mb_{\max}}.\end{aligned}$$

But this implies that $m'\ d \geq m\ d'$. Thus, because this inequality is independent of $b_{\max}$, it can hold at $b = b_{\max}$ only if it holds for all $b \in [b_{\max}, \beta_u]$. Moreover, we cannot have $m'\ d = m\ d'$ since this would imply that the two obstruction functions are equal for $b \in [b_{max}, b_{max} + \varepsilon)$. Hence we must have $m'\ d > m\ d'$, in which case $\mu_{\mathbf{E},b}(a)$ could not be live at $b = \beta_u$. Therefore this scenario does not happen, and so $b_{\max} = \beta_u$. This completes the proof that $\mu_{\mathbf{E},b}(a)$ is live at a for all $b \in J_{\mathbf{B}}$.

It remains to show that $\mathbf{E}$ is perfect. By Lemma 15 (ii), there is an exceptional class $\mathbf{E}' = (d', m'; \mathbf{m}')$ and an open subset J_b of $J_{\mathbf{E}}$ such that

$$\mu_{\mathbf{E}',b}(a) = \mu_{\mathbf{E},b}(a), \quad \forall\, b \in J_b.$$

Then we may write $\mathbf{m}' = \lambda\mathbf{m} + \mathbf{n}$ where $\mathbf{m} \cdot \mathbf{n} = 0$. Since

$$\frac{\mathbf{m}' \cdot \mathbf{w}(a)}{d' - m'b} = \frac{\lambda p}{d' - m'b} = \frac{p}{d - mb}$$

for all $b \in J_b$, we must have $d' = \lambda d,\ m' = \lambda m$ for some $\lambda > 0$. The identities

$$\begin{aligned}\mathbf{m}' \cdot \mathbf{m}' - 1 &= d'^2 - (m')^2 = \lambda^2(d^2 - m^2) = \lambda^2(pq - 1),\\ \mathbf{m}' \cdot \mathbf{m}' - 1 &= \lambda^2\mathbf{m} \cdot \mathbf{m} + \|\mathbf{n}\|^2 - 1 = \lambda^2 pq + \|\mathbf{n}\|^2 - 1\end{aligned}$$

then imply that $\|\mathbf{n}\|^2 = 1 - \lambda^2$. Therefore, unless $\mathbf{E}' = \mathbf{E}$ we must have $0 < \lambda < 1$. Further

$$\mathbf{E}' \cdot \mathbf{E} = d'd - m'm - \mathbf{m}' \cdot \mathbf{m} = \lambda(d^2 - m^2 - \mathbf{m} \cdot \mathbf{m}) = -\lambda.$$

But $\mathbf{E}' \cdot \mathbf{E}$ is an integer. It follows that $\mathbf{E}' = \mathbf{E}$, so that $\mathbf{E}$ is perfect as claimed.

Remark 43

(i) Notice that there may be no $b \in J_{\mathbf{E}}$ such that $\mu_{\mathbf{E},b}$ is live on the whole of the z-interval on which it is obstructive, since for each such b there may be classes with break points a' outside this interval (and hence with $\ell(a') \leq \ell(a)$) that are live near one end or other of this z-interval. For further discussion of this point, see Sect. 5.5 and the associated figures.

(ii) We will see in Proposition 49 that many of the classes that contribute to a staircase are themselves center-blocking classes. For example, Fig. 9 shows

that the class $\mathbf{E}'$ with center $170/29$, which is the $k = 0$ step with $\mathrm{end}_0 = 4$ of the staircase $\mathcal{S}^E_{u,0}$, obstructs at $\mathrm{acc}^{-1}(170/29)$, and hence is center-blocking. $\Diamond$

The next lemma is the key to the proofs of our results about the relation between blocking classes and staircases.

Lemma 44 *Let* $\mathbf{B} = \big(d, m; q\mathbf{w}(p/q)\big)$ *be a quasi-perfect class with* $m/d \neq 1/3$ *and define* acc^{-1} *to be the branch of the inverse whose image contains* m/d.

(i) *If* $\mathbf{B} = \big(d, m; q\mathbf{w}(p/q)\big)$ *is center-blocking with* $J_{\mathbf{B}} = (\beta_{\mathbf{B},\ell}, \beta_{\mathbf{B},u})$, *and* $I_{\mathbf{B}} = (\alpha_{\mathbf{B},\ell}, \alpha_{\mathbf{B},u}) = \mathrm{acc}(J_{\mathbf{B}})$, *then*

$$\begin{aligned} \mathrm{acc}^{-1}(\alpha_{\mathbf{B},\ell}) &= \frac{(1+\alpha_{\mathbf{B},\ell})d - 3q\alpha_{\mathbf{B},\ell}}{(1+\alpha_{\mathbf{B},\ell})m - q\alpha_{\mathbf{B},\ell}}, \\ \mathrm{acc}^{-1}(\alpha_{\mathbf{B},u}) &:= \frac{(1+\alpha_{\mathbf{B},u})d - 3p}{(1+\alpha_{\mathbf{B},u})m - p}. \end{aligned} \tag{2.3.5}$$

(ii) *Conversely, suppose given numbers* $z_\ell < p/q < z_u$ *such that* $\ell(p/q) < \ell(z)$ *for all* $z \in (z_\ell, z_u) \smallsetminus \{p/q\}$ *and*

$$\begin{aligned} \mathrm{acc}^{-1}(z_\ell) &= \frac{(1+z_\ell)d - 3qz_\ell}{(1+z_\ell)m - qz_\ell}, \\ \mathrm{acc}^{-1}(z_u) &:= \frac{(1+z_u)d - 3p}{(1+z_u)m - p}. \end{aligned} \tag{2.3.6}$$

Then $\mathbf{B}$ *is center-blocking, and we have* $I_{\mathbf{B}} = (z_\ell, z_u)$.

Proof Let $c(b) := \frac{(3-b)^2}{1-b^2} - 2$. Then, because $\mathrm{acc}(b) > 0$ and $0 \le b < 1$, we have

$$\begin{aligned} \frac{1+\mathrm{acc}(b)}{3-b} = \sqrt{\frac{\mathrm{acc}(b)}{1-b^2}} \quad &\Longleftrightarrow \quad (1+\mathrm{acc}(b))^2 = \mathrm{acc}(b)\frac{(3-b)^2}{1-b^2} = \mathrm{acc}(b)\big(c(b)+2\big), \\ &\Longleftrightarrow \quad \mathrm{acc}(b)^2 - c(b)\mathrm{acc}(b) + 1 = 0, \end{aligned} \tag{2.3.7}$$

which holds by the definition of $\mathrm{acc}(b)$ in (1.1.2). Therefore the function

$$b \mapsto V_b(\mathrm{acc}(b)) = \sqrt{\frac{\mathrm{acc}(b)}{1-b^2}}$$

is also given by the formula $b \mapsto \frac{1+\mathrm{acc}(b)}{3-b}$.

Since $\mathbf{B}$ is a quasi-perfect class, Lemma 16 shows that the obstruction function $z \mapsto \mu_{\mathbf{B},b}(z)$ on $I_{\mathbf{B}}$ is given by the formulas $z \mapsto \frac{qz}{d-mb}$ for $z < a$ and $z \mapsto \frac{p}{d-mb}$ for $z > a$. Thus because $V_b(\alpha_{\mathbf{B},\ell}) = \mu_{\mathbf{B},b}(\alpha_{\mathbf{B},\ell})$, the point $z = \alpha_{\mathbf{B},\ell}$ satisfies

$$\frac{1+\alpha_{\mathbf{B},\ell}}{3-b} = \frac{q\alpha_{\mathbf{B},\ell}}{d-mb}, \quad \text{where } b := \mathrm{acc}^{-1}(\alpha_{\mathbf{B},\ell}).$$

Similarly, the point $z = \alpha_{\mathbf{B},u}$ satisfies the equation

$$\frac{1+\alpha_{\mathbf{B},u}}{3-b} = \frac{p}{d-mb}, \quad \text{where } b := \mathrm{acc}^{-1}(\alpha_{\mathbf{B},u}).$$

Now rearrange these identities to obtain (2.3.5). This proves (i).

The identities in (2.3.7) show that $z = \mathrm{acc}(b)$ if and only if $(1+z)/(3-b) = V_b(z)$. Because the length $\ell(p/q)$ of the point p/q is minimal among all points in (z_ℓ, z_u), Lemma 16 implies that

$$\mu_{\mathbf{B},b}(z_\ell) = \frac{qz_\ell}{d-mb}, \quad \mu_{\mathbf{B},b}(z_u) = \frac{p}{d-mb}.$$

We saw in the proof of (i) above that

$$b = \frac{(1+z_\ell)d - 3qz_\ell}{(1+z_\ell)m - qz_\ell} \iff \frac{1+z_\ell}{3-b} = \frac{qz_\ell}{d-mb}.$$

Hence if $b = \mathrm{acc}^{-1}(z_\ell)$ is given by the formula $b = \frac{(1+z_\ell)d-3qz_\ell}{(1+z_\ell)m-qz_\ell}$, then

$$V_b(z_\ell) = \frac{1+z_\ell}{3-b} = \mu_{\mathbf{B},b}(z_\ell), \quad b := \mathrm{acc}^{-1}(z_\ell).$$

A similar argument with z_u shows that our hypothesis implies

$$V_b(z_u) = \frac{1+z_u}{3-b} = \mu_{\mathbf{B},b}(z_u), \quad b := \mathrm{acc}^{-1}(z_u)$$

Hence the constraint defined by the class **B** equals the volume obstruction at the two pairs $(z,b) = (z_\ell, \mathrm{acc}^{-1}(z_\ell))$, and $(z,b) = (z_u, \mathrm{acc}^{-1}(z_u))$ where $z_\ell < z_u$. It follows that **B** is a center blocking class that blocks the z-interval (z_ℓ, z_u), as claimed.

Corollary 45 *If $\mathcal{S}$ is a staircase in H_b that accumulates at a point a_∞ that is an endpoint of the blocked z-interval $I_{\mathbf{B}}$ for some quasi-perfect blocking class* **B***, then the two numbers a_∞, b can be expressed in terms of the same quadratic surd, i.e. there is $\sigma \in \mathbb{N}$ such that $a_\infty, b \in \mathbb{Q} + \mathbb{Q}\sqrt{\sigma}$.*

Proof This is an immediate consequence of Lemma 44 (i).

2.4 Pre-staircases and Blocking Classes

In this section we begin by discussing the structure of the staircases that we encounter, and then relate their properties to those of associated blocking classes.

Our aim is to clarify exactly what we need to prove in order to show that a particular sequence of Diophantine classes $(\mathbf{E}_k)$ does form a staircase.

Definition 46 We will say that a sequence $\mathcal{S} = \big(\mathbf{E}_k\big)_{k\geq 0}$ of quasi-perfect classes $\mathbf{E}_k := \big(d_k, m_k; q_k\mathbf{w}\,(p_k/q_k)\big)$ is a **pre-staircase** if it has the following properties:

- **(Recursion)** There is an integer $\sigma \geq 0$ such that $\sigma + 4$ is a perfect square, and each of the sequences $x_k := d_k, m_k, p_k, q_k$ satisfies the recursion

$$x_{k+1} = (\sigma + 2)x_k - x_{k-1} \quad \text{for all } k \geq 0, \tag{2.4.1}$$

- **(Relation)** there are integers R_0, R_1, R_2 such that the following linear relation holds

$$R_0 d_k = R_1 p_k + R_2 q_k \quad \text{for all } k \geq 0, \tag{2.4.2}$$

Moreover, if the classes $\mathbf{E}_k$ are perfect for all k, then we say that $\mathcal{S}$ is a **perfect pre-staircase.**

In this situation, the whole sequence of classes is determined by the first two centers $p_0/q_0,\ p_1/q_1$, since the other centers are then determined by the recursion, the d_k are determined by the linear relation (2.4.2), and then the m_k are determined by the linear Diophantine identity $3d_k = m_k + p_k + q_k$. Of course, for arbitrary choices of initial data, there is no guarantee that d_k, if so defined, is a positive integer, or that the quadratic Diophantine identity holds.

The following lemma explains the importance of the (Recursion) condition.

Lemma 47 *Let $x_k, k \geq 0$, be a sequence of integers that satisfy the recursion $x_{k+1} = (\sigma + 2)x_k - x_{k-1}$, where $\sigma + 4$ is a perfect square, and let $\lambda \in \mathbb{Q}[\sqrt{\sigma}]$ be the larger root of the equation $x^2 - (\sigma + 2)x + 1 = 0$. Then there is a number $X \in \mathbb{Q}[\sqrt{\sigma}]$ such that*

$$x_k = X\lambda^k + \overline{X}\,\overline{\lambda}^k, \tag{2.4.3}$$

where $\overline{a + b\sqrt{\sigma}} := (a - b\sqrt{\sigma})$, so that $\lambda\overline{\lambda} = 1$.

Proof If the monomials $x_k = c^k$ satisfy the recursion then we must have $c^2 - (\sigma + 2)c + 1 = 0$, so that $c = \big((\sigma + 2) \pm N\sqrt{\sigma}\big)/2$, where $N^2 = \sigma + 4$. Let λ be the larger solution, so that $\overline{\lambda}$ is the smaller one, and we have $\lambda\overline{\lambda} = 1$. Since (2.4.3) has a unique solution once given the seeds x_0, x_1, it follows that for each choice of constants A, B, the numbers

$$x_k := A\lambda^k + B\overline{\lambda}^k$$

form the unique solution with

$$x_0 = A + B, \qquad x_1 = A\lambda + B\overline{\lambda}.$$

Then $A, B \in \mathbb{Q}[\sqrt{\sigma}]$, and it is easy to check that $x_0, x_1 \in \mathbb{Q}$ only if we also have $B := \overline{A}$. This completes the proof.

Remark 48

(i) All the pre-staircases that we consider have $\sigma = (2n+1)(2n+5)$ for some $n \geq 0$ so that $\sigma + 4 = (2n+3)^2$, a quantity whose square root happens to equal the constant R_0 in (Relation). Note that, if $X = X' + X''\sqrt{\sigma}$ is as in Lemma 47 and the initial values are x_0, x_1, we have

$$\lambda = \frac{\sigma + 2 + (2n+3)\sqrt{\sigma}}{2}, \quad X' = \frac{x_0}{2}, \quad X'' = \frac{2x_1 - x_0(\sigma+2)}{2(2n+3)\sigma}. \tag{2.4.4}$$

(ii) It turns out that, as in Lemma 72 (i) below, some of our staircases can be extended by a 'class' $\big(d_{-1}, m_{-1}, q_{-1}\mathbf{w}(p_{-1}/q_{-1})\big)$ defined for $k = -1$ such that the (Recursion) (2.4.3) and (Relation) (2.4.2) conditions hold for all $k \geq -1$. The word 'class' above is in quotes because in some cases the numbers d_{-1}, m_{-1} are negative (though p_{-1}, q_{-1} are positive), so that the tuples have no geometric meaning. However, they can still be used for computational purposes. For example a quantity such as X in (2.4.3) can be computed from knowledge of the terms x_{-1}, x_0 instead of from x_0, x_1. Note in particular that if m_k, d_k both satisfy the recursion (2.4.1) for $k \geq -1$, then

$$m_0(d_1 + d_{-1}) = m_0 d_0(\sigma + 2) = d_0(m_{-1} + m_1),$$

which implies that

$$m_0 d_1 - m_1 d_0 = m_{-1} d_0 - m_0 d_{-1}. \tag{2.4.5}$$

This fact will simplify some calculations below. ◊

The following result shows that at least the tail end of a pre-staircase consists of center-blocking classes.

Proposition 49 *Suppose that $\mathcal{S} = (\mathbf{E}_k)$ is a pre-staircase as above, let λ be as in Lemma 47, and denote by D, M, P, Q the constants X defined by (2.4.3), where $x_k = d_k, m_k, p_k, q_k$ respectively. Suppose that $M/D \neq 1/3$ and that the centers p_k/q_k of $\mathcal{S}$ are all $> 3 + 2\sqrt{2}$. Then*

- *$P/Q = \lim p_k/q_k$, $M/D = \lim m_k/d_k$ and*

$$\operatorname{acc}\left(\frac{M}{D}\right) = \frac{P}{Q}. \tag{2.4.6}$$

- $\mathbf{E}_k$ *is a center-blocking class for sufficiently large* k*, and*
- *If in addition* $\mathcal{S}$ *is a perfect pre-staircase then* $b_\infty := M/D$ *is unobstructed.*

Proof The identities

$$\frac{P}{Q} = \lim \frac{p_k}{q_k}, \qquad \frac{M}{D} = \lim \frac{m_k}{d_k}$$

follow immediately from Lemma 47. Since $3d_k = m_k + p_k + q_k$ and $d_k^2 - m_k^2 = p_k q_k - 1$ we have $3D = M + P + Q$ and $D^2 - M^2 = PQ$. (Note that the 1 in the second identity disappears in the limit, after we divide by λ^{2k}.) By (2.3.7), we have $\mathrm{acc}(M/D) = P/Q$ exactly if

$$\frac{1 + \frac{P}{Q}}{3 - \frac{M}{D}} = \sqrt{\frac{\frac{P}{Q}}{1 - (\frac{M}{D})^2}}$$

or equivalently

$$\frac{P + Q}{\sqrt{PQ}} = \frac{3D - M}{\sqrt{D^2 - M^2}}.$$

But we saw above that the top and bottom entries on both sides are equal.

To prove that $\mathbf{E}_k$ is center-blocking, we must check that $\mu_{E_k, b_k}(p_k/q_k)$ is nontrivial, where $b_k := \mathrm{acc}^{-1}(p_k/q_k)$ and we choose the inverse so that M/D is in its range. Hence by Lemma 15 we must check that $|d_k b_k - m_k| < \sqrt{1 - b_k^2}$ for sufficiently large k.

To see this, suppose that $M/D < 1/3$. (For the other case, we simply choose $+$ instead of $-$ in (2.3.2).) Then (2.3.2) implies that

$$b_k = \frac{3 - \sqrt{\ell_k^2 - 8\ell_k}}{\ell_k + 1}, \qquad \text{where} \quad \ell_k = \frac{p_k}{q_k} + \frac{q_k}{p_k} + 2.$$

Since $\mathrm{acc}(M/D) = P/Q$, we know that

$$\frac{M}{D} = \frac{3 - \sqrt{L^2 - 8L}}{L + 1}, \quad L := \frac{P}{Q} + \frac{Q}{P} + 2.$$

But our assumptions imply that $\ell_k = L + O(\lambda^{-2k})$. Hence

$$b_k = \frac{M}{D} + O(\lambda^{-2k}).$$

Therefore

$$|d_k b_k - m_k| = |(D\lambda^k + O(\lambda^{-k}))(\frac{M}{D} + O(\lambda^{-2k})) - (M + O(\lambda^{-2k}))\lambda^k| = O(\lambda^{-k}).$$

Therefore, because $1 - b_k^2 = 1 - (M/D)^2 - O(\lambda^{-2k}) > 0$ for large k, the required inequality $|d_k b_k - m_k| < \sqrt{1 - b_k^2}$ holds for sufficiently large k. This proves the second point.

Finally, if the classes $\mathbf{E}_k$ are perfect then M/D is unobstructed by Lemma 27.

Remark 50

(i) Proposition 49 does not quite give an independent proof that any staircase in the manifold H_b must accumulate at $\mathrm{acc}(b)$, since it is not true that a staircase must be a pre-staircase. For example, we saw in Example 35 that the staircase in $H_{1/3}$ is not a pre-staircase since the classes do not satisfy a homogeneous linear relation. Further, a staircase need not be given by quasi-perfect classes, though no such examples are known.

(ii) It is very likely that, as in Lemma 41, all the classes $\mathbf{E}_k$ in the staircases defined in Sect. 3.1 are center-blocking, provided that their centers are in the range of the function $b \mapsto \mathrm{acc}(b)$. One could prove this in any particular case by being more careful with the estimates in the above lemma. ◊

Here is our most powerful staircase recognition criterion.

Theorem 51 *Let $\mathcal{S} = (\mathbf{E}_k)$ be a perfect pre-staircase with constants P, Q, D, M as in Proposition 49. Suppose in addition that at least one of the following conditions holds:*

(i) *There is $r/s > 0$ such that $M/D < r/s$,*

$$\frac{m_k^2 - 1}{d_k m_k} < \frac{M}{D} < \frac{s + m_k(rd_k - sm_k)}{r + d_k(rd_k - sm_k)}, \qquad \forall\, k \geq k_0,$$

and there is no overshadowing class at $(z, b_\infty) = (P/Q, M/D)$ of degree $d' < s/(r - sb_\infty)$ and with $m'/d' > r/s$.

(ii) *There is $r/s > 0$ such that $M/D > r/s$,*

$$\frac{m_k(sm_k - rd_k) - s}{d_k(sm_k - rd_k) - r} < \frac{M}{D} < \frac{m_k}{d_k} \qquad \forall\, k \geq k_0,$$

and there is no overshadowing class at $(z, b_\infty) = (P/Q, M/D)$ of degree $d' < s/(sb_\infty - r)$ and with $m'/d' < r/s$.

Then $\mathcal{S}$ is a staircase for $H_{M/D}$ that accumulates at P/Q.

Proof Proposition 49 implies that $\mathrm{acc}(M/D) = P/Q$ and that $H_{M/D}$ is unobstructed. Further $m_k/d_k \to M/D$ and $p_k/q_k \to P/Q$ by Lemma 47. Hence it remains to show that the staircase is live. If (i) holds this follows from Proposition 30, while if (ii) holds we argue as in Remark 31.

So far, the linear relation satisfied by a pre-staircase has played no role in our analysis except to specify the entries d_k, m_k of the classes $\mathbf{E}_k$. Our second main result in this section shows the relevance of this relation, using it to obtain the following blocking class recognition criterion.

Theorem 52 *Let* $\mathbf{B} = \big(d, m, q\mathbf{w}(p/q)\big)$ *be a quasi-perfect class with* $d \neq 3m$ *and suppose given two perfect pre-staircases* $\mathcal{S}_\ell, \mathcal{S}_u$ *that satisfy the following conditions:*

(i) *If the constants for* $\mathcal{S}_\ell$ *(resp.* $\mathcal{S}_u$*) are denoted* $D_\ell, M_\ell, P_\ell, Q_\ell$ *(resp.* D_u, M_u, P_u, Q_u*), then*

$$\frac{M_\ell}{D_\ell} \neq \frac{1}{3} \neq \frac{M_u}{D_u}, \qquad 3 + 2\sqrt{2} < \frac{P_\ell}{Q_\ell} < \frac{p}{q} < \frac{P_u}{Q_u}.$$

Moreover $\ell(p/q) < \ell(z)$ *for all* $z \neq p/q$ *in* $\big(P_\ell/Q_\ell, P_u/Q_u\big)$.

(ii) $\mathcal{S}_\ell$ *is ascending with linear relation* $R_0 d_k = R_1 p_k + R_2 q_k$ *where* $R_0 = d - 3m$, $R_1 = q - m$, $R_2 = -m$.

(iii) $\mathcal{S}_u$ *is descending with linear relation* $R_0 = d - 3m$, $R_1 = -m$, $R_2 = p - m$.

Then $\mathbf{B}$ *is a perfect blocking class that blocks the z-interval* $I_{\mathbf{B}} = (\alpha_\ell, \alpha_u)$ *where*

$$\alpha_\ell = \frac{P_\ell}{Q_\ell},\ \mathrm{acc}^{-1}(\alpha_\ell) = \frac{M_\ell}{D_\ell}, \quad \alpha_u = \frac{P_u}{Q_u},\ \mathrm{acc}^{-1}(\alpha_u) = \frac{M_u}{D_u} \tag{2.4.7}$$

Proof Since the hypotheses of Proposition 49 hold, we know that

$$\mathrm{acc}\Big(\frac{M_\ell}{D_\ell}\Big) = \frac{P_\ell}{Q_\ell}, \quad \mathrm{acc}\Big(\frac{M_u}{D_u}\Big) = \frac{P_u}{Q_u}.$$

Moreover, because we assumed that the pre-staircases are perfect, both M_ℓ/D_ℓ and M_u/D_u are unobstructed.

By Lemma 44 (ii), to see that $\mathbf{B}$ is a class that blocks the given interval $I_{\mathbf{B}}$ it suffices to check that the equations in (2.3.6) hold when d, p, m are defined by $\mathbf{B}$ and with $z_\ell = P_\ell/Q_\ell,\ z_u = P_u/Q_u$. Thus we must check that

$$\frac{M_\ell}{D_\ell} = \frac{3qP_\ell - (Q_\ell + P_\ell)d}{qP_\ell - (Q_\ell + P_\ell)m} = \frac{(3q-d)P_\ell - dQ_\ell}{(q-m)P_\ell - mQ_\ell} \quad \text{and}$$
$$\frac{M_u}{D_u} = \frac{3pQ_u - (Q_u + P_u)d}{pQ_u - (Q_u + P_u)m} = \frac{(3p-d)Q_u - dP_u}{(p-m)Q_u - mP_u}$$

Because

$$(q-m)P_\ell - mQ_\ell = R_1 P_\ell + R_2 Q_\ell = (d-3m)D_\ell,$$

the first identity will hold if $(d-3m)M_\ell = (3q-d)P_\ell - dQ_\ell$. But because $M_\ell = 3D_\ell - P_\ell - Q_\ell$ we have

$$\begin{aligned}(d-3m)M_\ell &= 3(d-3m)D_\ell - (d-3m)P_\ell - (d-3m)Q_\ell \\ &= 3\big((q-m)P_\ell - mQ_\ell\big) - (d-3m)P_\ell - (d-3m)Q_\ell \\ &= (3q-d)P_\ell - dQ_\ell\end{aligned}$$

as required. The proof of the second identity is similar; notice that again the coefficients in the denominator equal those in the linear relation for $\mathcal{S}_u$.

This shows that **B** is a blocking class that blocks the interval (α_ℓ, α_u). Since, as we saw in the first paragraph of this proof, H_{β_ℓ}, H_{β_u} are unobstructed, the class **B** is perfect by Proposition 42 (iii).

3 The Fibonacci Stairs, Its Cognates, and Beyond

The staircases $\mathcal{S}_n^E$ in Theorem 2 were found by trial and error using methods described in Sect. 5 and Remark 19. They occur for values of $b \in \big(1/5, 1/3\big)$. Once we began looking for staircases in other ranges of b, armed with the numerical knowledge of these stairs, nascent ideas about the importance of blocking classes, as well as the visual and computational tools explained in Sect. 5, we found many other examples. In Sect. 3.1 we describe three important sets of blocking classes, together with their associated staircases, and explain their relation to the staircases in Theorem 2. In Sect. 3.2, we explain our proof strategy.

We assume that the reader understands the definition of (quasi-)perfect classes (Definition 12), center-blocking class (Definition 37) and pre-staircase (Definition 46). The most important results are Proposition 49 and Theorem 51.

To put our work in context, recall from [17] that the Fibonacci stairs are given by a family of exceptional divisors

$$\big(g_k, 0; g_{k-1}\mathbf{w}(\frac{g_{k+1}}{g_{k-1}})\big) \tag{3.0.1}$$

where the $(g_k)_{k\geq 0}$ are the odd placed Fibonacci numbers $1, 2, 5, 13, 34, \cdots$. The continued fractions of the center points g_{k+1}/g_{k-1} divide naturally into two classes: the elements in the odd places have centers

$$5,\ [6;1,4],\ [6;1,5,1,4],\ \ldots, [6;1,\{5,1\}^k,4], \ldots$$

while those in the even places have centers

$$[6;2],\ [6;1,5,2],\ [6;1,5,1,5,2],\ \ldots, [6;1,\{5,1\}^k,5,2], \ldots.$$

Moreover, each of these classes form a pre-staircase in the sense of Definition 46, with recursion $x_{k+1} = 7x_k - x_{k-1}$ and linear relation $3d_k = p_k + q_k$.

The staircases that we describe below all have similar numerics, and limit at points $a_\infty = \lim p_k/q_k$ with 2-periodic continued fractions. There are other pre-staircases with continued fractions of higher periods; these will be discussed in our next paper.

3.1 The Main Theorems

We now describe three families of perfect center-blocking classes, $(\mathbf{B}_n^U)$, $(\mathbf{B}_n^L)$, and $(\mathbf{B}_n^E)$ together with their associated staircases. The classes $(\mathbf{B}_n^U)$, $(\mathbf{B}_n^E)$ are those in Theorems 1 and 5. We will begin with the classes $(\mathbf{B}_n^L)$, since, as explained in Remark 55, we can consider the Fibonacci staircase to be the initial (slightly anomalous) member of the corresponding family of ascending staircases $(\mathcal{S}_{\ell,n}^L)$. We will say that an ascending (respectively descending) staircase $\mathcal{S}_\ell$ (resp. $\mathcal{S}_u$) is associated to the blocking class $\mathbf{B}$ if it accumulates at the point $\alpha_{\mathbf{B},\ell}$ (resp. $\alpha_{\mathbf{B},u}$). Thus staircases labelled ℓ always ascend, while those labelled u always descend, regardless of whether the associated value of b is in $(0, 1/3)$ or $(1/3, 1)$.

Remark 53 Each of the staircases below consists of two intertwining sequences of classes as follows

- the first has centers $p_{n,k}/q_{n,k}$ for $k \geq 0$, with $\text{end}_n = 2n + 4$, and
- the second has centers $p'_{n,k}/q'_{n,k}$ for $k \geq 0$, with $\text{end}_n = (2n + 5, 2n + 2)$.

Each such sequence is a pre-staircase in the sense of Definition 46. When the first of the end entries occurs in an even place (as with the ascending staircase $\mathcal{S}_{\ell,n}^L$) then the fact that $2n + 4 < 2n + 5$ means that $p_{n,k}/q_{n,k} < p'_{n,k}/q'_{n,k}$.[9] Similarly, when the first of the end entries occurs in an odd place we have $p_{n,k}/q_{n,k} > p'_{n,k}/q'_{n,k}$.

Recall also from Definition 24 that we use the word 'staircase' rather loosely; thus it could refer to just one of these sequences of classes, or both, depending on context. ◊

Theorem 54 *The classes* $\mathbf{B}_n^L = \left(5n, n - 1; 2n\mathbf{w}((12n + 1)/(2n))\right), n \geq 1$, *with decreasing centers, are perfect and center-blocking, and have the following associated staircases* $\mathcal{S}_{\ell,n}^L, \mathcal{S}_{u,n}^L$ *for* $n \geq 1$*:*

- $\mathcal{S}_{\ell,n}^L$ *is ascending, with limit point* $a_{\ell,n,\infty}^L = [6; 2n + 1, \{2n + 5, 2n + 1\}^\infty]$, *and has*

[9] Remark 76 explains the order properties of points that are specified in terms of their continued fraction expansions.

$$\begin{array}{ll}
\text{(Centers)} & [6;\, 2n+1, \{2n+5, 2n+1\}^k, \text{end}_n], \\
& \qquad \text{end}_n = 2n+4 \quad \textit{or } (2n+5, 2n+2),\ k \geq 0; \\
\text{(Recursion)} & x_{n,k+1} = (\sigma_n + 2)x_{n,k} - x_{n,k-1}, \quad \sigma_n := (2n+1)(2n+5) \\
\text{(Relation)} & (2n+3)d_{n,k} = (n+1)\,p_{n,k} - (n-1)q_{n,k}.
\end{array}$$

- *$\mathcal{S}^L_{u,n}$ is descending, with limit point $a^L_{u,n,\infty} = [6;\, 2n-1, 2n+1, \{2n+5, 2n+1\}^\infty]$ and has*

$$\begin{array}{ll}
\text{(Centers)} & [6;\, 2n-1, 2n+1, \{2n+5, 2n+1\}^k, \text{end}_n], \\
\text{(Recursion)} & x_{n,k+1} = (\sigma_n + 2)x_{n,k} - x_{n,k-1}, \\
\text{(Relation)} & (2n+3)d_{n,k} = -(n-1)\,p_{n,k} + (11n+2)q_{n,k}
\end{array}$$

  with the same possibilities for end_n *and the same* σ_n.

The limit points $a^L_{\bullet,n,\infty}$ (with $\bullet = \ell$ or u) form a decreasing sequence in $(6, 7)$ *with limit* 6*, while the corresponding b-values lie in* $(0, 1/5)$ *and increase with limit* $1/5$.

Remark 55 The Fibonacci stairs has the same numerics as the case $n = 0$ of $\mathcal{S}^L_{\ell,n}$. However, there cannot be an associated blocking class since this would have to block a z-interval with lower endpoint α^F_ℓ equal to the limit point τ^4. But then $\text{acc}_L^{-1}(\alpha^F_\ell) = 0$ would be the upper endpoint of the corresponding b-interval, which is clearly impossible. Nevertheless, the obstruction given by the class $\mathbf{E}_0 := (3, 0;\, 2, 1^{\times 6})$ with break point $a = 7$ does go through the accumulation point $\big(\tau^4, V_0(\tau^4) = \tau^2\big)$, and this class can be considered as a substitute for the blocking class. In this paper, we will often ignore this distinction and will use the notation $\mathcal{S}^L_{\ell,0}$ to refer to the Fibonacci stairs. ◊

Here is the analogous result for the blocking classes in Theorem 1.

Theorem 56 *The classes* $\mathbf{B}^U_n = \big(n+3, n+2; \mathbf{w}(2n+6)\big)$, $n \geq 0$, *with increasing centers, are perfect and center-blocking, with the following associated staircases $\mathcal{S}^U_{\ell,n}$, $\mathcal{S}^U_{u,n}$, where σ_n and* end_n *are as in Theorem 54.*

- *for each $n \geq 1$, $\mathcal{S}^U_{\ell,n}$ has limit point $a^U_{\ell,n,\infty} = [\{2n+5, 2n+1\}^\infty]$, and has*

$$\begin{array}{ll}
\text{(Centers)} & [\{2n+5, 2n+1\}^k, \text{end}_n], \\
\text{(Recursion)} & x_{n,k+1} = (\sigma_n + 2)x_{n,k} - x_{n,k-1}, \\
\text{(Relation)} & (2n+3)d_{n,k} = (n+1)\,p_{n,k} + (n+2)q_{n,k}.
\end{array}$$

- *for each $n \geq 0$, $\mathcal{S}^U_{u,n}$ has limit point $a^U_{u,n,\infty} = [2n+7; \{2n+5, 2n+1\}^\infty]$, and has*

$$\begin{array}{ll}
\text{(Centers)} & [2n+7; \{2n+5, 2n+1\}^k, \text{end}_n], \\
\text{(Recursion)} & x_{n,k+1} = (\sigma_n + 2)x_{n,k} - x_{n,k-1}, \\
\text{(Relation)} & (2n+3)d_{n,k} = (n+2)\,p_{n,k} - (n+4)q_{n,k}
\end{array}$$

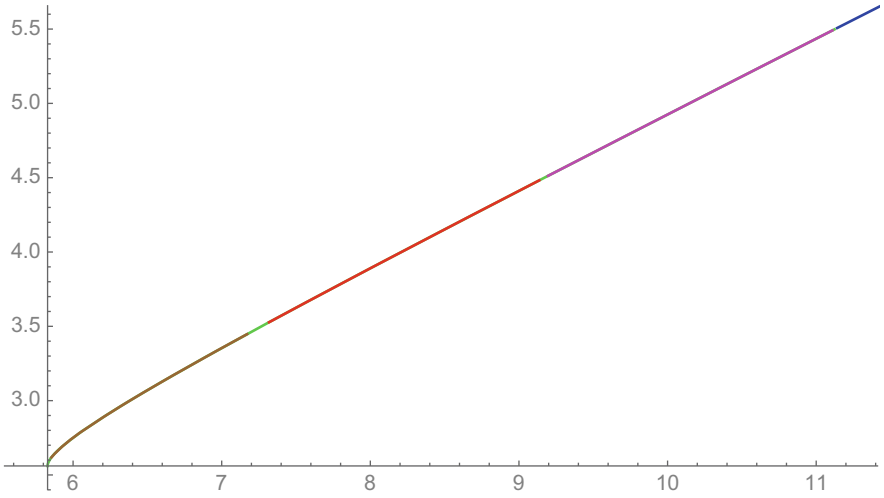


Fig. 4 Depicted are the intervals $J_{B_n^U}$ for $n = 0, \ldots, 2$ and part of the interval $J_{B_3^U}$, in brown, red, pink, and blue, respectively, for $1/3 < b < 0.795$, transplanted onto the accumulation point curve $b \mapsto (\mathrm{acc}(b), V_b(\mathrm{acc}(b)))$. The infinite staircases $\mathcal{S}^U_{\ell,n}$ accumulate at the left endpoints of the intervals and the infinite staircases $\mathcal{S}^U_{u,n}$ accumulate at their right endpoints. Note that $Stair$ is contained in the short green intervals between the intervals blocked by the $J_{B_n^U}$

The limit points $a^U_{\bullet,n,\infty}$ form increasing unbounded sequences in $(6, \infty)$, while the corresponding b-values lie in $(5/11, 1)$, where $5/11 = \mathrm{acc}_U^{-1}(6)$.

Figure 4 depicts the image under the parameterization $b \mapsto \big(\mathrm{acc}(b), V_b(\mathrm{acc}(b))\big)$ of the intervals $J_{B_n^U}$ for $n = 0, \ldots, 3$ and part of the interval $J_{B_4^U}$.

Remark 57 Theorem 56 lists descending staircases $\mathcal{S}^U_{u,n}$ for all $n \geq 0$, but ascending staircases $\mathcal{S}^U_{\ell,n}$ only for $n \geq 1$. There is an ascending staircase for $n = 0$ with numerics obtained by setting $n = 0$ in the sequence $\mathcal{S}^U_{\ell,n}$ above, and depicted in Fig. 2, but because the class $\mathbf{B}^U_0$ has center 6, it has limit point $a^U_{\ell,0,\infty} < 6$, and therefore, as we will explain in Remark 61 (iii) below, it is better to consider this staircase as part of a different family. Indeed. as pointed out in Corollary 60 (ii), the accumulation points $[6; 1, \{5, 1\}^\infty]$ for $\mathcal{S}^U_{\ell,0}$ and $[7; \{5, 1\}^\infty]$ for $\mathcal{S}^U_{u,0}$ are mutual images under the reflection Φ of Lemma 59; and we explain in Remark 61 how we expect that the associated families of staircases are organized. ♢

Finally, here is a sharper version of Theorems 2 and 5. Notice that the centers $6 - \frac{1}{2n+6}$ of the classes $\mathbf{B}^E_n$ increase with n.

Theorem 58 *The classes $\mathbf{B}^E_n = \big(5(n+3), n+4; (2n+6)\mathbf{w}(\frac{12n+35}{2n+6})\big)$, $n \geq 0$, with increasing centers, are perfect and center-blocking, and have the following associated staircases $\mathcal{S}^E_{\ell,n}, \mathcal{S}^E_{u,n}$, where σ_n and end_n are as in Theorem 54.*

- *for each* $n \geq 1$, $\mathcal{S}^E_{\ell,n}$ *is ascending, with limit point* $a^E_{\ell,n,\infty} = [5; 1, 2n+4, 2n+1, \{2n+5, 2n+1\}^\infty]$, *and has*

$$\begin{array}{ll} \text{(Centers)} & [5; 1, 2n+4, 2n+1, \{2n+5, 2n+1\}^k, \text{end}_n], \\ \text{(Recursion)} & x_{n,k+1} = (\sigma_n + 2)x_{n,k} - x_{n,k-1}, \\ \text{(Relation)} & (2n+3)d_{n,k} = (n+2)\,p_{n,k} - (n+4)q_{n,k} \end{array}$$

- *for each* $n \geq 0$, $\mathcal{S}^E_{u,n}$ *is descending, with limit point* $a^E_{u,n,\infty} = [5; 1, 2n+6, \{2n+5, 2n+1\}^\infty]$, *and has*

$$\begin{array}{ll} \text{(Centers)} & [5; 1, 2n+6, \{2n+5, 2n+1\}^k, \text{end}_n], \\ \text{(Recursion)} & x_{n,k+1} = (\sigma_n + 2)x_{n,k} - x_{n,k-1}, \\ \text{(Relation)} & (2n+3)d_{n,k} = -(n+4)\,p_{n,k} + (11n+31)q_{n,k}. \end{array}$$

The limit points $a^E_{\bullet,n,\infty}$ *form an increasing sequence in the interval* $\big(35/6 = [6; 1, 5], 6\big)$ *while the corresponding b-values lie in* $\big(1/5, 19/61\big)$ *where* $19/61 = \text{acc}_L^{-1}(35/6) < 1/3$.

The numerics of these three families are very similar, since they all contain repetitions of the same 2-periodic pair $(2n+5, 2n+1)$, which as is shown in Lemma 63, implies that the recursion is the same. We now observe that there are symmetries of the z-axis that relate the centers of the different families of blocking classes, as well as those of the corresponding staircase classes, as follows.

In the following we denote by $CF(x)$ the continued fraction expansion of a rational number $x \geq 1$, and, by abuse of language, we sometimes identify z with $CF(z)$. For example, every[10] rational $z \in (6, 7)$ has continued fraction of the form $z = CF(z) = [6; k, CF(x)]$ for some integer $k \geq 0$ and rational number $x \geq 1$.

Lemma 59

(i) *The fractional linear transformation*

$$\Psi : (6, \infty) \to (6, \infty) : \quad w \mapsto \frac{6w - 35}{w - 6}$$

reverses orientation and has the following properties:

(a) $\Psi \circ \Psi = 1, \qquad \Psi(7) = 7;$

(b) $\Psi([6; k, CF(x)]) = [6+k; CF(x)] \in (7, \infty)$ *for all* $z = [6; k, CF(x)] \in (6, 7)$*;*

(ii) *The fractional linear transformation*

[10] This is true even for numbers such as $13/2 = [6; 2]$ since in this case we can also write $z = [6; 1, 1]$.

$$\Phi : (\frac{35}{6}, \infty) \to (\frac{35}{6}, \infty) : \quad w \mapsto \frac{35w - 204}{6w - 35}$$

reverses orientation and has the following properties:

(a) $\Phi \circ \Phi = 1$, $\Phi(6) = 6$ *and*

$$\Phi(7) = \frac{41}{7} = [5; 1, 6], \qquad \Phi(8) = \frac{76}{13} = [5; 1, 5, 2].$$

(b) *every* $z \in \left(\frac{76}{13}, \frac{41}{7}\right)$ *has* $CF(z) = [5; 1, 5, 1, CF(x)]$ *for* $x > 1$ *and*

$$\Phi([5; 1, 5, 1, CF(x)]) = [7; CF(x)] \in (7, 8), \quad \textit{for all } x > 1.$$

(iii) *The bijective map*

$$Sh : (1, \infty) \to (5, 6), \quad z \mapsto \frac{6z - 1}{z}$$

preserves orientation, and for each rational $x > 1$ *and* $k \geq 1$ *we have*

$$Sh(z) = Sh([k + 5; CF(x)]) = [5; 1, k + 4, CF(x)], \quad \textit{if } z = [k + 5; CF(x)] \in (6, \infty).$$

Further, $\Phi \circ \Psi = Sh : (6, \infty) \to (35/6, 6)$.

Proof We prove (i). (a) follows immediately from the definition. To prove (b), observe that

$$w = [6; k + 1, CF(x)] = 6 + \frac{1}{k + 1 + \frac{1}{x}} = \frac{(6k + 7)x + 6}{(k + 1)x + 1}$$

which gives

$$x = \frac{w - 6}{(6k + 7) - (k + 1)w}$$

Therefore we have

$$\begin{aligned} z = [k + 7; CF(x)] &= k + 7 + \frac{1}{x} \\ &= k + 7 + \frac{(6k + 7) - (k + 1)w}{w - 6} \\ &= \frac{6w - 35}{w - 6} = \Psi(w) \end{aligned}$$

as claimed. This proves (i).

The proofs of (ii), (iii) are very similar and are left to the reader.

Corollary 60

(i) *For each $n \geq 1$, the involution Ψ takes the center $2n+6$ of the blocking class $\mathbf{B}_n^U$ to the center $[6; 2n]$ of the blocking class $\mathbf{B}_n^L$. It also takes the centers of the associated staircases $\mathcal{S}_{\ell,n}^U, \mathcal{S}_{u,n}^U$ to those of $\mathcal{S}_{u,n}^L, \mathcal{S}_{\ell,n}^L$, interchanging increasing and decreasing staircases. Further, Ψ also takes the centers of the decreasing staircase $\mathcal{S}_{u,0}^U$ to the centers of the Fibonacci stairs.*

(ii) *The involution Φ fixes the center 6 of the blocking class $\mathbf{B}_0^U$ and interchanges the steps of the two associated staircases.*

(iii) *For each $n \geq 1$, the shift Sh takes the center $2n+6$ of the blocking class $\mathbf{B}_n^U$ to the center $[5; 1, 2n+5]$ of the blocking class $\mathbf{B}_n^E$. It also takes the centers of the associated staircases $\mathcal{S}_{\ell,n}^U, \mathcal{S}_{u,n}^U$ to those of $\mathcal{S}_{\ell,n}^E, \mathcal{S}_{u,n}^E$, preserving the direction of the staircases.*

Proof This is an immediate consequence of Lemma 59. Note that by Theorem 1 the ascending staircase associated to $\mathbf{B}_0^U$ has steps at $[5; 1, \{5, 1\}^k, \text{end}_0]$ while the descending staircase has steps at $[7; \{5, 1\}^{k-1}, \text{end}_0]$. Further, for these values of z, Φ moves the first entry in end_0 from an even place to an odd place, and hence, in accordance with Remark 53, takes an ascending stair to a descending one.

Remark 61

(i) Though they might seem similar, the two involutions Φ and Ψ have very different effects. As shown by part (ii) of Corollary 60, Φ fixes the center of the blocking class $\mathbf{B}_0^U$ and interchanges its two staircases, both of which exist for values of $b > 1/3$. On the other hand, the fixed point of Ψ is not the center of a blocking class, and it takes staircases for $b > 1/3$ to staircases for $b < 1/3$. Similarly, Sh takes staircases for $b > 1/3$ with centers in $(6, \infty)$ to staircases for $b < 1/3$ with centers in $(35/6, 6)$. However, the transformation $\text{Sh}^2 := Sh \circ Sh$ should take staircases for $b > 1/3$ (resp. $b < 1/3$) to staircases for b in the same interval.

(ii) The maps $\Phi, \Psi, Sh = \Phi \circ \Psi$ generate a set of fractional linear transformations that act on the centers of the blocking classes and of the staircase steps. Corollary 60 shows that there has to be some associated action on the other two entries d, m of a perfect blocking class. However, as yet this is not fully understood.

(iii) We conjecture that every 2-periodic staircase for $b \in (1/3, 1)$ is one of the following:

- the staircases $\mathcal{S}_{\bullet,n}^U$ in Theorem 56 with centers in $(6, \infty)$;
- their images under Φ with centers in $(35/6, 6)$;
- the images of these two families under $Sh^{2i}, i = 1, 2 \ldots$, with centers in the disjoint intervals $Sh^2\big((35/6, 6)\big) = (1189/204, 35/6)$, $Sh^4\big((35/6, 6)\big)$, and so on.

Similarly, the 2-periodic staircase for $b \in [0, 1/3)$ should be formed from the staircases $\Psi\big(\mathcal{S}_{\bullet,n}^U\big)$ with centers in $(6, 7)$ by involutions and shifts. This idea,

together with many other related results, will be developed in our next paper. We have also found many potential staircases of period $2m$, $m > 1$, and have conjectural description of them as well. In these staircases the repeated portion of the continued fraction expansion of the centers p_k/q_k has length $2m$, and turns out to consist of a sequence of integers of the form $(2n+1, 2n+3, \ldots, 2n+4m-1)$ in some order. Thus these staircases are still in some way derived from the Fibonacci staircases. ◊

3.2 *Proof of Theorems 56 and 1*

Here is the strategy for proving Theorems 54, 56 and 58.

- **Step 1:** Check that the classes $\mathbf{B}_n$ are quasi-perfect.
- **Step 2:** Prove that the classes in the two associated families $\mathcal{S}_\ell, \mathcal{S}_u$ are perfect.
- **Step 3:** Check that the conditions in Theorem 52 hold, and conclude that each $\mathbf{B}_n$ is a perfect blocking class.
- **Step 4:** Check that each pre-staircase $\mathcal{S}_\ell, \mathcal{S}_u$ satisfies one of the conditions in Theorem 51, and conclude that both are staircases with the given properties.

We will begin by proving Theorem 56, since this is the most straightforward. As we will see, in the cases at hand only Steps 2 and 4 require significant argument. The most computational part of the argument is the proof that the pre-staircase classes are perfect, rather than just quasi-perfect, which involves showing that they reduce correctly under Cremona moves. These proofs are all given in Sect. 3.4. As we will see there, the proofs for the different pre-staircases are closely related.

Remark 62 The arguments given below reduce many of the proofs to checking some numerical identities that are polynomial in n and of small degree d (less than 10 or so). Therefore if one verifies these identities by computer for $d+1$ values of n, then they will hold for all n. ◊

Step 1: For $n \geq 0$, the class

$$\mathbf{B}_n = \left(n+3, n+2; 1^{\times(2n+6)}\right)$$

has $d = n+3$, $m = n+2$, $p = 2n+6$, $q = 1$ and evidently satisfies the Diophantine conditions $3d = m + p + q$, $d^2 - m^2 = pq - 1$. Hence it is quasi-perfect with center $2n+6$.

Step 2: We begin by establishing the following lemma.

Lemma 63

(i) *For $k \geq 1$, let $[2n+5; 2n+1, \{2n+5, 2n+1\}^{k-1}, 2n+4]$ have (integral) weight expansion*

$$\left(a_{n,k}^{\times(2n+5)}, b_{n,k}^{\times(2n+1)}, (a_{n,k} - (2n+1)b_{n,k})^{\times(2n+5)}, \ldots, 1^{\times(2n+4)}\right)$$

and define

$$a_{n,0} = 1, \quad b_{n,0} = -1, \quad a_{n,-1} = 2n + 2. \tag{3.2.1}$$

Then

(a) $b_{n,1} = 2n + 4$, $a_{n,1} = (2n + 1)(2n + 4) + 1$.
(b) *we have*

$$b_{n,k+1} = (2n + 5)a_{n,k} + b_{n,k}, \quad a_{n,k} = (2n + 1)b_{n,k} + a_{n,k-1} \qquad \text{if } k \geq 0. \tag{3.2.2}$$

(c) *For each n, the quantities* $a_{n,\bullet}$, $b_{n,\bullet}$ *satisfy the recursion relation*

$$x_{n,k+1} = (4n^2 + 12n + 7)x_{n,k} - x_{n,k-1}, \quad k \geq 1. \tag{3.2.3}$$

(ii) *For* $k \geq 1$, *let* $[2n + 5; 2n + 1, \{2n + 5, 2n + 1\}^{k-1}, 2n + 5, 2n + 2]$ *have (integral) weight expansion*

$$\big(a'^{\times(2n+5)}_{n,k}, b'^{\times(2n+1)}_{n,k}, (a'_{n,k} - (2n + 1)b'_{n,k})^{\times(2n+5)}, \ldots, (2n + 2)^{\times(2n+5)}, 1^{\times(2n+2)}\big).$$

and define

$$a'_{n,0} = 2n + 2, \quad b'_{n,0} = 1, \quad a'_{n,-1} = 1 \tag{3.2.4}$$

Then

$$b'_{n,1} = (2n+2)(2n+5)+1, \qquad a'_{n,1} = (2n+1)\big((2n+2)(2n+5)+1\big)+(2n+2);$$

and (3.2.2), (3.2.3) *hold for* $a'_{n,\bullet}$, $b'_{n,\bullet}$ *as in* (i).

Proof First we consider case (i). As explained in Example 8, the integral weight expansion $(x_0^{\times\ell_0}, x_1^{\times\ell_1}, \ldots, x_N^{\times\ell_N})$ of $p/q = [\ell_0; \ell_1, \ldots, \ell_N]$ satisfies the identities $x_{k+1} = x_{k-1} - \ell_k x_k$. Assuming that one knows the multiplicities $[\ell_0; \ell_1, \ldots, \ell_N]$, one can read this identity in either direction. If one knows p, q then $x_0 = q$ so that $x_1 = p - \ell_0 q$, and then one can determine x_2, x_3, and so on. On the other hand, since we are dealing with the integral expansion the last weight is 1, and hence we have $x_N = 1$, $x_{N-1} = \ell_N$ which determines x_{N-2}, x_{N-3} and so on. The latter is the relevant procedure here since going from $a_{n,k}, b_{n,k}$ to $a_{n,k+1}, b_{n,k+1}$ means adding an extra pair $(2n + 5, 2n + 1)$ to the continued fraction. For example, the integral weight expansion of $[2n + 5; 2n + 1, 2n + 4]$ is

$$(a_{n,1}^{\times(2n+5)}, b_{n,1}^{\times(2n+1)}, 1^{\times(2n+4)}).$$

This proves (i:a), and (i:b) holds by similar reasoning. One can check that the relation also holds when $k = 0$ with the given definitions of $a_{n,0}$, $b_{n,0}$ and $a_{n,-1}$.

Finally, we prove by induction on $k \geq 1$ that the numbers $a_{n,k}, b_{n,k+1}$ satisfy (3.2.3). The base case for $a_{n,1}$ follows from the definitions of $a_{n,0}, a_{n,-1}$ in (2.4.1), while the claim for $b_{n,2}$ holds because

$$\begin{aligned} b_{n,2} &= (2n+5)\big((2n+1)(2n+4)+1\big) + (2n+4) \quad \text{by (b)} \\ &= (2n+4)(4n^2+12n+6) + (2n+5) \\ &= (4n^2+12n+7)(2n+4) + 1 = (4n^2+12n+7)b_{n,1} - b_{n,0}. \end{aligned}$$

The inductive step then follows from the linear relations $b_{n,k+1} = (2n+5)a_{n,k} + b_{n,k}$ and $a_{n,k+1} = (2n+1)b_{n,k+1} + a_{n,k}$ which were proved in (i)(b). This proves (i).

The proof of (ii) is similar, and is left to the reader.

Lemma 64 *Let $\mathcal{S}$ be a pre-staircase with linear relation $(2n+3)d_k = R_1 p_k + R_2 q_k$, where $R_1, R_2 \in \mathbb{Z}[n]$, the ring of polynomials in n with integer coefficients, and with steps at the points*

$$p_k/q_k = [\ell_0; \ell_1, \dots, \ell_r, \{2n+5, 2n+1\}^k, \mathrm{end}_n], \quad \mathrm{end}_n := 2n+4, \ell_r \in \mathbb{Z}[n],$$

and other coefficients d_k, m_k. Further, let $\mathcal{S}'$ be the corresponding pre-staircase with the same linear relation $(2n+3)d'_k = R_1 p'_k + R_2 q'_k$ and steps p'_k/q'_k defined as above but with $\mathrm{end}_n = (2n+5, 2n+2)$. Then

(i) *$p_k, q_k \in \mathbb{Z}[n]$, and $(2n+3)$ divides $R_1 p_k + R_2 q_k$ if and only if it divides $R_1 p'_k + R_2 q'_k$. Moreover, this holds for all k if and only if it holds for $k = 0$.*
(ii) *$m_1 d_0 - m_0 d_1 = m'_1 d'_0 - m'_0 d'_1$.*

Proof By Lemma 63 (i)

$$q_k \mathbf{w}(p_k/q_k) = \big(\dots,\ (\ell_r b_{n,k+1} + a_{n,k}^{\times \ell_{r-1}},\ b_{n,k+1}^{\times \ell_r}, a_{n,k}^{\times 2n+5}, \dots\big).$$

Therefore, by induction on r, one can see that there are polynomials $C_1, \dots, C_4$ that depend only $\ell_0, \dots, \ell_r$ such that

$$q_k = C_1 b_{n,k+1} + C_2 a_{n,k}, \quad p_k = C_3 b_{n,k+1} + C_4 a_{n,k}.$$

Similarly, we have

$$q'_k = C_1 b'_{n,k+1} + C_2 a'_{n,k}, \quad p'_k = C_3 b'_{n,k+1} + C_4 a'_{n,k},$$

with the same C_i. Therefore, there are elements $N_5, N_6 \in \mathbb{Z}[n]$ such that

$$R_1 p_k + R_2 q_k = N_5 b_{n,k+1} + N_6 a_{n,k}, \quad R_1 p'_k + R_2 q'_k = N_5 b'_{n,k+1} + N_6 a'_{n,k}$$

Next note that to prove (i), it suffices to check that it holds for $k = 0, 1$ since the general case then follows by the recursion. Hence to prove (i) we must check that

$(2n + 3)$ divides $N_5 b_{n,k+1} + N_6 a_{n,k}$ for $k = 0, 1$ if and only if it also divides $N_5 b'_{n,k+1} + N_6 a'_{n,k}$ for $k = 0, 1$.

But by Lemma 63, modulo $(2n + 3)$ we have

$$a_{n,0} \equiv 1, b_{n,1} \equiv 1, \quad a_{n,1} \equiv -1, b_{n,2} \equiv -1, \qquad a'_{n,0} \equiv -1, \ b'_{n,1} \equiv -1, \ a'_{n,1} \equiv 1, b'_{n,2} \equiv 1.$$

Hence in both cases we need precisely that $N_5 + N_6$ is divisible by $2n + 3$. Since this identity follows already from the case $k = 0$, this proves (i).

To prove (ii), notice that given p_k, q_k we define d_k by the staircase relation. Therefore we may write $d_k = C_7 a_{n,k} + C_8 b_{n,k+1}$, where C_7, C_8 are (possibly rational, but in practice polynomial) functions of n that depend also on $\ell_0, \ldots, \ell_r$. When $k = 0, 1$, (3.2.2) gives an explicit linear expression for $b_{n,k+1}$ in terms of a_n, b_n, with coefficients that also depend on n (but are independent of k). Therefore, for $k = 0, 1$ we can write $d_k = N_1 a_{n,k} + N_2 b_{n,k}$ where N_1, N_2 are functions of $n, \ell_0, \ldots, \ell_r$. Since $m_k := 3d_k - p_k - q_k$, a similar argument shows that there are functions $N_1, \ldots, N_4$ of $n, \ell_0, \ldots, \ell_r$ such that the first two lines below hold.

$$\begin{aligned}
d_0 &= N_1 a_{n,0} + N_2 b_{n,0}, \quad & m_0 &= N_3 a_{n,0} + N_4 b_{n,0} \\
d_1 &= N_1 a_{n,1} + N_2 b_{n,1}, \quad & m_1 &= N_3 a_{n,1} + N_4 b_{n,1} \\
d'_0 &= N_1 a'_{n,0} + N_2 b'_{n,0}, \quad & m'_0 &= N_3 a'_{n,0} + N_4 b'_{n,0} \\
d'_1 &= N_1 a'_{n,1} + N_2 b'_{n,1}, \quad & m'_1 &= N_3 a'_{n,1} + N_4 b'_{n,1}.
\end{aligned}$$

The second two lines then hold with the same constants since (3.2.2) also holds in this case. Hence

$$\begin{aligned}
m_1 d_0 - m_0 d_1 &= (N_2 N_3 - N_1 N_4)(a_{n,1} b_{n,0} - a_{n,0} b_{n,1}) \\
m'_1 d'_0 - m'_0 d'_1 &= (N_2 N_3 - N_1 N_4)(a'_{n,1} b'_{n,0} - a'_{n,0} b'_{n,1}).
\end{aligned}$$

Therefore it suffices to check that

$$a_{n,1} b_{n,0} - a_{n,0} b_{n,1} = a'_{n,1} b'_{n,0} - a'_{n,0} b'_{n,1}.$$

But by Lemma 63 both sides equal $-(4n^2 + 12n + 9)$.

Lemma 65 *Suppose that the positive integers d_k, m_k, p_k, q_k are defined for $k \geq 0$ and satisfy the recursion*

$$x_{k+1} = R x_k - x_{k-1}, \quad k \geq 1.$$

Then the identity $d_k^2 - m_k^2 = p_k q_k - 1$ holds for all k, if and only if it holds for $k = 0, 1$, and also

$$2(d_0 d_1 - m_0 m_1) = (p_1 q_0 + p_0 q_1) - R. \tag{3.2.5}$$

Hence it holds for all k if and only if it holds for $k = 0, 1, 2$.

Proof Abbreviate $x := x_k$, $x' := x_{k-1}$ so that $x_{k+1} = Rx - x'$. Then observe that

$$\Big(d^2 - m^2 = pq - 1, \quad (d')^2 - (m')^2 = p'q' - 1\Big) \Longleftrightarrow \tag{3.2.6}$$
$$\Big((Rd - d')^2 - (Rm - m')^2 = (Rp - p')(Rq - q') - 1\Big),$$

exactly if

$$2(dd' - mm') = (pq' + p'q) - R,$$

But if the latter equation holds, we have

$$\begin{aligned} 2\big((Rd - d')d - (Rm - m')m\big) &= 2R(d^2 - m^2) - 2(dd' - mm') \\ &= 2R(pq - 1) + R - (pq' + p'q) \\ &= -R + (Rp - p')q + (Rq - q')p. \end{aligned}$$

Hence, arguing inductively, we see that (3.2.5) implies

$$2(d_{k+1}d_k - m_{k+1}m_k) = -R + (p_kq_{k+1} + p_{k+1}q_k), \qquad \text{for all } k \geq 0.$$

Thus, the calculation in (3.2.6) implies that $d_k^2 - m_k^2 = p_kq_k - 1$ holds for all $k \geq 0$.

Lemma 66

(i) *For each $n \geq 1$, the classes in the ascending pre-staircase $\mathcal{S}^U_{\ell,n}$ are integral and satisfy the linear Diophantine identity.*
(ii) *For each $n \geq 0$, the classes in the descending pre-staircase $\mathcal{S}^U_{u,n}$ are integral and satisfy the linear Diophantine identity.*

Proof First consider case (i) with ending $\mathrm{end}_n = 2n + 4$. Then $\mathbf{E}^U_{\ell,n,0}$ has center $2n + 4$, and one can calculate the corresponding degree d by using the staircase relation

$$(2n + 3)d^U_{\ell,n,k} = (n + 1)p^U_{\ell,n,k} + (n + 2)q^U_{\ell,n,k}.$$

When $k = 1$ we may use Lemma 63 (i) to obtain

$$q^U_{\ell,n,1} = a_{n,1}, \qquad p^U_{\ell,n,1} = (2n + 5)a_{n,1} + b_{n,1}.$$

We compute $d^U_{\ell,n,1}$ using the relation, and then define $m^U_{\ell,n,1}$ by the linear Diophantine identity. This gives the following:

$$\big(d^U_{\ell,n,0}, m^U_{\ell,n,0}, p^U_{\ell,n,0}, q^U_{\ell,n,0}\big) = \big(n + 2, n + 1, 2n + 4, 1\big) \tag{3.2.7}$$

$$d^U_{\ell,n,1} = 4n^3 + 20n^2 + 30n + 13, \quad m^U_{\ell,n,1} = 4n^3 + 16n^2 + 18n + 5.$$
$$p^U_{\ell,n,1} = 8n^3 + 40n^2 + 62n + 29, \quad q^U_{\ell,n,1} = 4n^2 + 10n + 5,$$

The tuples for $k \geq 2$ are then defined using the recursion. Since $d^U_{\ell,n,k}$ is an integer for $k = 0, 1$, the recursion implies that it is an integer for all k. The tuple satisfies the linear Diophantine relation by definition.

The claims when $\mathrm{end}_n = (2n + 5, 2n + 2)$ follow by arguing similarly, using Lemma 64 (i). This proves (i).

The steps of $\mathcal{S}^U_{u,n}$ have centers $[2n + 7; \{2n + 5, 2n + 1\}^k, \mathrm{end}_n]$ for $k \geq 0$, and the relation is

$$(2n + 3)d^U_{u,n,k} = (n + 2)\,p^U_{u,n,k} - (n + 4)q^U_{u,n,k}.$$

Since the staircase $\mathcal{S}^U_{\ell,n}$ has steps with centers

$$[\{2n + 5, 2n + 1\}^k, \mathrm{end}_n] = q^U_{\ell,n,k}\ \mathbf{w}\Big(\frac{p^U_{\ell,n,k}}{q^U_{\ell,n,k}}\Big),$$

it follows that[11]

$$q^U_{u,n,k} = p^U_{\ell,n,k}, \quad p^U_{u,n,k} = (2n + 7)\,p^U_{\ell,n,k} + q^U_{\ell,n,k}.$$

Therefore

$$\begin{aligned} R_1 p^U_{u,n,k} + R_2 q^U_{u,n,k} &= (n + 2)\big((2n + 7)\,p^U_{\ell,n,k} + q^U_{\ell,n,k}\big) - (n + 4)\,p^U_{\ell,n,k} \\ &\equiv (n + 1)\,p^U_{\ell,n,k} + (n + 2)q^U_{\ell,n,k} \pmod{(2n + 3)}, \end{aligned}$$

and hence is divisible by $2n + 3$ by (i). Again, the linear Diophantine relation holds by the definition of $m^U_{u,n,k}$. For later use, we note that

$$\begin{aligned} \big(d^U_{u,n,0}, m^U_{u,n,0}\big) &= \big(2n^2 + 11n + 14,\ \ 2n^2 + 9n + 9\big) \\ d^U_{u,n,1} &= 8n^4 + 68n^3 + 202n^2 + 245n + 100, \\ m^U_{u,n,1} &= 8n^4 + 60n^3 + 158n^2 + 171n + 63. \end{aligned} \tag{3.2.8}$$

This completes the proof.

To complete Step 2 we must show that the classes are perfect. This proof is deferred to Sect. 3.4; see Proposition 79. The quadratic Diophantine equality

[11] In general, if $[\ell_0; \ell_1, \ldots,] = p/q$ then $[s; \ell_0, \ell_1, \ldots,] = (sp + q)/p$.

then follows (though one could also prove it by an inductive argument based on Lemma 65).

Step 3: We must check that the conditions in Theorem 52 hold. The fact that $3 + 2\sqrt{2} < P_\ell/Q_\ell < p/q < P_u/Q_u$ follows immediately from the continued fraction expansions of P_ℓ/Q_ℓ, P_u/Q_u. See Remark 76, which explains the somewhat subtle ordering of numbers that are given by continued fractions, and Remark 53 which explains how to check whether a staircase ascends or descends. The claim that $\ell(p/q) < \ell(z)$ for all $z \in \left(P_\ell/Q_\ell, P_u/Q_u\right)$ is very straightforward in the case of the staircases $\mathcal{S}^U_{\bullet,n}$, since p/q is an integer and $\ell(p/q) = p$. We also must check that $M_\bullet/D_\bullet \neq 1/3$ for $\bullet = \ell, u$. Since $M_\bullet$, $D_\bullet$ are both numbers of the form $a + b\sqrt{\sigma}$ where $a, b \in \mathbb{Q}$ it suffices to check that rational part a of $3M_\bullet$ is not equal to that of $D_\bullet$. Therefore, by (2.4.4) it suffices to check that in each staircase $3m_0 \neq d_0$, a fact that is immediate from the formulas in (3.2.7) and (3.2.8). Thus condition (i) in Theorem 52 holds, and conditions (ii) and (iii) can be verified by comparing the given linear relations with the parameters of $\mathbf{B}^U_n$.

Step 4: By Step 2 and Sect. 3.4 we know that for each n the pre-staircase classes are perfect for large k, but it remains to check that they are live at the appropriate limiting value for b. This requires considerable work, and again we begin with a general result.

Lemma 67 *Consider a pre-staircase with classes* $\left(d_k, m_k; q_k \mathbf{w}(p_k/q_k)\right)$, *where the ratios* $b_k := m_k/d_k$ *have limit* b_∞, *and let the constants* $D, D', D'', M, M', M'', \sigma := \sigma_n$ *be as in* (2.4.4), *with* $x_k = d_k, m_k$ *respectively.*

(i) *Suppose that* $M\overline{D} - \overline{M}D \neq 0$. *Then the* b_k *are strictly increasing iff*

$$M\overline{D} - \overline{M}D = 2\sqrt{\sigma_n}(M''D' - M'D'') > 0,$$

and otherwise they are strictly decreasing.

(ii) *if* $M''D' - M'D'' > 0$, $b_\infty < r/s \leq 1$, *and*

$$|M''D' - M'D''| \leq \frac{sD - rM}{2\sqrt{\sigma}\,|rD - sM|}, \tag{3.2.9}$$

then there is k_0 *such that*

$$\frac{m_k}{d_k} \leq b_\infty = \frac{M}{D} \leq \frac{s + m_k(rd_k - sm_k)}{r + d_k(rd_k - sm_k)}, \quad \text{for} \quad k \geq k_0.$$

(iii) *if* $M''D' - M'D'' < 0$, $b_\infty > r/s > 0$, *and* (3.2.9) *holds, then there is* k_0 *such that*

$$b_k := \frac{m_k(sm_k - rd_k) - s}{d_k(sm_k - rd_k) - r} \leq b_\infty = \frac{M}{D} \leq \frac{m_k}{d_k}, \quad \text{for} \quad k \geq k_0,$$

Proof Write

$$m_k = M\lambda^k + \overline{M}\,\overline{\lambda}^k, \quad d_k = D\lambda^k + \overline{D}\,\overline{\lambda}^k,$$

as in Lemma 47. Then

$$b_k := \frac{m_k}{d_k} = \frac{M\lambda^k + \overline{M}\,\overline{\lambda}^k}{D\lambda^k + \overline{D}\,\overline{\lambda}^k} < \frac{M\lambda^{k+1} + \overline{M}\,\overline{\lambda}^{k+1}}{D\lambda^{k+1} + \overline{D}\,\overline{\lambda}^{k+1}} = \frac{m_{k+1}}{d_{k+1}}$$

if and only if

$$\big(M\lambda^k + \overline{M}\,\overline{\lambda}^k\big)\big(D\lambda^{k+1} + \overline{D}\,\overline{\lambda}^{k+1}\big) < \big(M\lambda^{k+1} + \overline{M}\,\overline{\lambda}^{k+1}\big)\big(D\lambda^k + \overline{D}\,\overline{\lambda}^k\big).$$

This is equivalent to

$$\overline{M}D\lambda + M\overline{D}\,\overline{\lambda} \;<\; \overline{M}D\overline{\lambda} + M\overline{D}\lambda,$$

and hence to the condition

$$(M\overline{D} - \overline{M}D)\lambda \;>\; (M\overline{D} - \overline{M}D)\overline{\lambda}.$$

Since $\lambda > \overline{\lambda} > 0$, this holds exactly if $M\overline{D} - \overline{M}D > 0$.

Now write $M = M' + M''\sqrt{\sigma}$, $D = D' + D''\sqrt{\sigma}$ as in Remark 48. Then

$$\begin{aligned} M\overline{D} - \overline{M}D &= (M' + M''\sqrt{\sigma})(D' - D''\sqrt{\sigma}) - (M' - M''\sqrt{\sigma})(D' + D''\sqrt{\sigma}) \\ &= 2\sqrt{\sigma}(M''D' - M'D''). \end{aligned}$$

This proves (i).

To prove (ii) we must check that

$$\begin{aligned} &rM + M(D\lambda^k + \overline{D}\,\overline{\lambda}^k)\big(r(D\lambda^k + \overline{D}\,\overline{\lambda}^k) - s(M\lambda^k + \overline{M}\,\overline{\lambda}^k)\big) \\ &\qquad \le sD + D(M\lambda^k + \overline{M}\,\overline{\lambda}^k)\big(r(D\lambda^k + \overline{D}\,\overline{\lambda}^k) - s(M\lambda^k + \overline{M}\,\overline{\lambda}^k)\big). \end{aligned}$$

The coefficients of λ^{2k} are the same on both sides, while, after a little manipulation, the constant term gives the inequality

$$(M\overline{D} - \overline{M}D)(rD - sM) \le sD - rM.$$

Since $b_\infty := M/D < r/s \le 1$ by assumption, we have $rD - sM > 0$, and $sD - rM > 0$ since $D > M$, $s \ge r$. The further assumption

$$M\overline{D} - \overline{M}D = 2\sqrt{\sigma}(M''D' - M'D'') > 0,$$

shows that the above inequality is equivalent to requiring that $|M''D' - M'D''| \leq \frac{sD-rM}{2\sqrt{\sigma}(sM-rD)}$, as claimed.

This completes the proof of (ii). The proof of (iii) is almost identical, and is left to the reader.

Remark 68

(i) The inequalities in Lemma 67 can be simplified because, in the notation of (2.4.4) we have $X = X' + X''\sqrt{\sigma_n}$ where

$$X' = \frac{x_0}{2}, \quad X'' = \frac{2x_1 - x_0(\sigma_n + 2)}{2(2n+3)\sigma_n}, \quad \sigma_n = (2n+1)(2n+5). \tag{3.2.10}$$

Hence if the values of $d_{n,k}, m_{n,k}$ for $k = 0, 1$ are denoted d_k, m_k, we have

$$D' = \frac{d_0}{2}, \quad D'' = \frac{2d_1 - d_0(\sigma_n+2)}{2(2n+3)\sigma_n}, \quad M' = \frac{m_0}{2}, \quad M'' = \frac{2m_1 - m_0(\sigma_n+2)}{2(2n+3)\sigma_n}$$

so that

$$M''D' - M'D'' = \frac{1}{2(2n+3)\sigma_n}\big(m_1d_0 - m_0d_1\big). \tag{3.2.11}$$

Therefore, we just need to check the sign of $m_1d_0 - m_0d_1$, and the following reformulation of (3.2.9):

$$\frac{|m_1d_0 - m_0d_1|}{2n+3} \leq \frac{\sqrt{\sigma_n}(sD - rM)}{|sM - rD|}. \tag{3.2.12}$$

Note that this inequality only involves the entries in the first two terms of the pre-staircase.

(ii) By Lemma 64 (ii), the quantity $m_1d_0 - m_0d_1$ does not depend on the chosen end end_n. Further, the ratio M/D is also independent of the end, since P/Q is by definition, and $M/D = \mathrm{acc}^{-1}(P/Q)$. Therefore the right hand side of the inequality (3.2.11) is also independent of the choice of end. Hence our arguments below apply with both choices.

(iii) We observe that in all cases encountered below the expression $m_1d_0 - m_0d_1$ is a multiple of $2n+3$. Further, the calculations can sometimes be simplified by noting that the sign of $m_1d_0 - m_0d_1$ determines whether m_0/d_0 is greater or less than m_1/d_1. Since the function $x \mapsto \frac{s-rx}{|sx-r|}$ is monotonic, it is sometimes enough to calculate the simpler expression $\frac{sd_0 - rm_0}{|sm_0 - rd_0|}$ instead of $\frac{sD-rM}{|sM-rD|}$.

(iv) If $\mathcal{S} = (\mathbf{E}_k)$ is a pre-staircase whose steps have centers p_k/q_k, then we saw in Proposition 49 that the classes $\mathbf{E}_k$ are perfect blocking classes for large k. The corresponding family of blocked b-intervals $J_{\mathbf{E}_k}$ cannot include the limit b_∞ and if the b values are $> 1/3$ (resp. $< 1/3$) these intervals ascend if and only if the centers ascend (resp. descend). Although, as we saw in Remark 39, the

values $b_k = m_k/d_k$ do not have to lie in these blocked b-intervals, they cannot be too far away. Hence one would expect them to increase for ascending (resp. descending) staircases with b-values $> 1/3$ (resp. $< 1/3$.) However, this does not seem to be true. For example the ascending staircases $\mathcal{S}^U_{\ell,n}$ considered in Lemma 69 have b-values in $(1/3, 1)$ but yet the sequence $b^U_{\ell,n,k}$ decreases. Since this implies that $b^U_{\ell,n,k} > b^U_{\ell,n,\infty}$, this implies that $b^U_{\ell,n,k} \notin J_{\mathbf{E}^U_{n,k}}$. A similar phenomenon happens with the ascending staircases $\mathcal{S}^L_{\ell,n}, \mathcal{S}^E_{\ell,n}$ with b-values in $(0, 1/3)$ and yet increasing $b^L_{\ell,n,k}, b^E_{\ell,n,k}$. On the other hand, the descending staircases do not seem to exhibit this behavior, though in these cases it can require more work to eliminate potential overshadowing classes. ◊

Lemma 69 *The pre-staircase $\mathcal{S}^U_{\ell,n}$ satisfies condition (ii) in Theorem 51 for all $n \geq 1$.*

Proof By Remark 68 (ii) we may suppose that $\mathrm{end}_n = 2n + 4$. Then, using the notation in that remark, we have from Lemma 66 that $d_0 := d^U_{\ell,n,0} = n + 2$, $m_0 := m^U_{\ell,n,0} = n + 1$ and

$$d_1 = 4n^3 + 20n^2 + 30n + 13, \quad m_1 = 4n^3 + 16n^2 + 18n + 5.$$

Hence

$$\begin{aligned} m_1 d_0 - m_0 d_1 &= (4n^3 + 16n^2 + 18n + 5)(n + 2) - (4n^3 + 20n^2 + 30n + 13)(n + 1) \\ &= -(2n + 3). \end{aligned}$$

Therefore, by Lemma 67 (iii), we must check that (3.2.12) holds for suitable r, s, i.e. for each n we need

$$\frac{|m_1 d_0 - m_0 m_1|}{2n + 3} = 1 < \frac{\sqrt{\sigma_n}(sD - rM)}{sM - rD}.$$

But the function $x \mapsto \frac{s - rx}{sx - r}$ decreases for $x \in (r/s, 1]$ with minimum value 1. Hence for each n, this inequality holds for all $r/s < M/D = b^U_{\ell,n,\infty}$.

Finally, we show that there is no overshadowing class of degree $d' < sD/(sM - rD)$. Fix n and abbreviate $b_n := b^U_{\ell,n,\infty}$. We first claim that $b_n > 1/2$ for all $n \geq 1$. Indeed, $a^U_{\ell,1,\infty} = [7; 3; \{7, 3\}^\infty] > 7$ while one can calculate using (2.3.2) that $\mathrm{acc}_U^{-1}(7) \approx 0.596$. Hence, for each n we may choose $r/s > 0$ so that $s/(sb_n - r) < 2$, or equivalently so that $0 < 2r < s(2b_n - 1)$. Thus any overshadowing class would have to have degree 1, and hence does not exist.

It remains to consider the descending pre-staircases $\mathcal{S}^U_{u,n}$. Again, it suffices by Remark 68 (ii) to treat the case $\mathrm{end}_n = 2n + 4$. It is convenient to deal first with the case $n = 0$.

Example 70 When $n = 0$ the classes $\mathbf{E}^U_{u,0,k}$ for $k = 0, 1$ have centers $[7; 4]$ and $[7; 5, 1, 4]$, and hence have parameters

$$(d, m, p, q) = (14, 9, 29, 4), \quad (100, 63, 208, 29) \quad \text{with } \sigma = 5.$$

Since

$$X := X' + X''\sqrt{\sigma} := \frac{x_0}{2} + \frac{2x_1 - x_0(\sigma+2)}{2(2n+3)\sigma}\sqrt{\sigma},$$

we have

$$D = 7 + \frac{17}{5}\sqrt{5} \approx 14.6, \qquad M = \frac{9}{2} + \frac{21}{10}\sqrt{5} \approx 9.2,$$

Hence $b_\infty = M/D \approx 0.6297$, and

$$m_1 d_0 - m_0 d_1 = 63 \cdot 14 - 9 \cdot 100 = -18,$$

so that the sequence $m^U_{u,0,k}/d^U_{u,0,k}$ decreases as expected. Since $b_\infty > 1/2$ we may take $r = 1, s = 2$, and one can check that this choice gives the estimate $\frac{\sqrt{5}(2D-M)}{(2M-D)} > 6$ required by (3.2.12).

Further, since $b_\infty > 5/8$ we have $2/(2b_\infty - 1) < 2/0.25 < 8$, so that we only need rule out overshadowing classes of degree $d' < 8$ with $m' < d'/2$. But if $d' = 7$ this means that $m' \le 3$, so that $d'b_\infty - m' > 1$. Hence, by Lemma 15 (i), such a class cannot be obstructive for b in a nonempty interval of the form $(b_\infty, b_\infty + \varepsilon)$, and so is not overshadowing. A similar argument rules out the cases $d' = 2, 4, 5, 6$, while if $d' = 3$ we must have $m' \le 1$ so that $\mathbf{E}' = (3, 1; 2, 1^{\times 5})$ or $(3, 0; 2, 1^{\times 6})$. But the centers of both these classes are less than the accumulation point a_∞, so that they cannot overshadow.[12] Thus no overshadowing classes satisfy the given restrictions. It follows that $\mathcal{S}^U_{u,0}$ is a staircase, by Theorem 51 (ii). ♢

Lemma 71 *For each $n \ge 0$, the pre-staircase $\mathcal{S}^U_{u,n}$ satisfies condition (ii) in Theorem 51.*

Proof Again we carry out the argument for $\text{end}_n = 2n + 4$, leaving the other case to the reader. Since the case $n = 0$ was treated in Example 70, we will suppose that $n \ge 1$. By (3.2.8) we have

$$\begin{aligned}(d_0, m_0) &= \big(2n^2 + 11n + 14,\ \ 2n^2 + 9n + 9\big) = \big((2n+7)(n+2),\ \ (2n+3)(n+3)\big)\\ d_1 &= 8n^4 + 68n^3 + 202n^2 + 245n + 100\\ &= (4n^2 + 12n + 7)d_0 + 2 = (\sigma_n + 2)d_0 + 2,\end{aligned}$$

[12] Notice that when $z > 6$ the obstruction from $(3, 1; 2, 1^{\times 5})$ is no longer given by $z \mapsto \frac{1+z}{3-b}$.

$$m_1 = 8n^4 + 60n^3 + 158n^2 + 171n + 63 = (\sigma_n + 2)m_0.$$

Therefore

$$m_1 d_0 - m_0 d_1 = -2m_0 < 0,$$

so that by (3.2.11) the b_k decrease, and we are case (iii) of Lemma 67. Further,

$$\frac{M}{D} \le \frac{m_0}{d_0} = \frac{2n^2 + 9n + 9}{2n^2 + 11n + 14} = \frac{1}{1+y}, \quad \text{where } y := \frac{2n+5}{2n^2+9n+9}.$$

Therefore, because the function $x \mapsto \frac{2-x}{2x-1}$ is decreasing for $1/2 < x \le 1$, we have

$$\begin{aligned} \frac{2D - M}{2M - D} &> \frac{2d_0 - m_0}{2m_0 - d_0} = \frac{1+2y}{1-y} \\ &\ge 1 + 3y = \frac{2n^2 + 15n + 24}{(2n+3)(n+3)} > \frac{2(n+4)}{2n+3}. \end{aligned}$$

Therefore, because $2n + 2 < \sqrt{\sigma_n} < 2n + 3$, we have

$$\frac{|m_1 d_0 - m_0 d_1|}{2n+3} = 2(n+3) \le 2(n+1)\frac{2(n+4)}{2n+3} \le \sqrt{\sigma_n}\frac{2D - M}{2M - D}$$

where the first inequality uses $(n + 3)(2n + 3) \le 2(n + 1)(n + 4)$ for all $n \ge 1$.

This shows that condition (3.2.12), or equivalently (3.2.9), holds with $r/s = 1/2$.

When $n = 1$ we can use (3.2.10) and the above formulas for (d_i, m_i) to calculate $b_\infty \approx 0.738 > 0.7$. Hence $2/(2b_\infty - 1) < 5$. Therefore it remains to show there are no overshadowing classes of degree $d' \le 4$ and $m'/d' < 1/2$. But the arguments in Example 70 still apply when $n > 0$. This completes the proof.

Proof of Theorem 56 This follows from Lemmas 66, 69, and 71, and Corollary 84, using the reasoning explained at the beginning of Sect. 3.2.

Proof of Theorem 1 Most of the claims in this theorem are restated in Theorem 56, and hence are proved above. It remains to check the explicit formulas for the endpoints of $J_{\mathbf{B}_n^U}$ and $I_{\mathbf{B}_n^U}$:

$$\begin{aligned} \beta_\ell(n) &= \frac{(2n^2 + 6n + 3) - \sqrt{\sigma_n}}{2n^2 + 6n + 2}, \quad \frac{(n+3)\left(3n + 7 + \sqrt{\sigma_n}\right)}{5n^2 + 30n + 44} = \beta_u(n), \\ \alpha_\ell(n) &= \frac{\sigma_n + (2n+3)\sqrt{\sigma_n}}{2(2n+1)}, \quad 6 + \frac{\sigma_n + (2n+3)\sqrt{\sigma_n}}{2(2n+5)} = \alpha_u(n). \end{aligned}$$

By Theorem 52, $\beta_\ell(n) = M/D$ where M, D are calculated for $\mathcal{S}^U_{\ell,n}$, while $\beta_u(n) = M/D$ where M, D are calculated for $\mathcal{S}^U_{u,n}$. Thus by Lemma 69, and taking $\text{end}_n =$

$2n+4$ we have

$$\beta_\ell(n) = \frac{(n+1)(2n+3)\sqrt{\sigma_n} + (4n^3+16n^2+17n+3)}{(n+2)(2n+3)\sqrt{\sigma_n} + (4n^3+20n^2+29n+12)},$$

and one can check that this does simplify to the desired expression. Similarly, Lemma 71 shows that

$$\beta_u(n) = \frac{(n+3)(2n+3)\big((2n+3)\sqrt{\sigma_n} + (\sigma_n+2)\big)}{(n+2)(2n+7)\big((2n+3)\sqrt{\sigma_n} + (\sigma_n+2)\big)+4},$$

and one can again check that this does simplify to the desired expression.

The closed formulas for $\alpha_\ell(n), \alpha_u(n)$ can be obtained in a similar way from the ratios P/Q, or by simply calculating $\mathrm{acc}(\beta_\ell(n))$, $\mathrm{acc}(\beta_u(n))$. Further details are left to the interested reader.

3.3 *Proof of Theorems 54, 58, 2 and 5*

Since the statement of Theorem 58 includes those of Theorems 2 and 5, it suffices to prove the first two theorems.

We will first consider the staircases associated to the classes

$$\mathbf{B}_n^L = \Big(5n, n-1; 2n\mathbf{w}(\frac{12n+1}{2n})\Big), n \geq 1.$$

(Notice that the case $n=0$ makes no sense here.) We follow the procedure laid out at the beginning of Sect. 3.2. Although we could make the arguments completely afresh, ignoring the connection between the classes $\mathbf{B}_n^U$ and $\mathbf{B}_n^L$, we will use the fact from Corollary 60 that the reflection $\Psi : w \mapsto \frac{6w-35}{w-6}$ takes the centers of the classes in the staircases $\mathcal{S}^U_{\bullet,n}$ to those of $\mathcal{S}^L_{\bullet',n}$, where $\ell' := u, u' := \ell$, i.e. it interchanges the ascending and descending staircases.

We saw in Example 8 that the integers p, q calculated from a continued fraction expansion $[\ell_0; \ldots, \ell_N]$ are always coprime. Further, if $\gcd(p,q) = 1$ then $\gcd(6p-35q, p-6q) = 1$ since any common divisor of $6p-35q$, $p-6q$ also divides $6p-35q-6(p-6q) = q$ and hence also divides p. Since $\Psi^{-1} = \Psi$, this proves the following:

$$\Big(\gcd(p,q)=1,\ \Psi(\frac{p}{q}) = \frac{\widehat{p}}{\widehat{q}}\Big) \Longleftrightarrow \Big(\gcd(\widehat{p},\widehat{q})=1,\ \widehat{p} = 6p-35q, \widehat{q} = p-6q\Big). \tag{3.3.1}$$

To prove Step 1 we simply need to check that the classes $\mathbf{B}_n^L$ satisfy the Diophantine equations (2.1.2), which can be done by an easy computation.

We move on to Step 2, which claims that the staircase classes are perfect. The next lemma shows that they are well defined (i.e. that the degree d is always an integer), and satisfy the linear Diophantine identity. The proof that they are perfect is given in Proposition 86. One could also use Lemma 65 to check the quadratic Diophantine identity; however, since this follows from Proposition 86 we do not do that here. Note that there is an analog of Lemma 72 when $\text{end}_n = (2n+5, 2n+2)$, but in view of Lemma 64 we do not need it.

Lemma 72

(i) *For $n \geq 2$ and $\text{end}_n = 2n+4$, the pre-staircase $\mathcal{S}^L_{u,n}$ can be extended by the class for $k = -1$ with coefficients*

$$\big(5(n-1), (n-2); (2n-2)\mathbf{w}(\frac{12n-11}{2n-2})\big). \tag{3.3.2}$$

(ii) *The classes in $\mathcal{S}^L_{\ell,n}$ and $\mathcal{S}^L_{u,n}$ are well defined and satisfy the linear Diophantine identity for $n \geq 1$ and all k.*

Proof The classes in $\mathcal{S}^L_{u,n}$ have centers

$$\begin{aligned} a^L_{u,n,k} &= [6; 2n-1, 2n+1, \{2n+5, 2n+1\}^k, \text{end}_n], \qquad n \geq 1, k \geq 0, \\ &= \Psi([2n+5; 2n+1, \{2n+5, 2n+1\}^k, \text{end}_n] = \Psi(a^U_{\ell,n,k+1}), \end{aligned}$$

where the second identity holds by Lemma 59 (i:b). Notice here that the term in k for $\mathcal{S}^L_{u,n}$ corresponds to the term in $k+1$ for $\mathcal{S}^U_{\ell,n}$. Therefore, there should be a term in $\mathcal{S}^L_{u,n}$ with center $\Psi(\text{end}_n)$. When $\text{end}_n = 2n+4$, the center is

$$\Psi(2n+4) = \frac{6(2n+4)-35}{2n+4-6} = \frac{12n-11}{2n-2}.$$

Therefore, because the relation in $\mathcal{S}^L_{u,n}$ is

$$(2n+3)d_{n,k} = -(n-1)p_{n,k} + (11n+2)q_{n,k},$$

this entry is $\big(5(n-1), n-2; (2n-2)\mathbf{w}(\frac{12n-11}{2n-2})\big)$. This is a quasi-perfect class, as can be verified by direct calculation. Since the (Relation) holds by definition, to complete the proof that this class can be added to $\mathcal{S}^L_{u,n}$ as in Remark 48 (ii), we must check that the (Recursion) holds for the triple $k = -1, 0, 1$, i.e. that

$$x_1 = (\sigma+2)x_0 - x_{-1}, \qquad x = d, p, q, m.$$

Since $d_\bullet, m_\bullet$ are linear combinations of $p_\bullet, q_\bullet$, it suffices to prove this for $p_\bullet, q_\bullet$. But by definition p_k, q_k are linear functions of p'_k, q'_k, where $p'_k/q'_k = [\{2n+5, 2n+$

$1\}^{k+1}, 2n+4]$, and p'_k, q'_k satisfy the recursion by Lemma 63 (i). This justifies our extending $\mathcal{S}^L_{u,n}$ by this class, and hence proves (i).

Since the linear Diophantine equality is always satisfied by definition of m, to prove (ii) we must check that d as defined by the staircase relation is integral for each n. By Lemma 64 (i) it suffices to do this in the case $\mathrm{end}_n = 2n+4$, and for only two adjacent values of k (either $k = 0, 1$ or $k = -1, 0$). In the case of $\mathcal{S}^L_{u,n}$ with $n \geq 2$, we have the integral vectors

$$\begin{aligned}(d^L_{u,n,-1}, m^L_{u,n,-1}, p^L_{u,n,-1}, q^L_{u,n,-1}) &= \big(5(n-1), n-2, 12n-11, 2n-2\big); \qquad (3.3.3)\\ (d^L_{u,n,0}, m^L_{u,n,0}, p^L_{u,n,0}, q^L_{u,n,0}) &= \big(20n^3+40n^2+6n-1, 4n^3+4n^2-6n-1,\\ &\qquad 48n^3+100n^2+22n-1, 8n^3+16n^2+2n-1\big).\end{aligned}$$

Further, when $n = 1$, although the tuple $(0, -1; 1, 0)$ obtained by setting $n = 1$ in (3.3.2) makes no sense geometrically, it does make numerical sense, and so $d^L_{u,1,k}$ is integral for each k as well.

As for $\mathcal{S}^L_{\ell,n}$ with $\mathrm{end}_n = 2n+4$, one can check that its initial classes are

$$\begin{aligned}(d^L_{\ell,n,0}, m^L_{\ell,n,0}) &= \big(10n^2+25n+13,\ 2n^2+3n\big) \qquad (3.3.4)\\ (d^L_{\ell,n,1}, m^L_{\ell,n,1}) &= \big(40n^4+220n^3+422n^2+331n+89,\ 8n^4+36n^3+50n^2+21n\big).\end{aligned}$$

Therefore $d^L_{\ell,n,k}$ is integral for all k, n as claimed.

Step 3 again follows immediately from the definitions. (For more details about how to check this, see this step in Sect. 3.2; it occurs just after Lemma 66.) Hence it remains to prove the analog of Lemmas 69 and 71. Again it suffices to carry out the argument only for the case $\mathrm{end}_n = 2n+4$ since the other case will follow by Remark 68 (ii).

Lemma 73 *For all $n \geq 0$, the pre-staircase $\mathcal{S}^L_{\ell,n}$ satisfies condition (i) in Theorem 51 with appropriate r/s.*

Proof We first check the sign of $m_1 d_0 - m_0 d_1$. By (3.3.4) we have

$$\begin{aligned}m_1 d_0 - m_0 d_1 &= \big(8n^4+36n^3+50n^2+21n\big)\big(10n^2+25n+13\big)\\ &\qquad - \big(40n^4+220n^3+422n^2+331n+89\big)\big(2n^2+3n\big)\\ &= 4n^2+6n = (2n+3)2n > 0 \qquad \text{since } n > 0,\end{aligned}$$

Therefore we are in case (ii) of Lemma 67, and, by (3.2.12), need to find $r/s > M/D$ so that

$$\frac{\sqrt{\sigma}(sD - rM)}{rD - sM} > 2n, \qquad (3.3.5)$$

where $M := M_\ell(n)$, $D := D_\ell(n)$ are the values appropriate for $\mathcal{S}^L_{\ell,n}$, and $\sigma := \sigma_n$. Since $(2n+2)^2 < \sigma = (2n+1)(2n+5) < (2n+3)^2$, it suffices to have $\frac{sD-rM}{rD-sM} > 1$, or equivalently $(s-r)D > (s-r)M$. But this holds for all $0 < r < s$, showing that (3.3.5) holds for all such r, s. Therefore, for each n we are free to choose $1 > r/s > M/D$ to minimize the degree $d' < sD/(rD - sM)$ of a potentially overshadowing class.

We now claim that $M/D < 1/5$. One could simply prove this by direct calculation. Alternatively, recall that we have already proved that this staircase is associated to the blocking class $\mathbf{B}^L_n$ so that M/D is one of the endpoints (in fact the upper one) of the blocked interval $J_{\mathbf{B}^L_n}$, which as one can readily check is a subset of $(0, 1/5)$. Hence we can choose any $r/s > 1/5$. In particular, if we take $r/s = 7/10$ then

$$\frac{sD}{rD - sM} = \frac{s}{r - sM/D} < \frac{s}{r - s/5} = 2.$$

But the only exceptional classes of degree $d = 1$ are $L - E_i - E_j$, which do not overshadow. Hence there are no overshadowing classes, and the proof is complete.

Lemma 74 *For all $n \geq 1$, the pre-staircase $\mathcal{S}^L_{u,n}$ satisfies condition (i) in Theorem 51.*

Proof As we pointed out in (2.4.5) in cases when we can extend the staircase by a term with $k = -1$ we have $m_0 d_1 - m_1 d_0 = m_{-1} d_0 - m_0 d_{-1}$. Therefore we can apply (3.2.11) using the terms with $k = -1, 0$. By (3.3.2) and (3.3.3) we have

$$\begin{aligned} m_0 d_{-1} - m_{-1} d_0 &= (4n^3 + 4n^2 - 6n - 1)5(n-1) - (n-2)(20n^3 + 40n^2 + 6n - 1) \\ &= 24n^2 + 38n + 3 = (2n+3)(12n+1) > 0. \end{aligned}$$

Hence, by Lemma 67 (i) and (3.2.11), for each n the sequence $(b_{n,k})_{k \geq -1}$ increases, which is what we would expect from a descending staircase with b-values $< 1/3$ (but see Remark 68 (iv)).

Next, we must check to see if (3.2.12) holds with appropriate r/s. Thus we need

$$\frac{m_0 d_{-1} - m_{-1} d_0}{\sqrt{\sigma_n}(2n+3)} = \frac{12n+1}{\sqrt{\sigma_n}} < \frac{s - r\frac{M}{D}}{r - s\frac{M}{D}}. \tag{3.3.6}$$

Since $2n + 2 < \sqrt{\sigma_n}$, we have $\frac{12n+1}{\sqrt{\sigma_n}} < 6$ for all $n \geq 1$. However there are better upper bounds for small n:

$$\frac{13}{\sqrt{\sigma_1}} < 2.84, \quad \frac{25}{\sqrt{\sigma_2}} < 3.73, \quad \frac{37}{\sqrt{\sigma_3}} < 4.22 \tag{3.3.7}$$

Notice also that the values $b_n := b^L_{u,n,\infty} = M/D$ increase with n because $\mathrm{acc}(b_n) = a^L_{u,n,\infty}$ decreases with n (as one can see because the centers of the

associated blocking classes $\mathbf{B}_n^L$ decrease). Further, for each n we have

$$\mathrm{acc}_L^{-1}\Big(\frac{12(n-1)+1}{2(n-1)}\Big) < b_n < \mathrm{acc}_L^{-1}\Big(\frac{12n+1}{2n}\Big),$$

since $\frac{12n+1}{2n}$ is the center point of the blocking class $\mathbf{B}_n^L$ and acc_L^{-1} reverses orientation. Therefore

$$0 < b_1 < \mathrm{acc}_L^{-1}\Big(\frac{13}{2}\Big) < b_2 < \mathrm{acc}_L^{-1}\Big(\frac{25}{4}\Big) < b_3 < \mathrm{acc}_L^{-1}\Big(\frac{37}{6}\Big) < b_4, \tag{3.3.8}$$

where, using (2.3.2) to compute $\mathrm{acc}_L^{-1}(\frac{12n+1}{2n})$ for $n = 1, 2, 3$, we find they are approximately 0.064, 0.121, and 0.144.

It turns out that if we choose $r/s = 3/10$, we can satisfy the inequality $\frac{12n+1}{\sqrt{\sigma_n}} < \frac{s-rb_n}{r-sb_n}$ for all n, while the requirement in Lemma 15 (i) that an obstructive class $\mathbf{E}' = (d', m'; \mathbf{m}')$ at b_n have $|b_n d' - m'| < 1$ has no suitable solutions if $m'/d' > 3/10$ and $d' \le \frac{10}{3-10b_n}$. Thus there is no overshadowing class.

Here are the details. Notice that when $0 \le x < r/s < 1$ the function $x \mapsto \frac{s-rx}{r-sx}$ increases. Therefore $\frac{10-3b}{3-10b} > 3$ for all b, which in view of (3.3.7) gives the bound in (3.3.6) for $n = 1$. In fact, one can check that (3.3.6) holds for all n because, using the bounds (3.3.8) for b_i and the bounds in (3.3.7) for $\frac{12n+1}{\sqrt{\sigma_n}} \le 6$, we have

$$\frac{10-3b_2}{3-10b_2} > 4, \qquad \frac{10-3b_3}{3-10b_3} > 5, \qquad \frac{10-3b_n}{3-10b_n} > 6, \ \forall\ n \ge 4.$$

Next observe that the maximum degree of d' of a potentially overshadowing class is $< \frac{10}{3-10b_n} < 10$ because $b_n = M/D < 1/5 < 3/10$. If $d' = 7, 8, 9$ the inequality $m'/d' > 3/10$ implies that $m' \ge 3$, so that

$$|bd' - m'| = m' - d'b \ge 3 - 9/5 > 1.$$

Therefore there are no solutions of these degrees. Similarly, one can check that there are no solutions if $d' = 4, 5$, so that $m' \ge 2$, or if $d' = 3$, $m' = 2$. If $d' \le 3$ with $m' = 1$ then the only possible solution is $\mathbf{E}' = (3, 1; 2, 1^{\times 5})$, but as pointed out in Example 32 this class is not even obstructive at its break point $z = 6$ when $b < 1/5$.

However, if $d' = 6$, then the inequality $m'/d' > 3/10$ is satisfied with $m' = 2, 3$, and if $m' = 2$ there are $b < 1/5$ such that $|6b - 2| < \sqrt{1 - b^2}$. We rule this case out as follows. The break point a' of any overshadowing class $\mathbf{E}'$ must be > 6, and hence have $\ell(a') \ge 7$. Further, because $\mathbf{w}(a')$ starts with a block of length at least 6, it follows from Lemma 18 (ii) that the first 6 entries of $\mathbf{m}'$ are $k^{\times 5}, k'$ where $|k' - k| \le 1$. The identity $\sum m_i = 3d' - m' - 1 = 15$ shows that $k = 1$ or 2, while the quadratic identity shows that $k \ne 1$. Therefore we would have to have $\mathbf{E}' = (6, 2; 3, 2^{\times 6})$, with break point 7; to see this notice that 7 is the unique point

$a > 6$ with $\ell(a) = 7$. Arguing as in the proof of Lemma 16, we find that

$$\mu_{\mathbf{E}',b}(z) = \frac{1+2z}{6-2b} < \frac{1+z}{3-b}, \qquad \text{for } 6 < z < 7.$$

But by (2.3.7) when $z = \mathrm{acc}(b) \in (6, 7)$ we have $V_b(z) = \frac{1+z}{3-b}$. Hence the constraint $\mu_{\mathbf{E}',b}$ never meets the volume constraint at the accumulation point $z = \mathrm{acc}(b)$ and so cannot be an overshadowing class. This completes the proof.

Proof of Theorem 54 This holds by Lemmas 72–74 together with Proposition 86, using the proof scheme explained at the beginning of Sect. 3.2.

Now consider the staircases $\mathcal{S}^E_{\bullet,n}$ with blocking classes

$$\mathbf{B}^E_n = \big(5(n+3), n+4; (2n+6)\mathbf{w}(\frac{12n+35}{2n+6})\big).$$

Since we are given simple formulas for the components of $\mathbf{B}^E_n$, one can readily check that these are Diophantine classes, and hence quasi-perfect.

For Step 2, we must check that the staircase classes are quasi-perfect. The argument in Lemma 72 adapts in a straightforward way. In particular, as we show in Lemma 75 below, the staircase $\mathcal{S}^E_{\ell,n}$ can again be extended by a class defined for $k = -1$, which simplifies the calculations. We will leave it to the reader to check the further details of this step.

Step 3 involves checking that the relations in the pre-staircases are those associated to $\mathbf{B}^E_n$ as described in Theorem 52, and so follows by inspection. However, the last step, which establishes that the staircase classes are live at the limiting b-value, is somewhat more tricky, specially in the case $\mathcal{S}^E_{u,n}$. First consider the ascending pre-staircases $\mathcal{S}^E_{\ell,n}$. By Corollary 60 they are the image by the shift Sh of the corresponding staircases $\mathcal{S}^U_{\ell,n}$, and have b-values in the range $(1/5, 1/3)$.

Lemma 75 *The pre-staircases $\mathcal{S}^E_{\ell,n}$, $n \geq 1$, satisfy condition (i) in Theorem 51.*

Proof We first claim that, as with the staircase $\mathcal{S}^L_{u,n}$ considered in Lemma 72, when $n \geq 1$ we can extend each $\mathcal{S}^E_{\ell,n}$ by a class $\mathbf{E}^E_{\ell,n,-1}$ with center $Sh(2n+4) = \frac{12n+23}{2n+4}$ and hence with parameters

$$(5n+10, n+3; 12n+23, 2n+4).$$

To see this, note that by transforming the formulas for p/q in (3.2.7) by Sh and then using the linear relation (or by direct computation), the parameters of this staircase are

$$(d_{-1}, m_{-1}) = (5n+10, n+3)$$
$$(d_0, m_0) = (20n^3 + 100n^2 + 156n + 74, 4n^3 + 24n^2 + 44n + 24).$$

As in Lemma 72 it is straightforward to check that entries p_k, q_k of these staircases satisfy the recursion for $k = -1, 0, 1$, so that d_k, m_k do as well. This justifies our extension of the staircase by a term with $k = -1$.

We may now calculate

$$\frac{m_0 d_{-1} - m_{-1} d_0}{2n+3} = 2n + 6 > 0.$$

Therefore, by (2.4.5), and (3.2.11), we are in case (ii) of Lemma 67. Next notice that, taking $r = 1, s = 2$, we have

$$\frac{m_0 d_{-1} - m_{-1} d_0}{\sqrt{\sigma_n}(2n+3)} \le \frac{2n+6}{2n+2} \le 2 < \frac{2D - M}{D - 2M}, \qquad \forall\, n \ge 1,\ 0 \le \frac{M}{D} < \frac{1}{2}.$$

Therefore, because in fact $M/D < 1/3$ it remains to show there is no overshadowing class with $d' < 2/(1 - 2/3) = 6$ and $m'/d' > 1/2$. Thus we must have $(d', m') = (3, 2), (4, 3), (5, 3)$ or $(5, 4)$. But no such classes can be live at $b < 1/3$ since in none of these cases is $|d'b - m'| < 1$, which is a necessary condition for a class to be obstructive by Lemma 15 (i). This completes the proof.

It requires somewhat more work to rule out overshadowing classes for the staircases $\mathcal{S}^E_{u,n}$. The following remark will be helpful in understanding the possible break points of such a class since it explains how the order on $\mathbb{Q}$ interacts with continued fraction expansions.

Remark 76 Each rational number $a > 1$ has a continued fraction representation $[\ell_0; \ell_1, \dots, \ell_N]$, where ℓ_i is a positive integer, and $N > 0$ unless $a \in \mathbb{Z}$. By convention, the last entry in a continued fraction is always taken to be > 1, since $[\ell_0; \ell_1, \dots, \ell_N, 1] = [\ell_0; \ell_1, \dots, \ell_N + 1]$; for example

$$[1; 3, 1] = 1 + \frac{1}{3 + \frac{1}{1}} = 1 + \frac{1}{4} = [1; 4].$$

If N is odd, then $[\ell_0; \ell_1, \dots, \ell_N] > [\ell_0; \ell_1, \dots, \ell_N + 1]$, in other words, if the last place is odd, increasing this entry decreases the number represented. For example, because $2/13 < 3/19$, we have

$$[1; 3, 6, 2] = 1 + \frac{1}{3 + \frac{1}{6 + \frac{1}{2}}} = 1 + \frac{1}{3 + \frac{2}{13}} > [1; 3, 6, 3] = 1 + \frac{1}{3 + \frac{1}{6 + \frac{1}{3}}} = 1 + \frac{1}{3 + \frac{3}{19}},$$

Further, although increasing an even place usually increases the number represented, this rule does not hold if one increases an even place[13] from 0 to a positive

[13]This place would necessarily be the last; and notice that the initial place is labelled 0 and hence is considered to be even.

number. For example

$$[1;4] = 1 + \frac{1}{4} > [1;4,2] = 1 + \frac{1}{4 + \frac{1}{2}} = 1 + \frac{2}{9} > [1;4,1] = [1;5] = 1 + \frac{1}{5}. \tag{3.3.9}$$

Similarly, if one increases an odd place from 0, the corresponding number increases. In particular,

$$[\ell_0; \ell_1, \ldots, \ell_N] > [\ell_0; \ell_1, \ldots, \ell_N + 1, \ldots] \quad \text{if } N \text{ is odd, } \ell_N \geq 2 \tag{3.3.10}$$

whatever the last entries are. ◊

Example 77 Consider the pre-staircase $\mathcal{S}^E_{u,0}$. Its first two steps (with $\text{end}_0 = 4$) are

$$\big(73, 20; 29\mathbf{w}(\frac{170}{29})\big), \quad \big(524, 145; 208\mathbf{w}(\frac{1219}{208})\big).$$

Using (2.4.4), one can calculate

$$d_0 = 73, \quad m_0 = 20, \quad d_1 = 524, \quad m_1 = 145,$$
$$D = \frac{73}{2} + \frac{179}{10}\sqrt{5}, \qquad M = 10 + 5\sqrt{5}, \qquad \frac{M}{D} = b_\infty \approx 0.27677.$$

Therefore

$$\frac{m_1 d_0 - m_0 d_1}{3} = 35 < \sqrt{5}\frac{3D - M}{D - 3M}.$$

Thus we are in case (ii) of Lemma 67 with $r = 1, s = 3$, and must check that there are no overshadowing classes with

$$d' < \frac{3}{1 - 3b_\infty} < 18, \quad m'/d' > 1/3, \quad \text{and} \quad |m' - d' b_\infty| < 1.$$

It is straightforward to check that the only possibilities for (d', m') are

$$(d', m') = (11, 4), \quad (8, 3), \quad (5, 2), \quad (4, 2), \quad (2, 1), \quad (1, 1).$$

For example, the pairs $(17, 6)$, $(14, 5)$ satisfy the first two requirements, but not the third.

If a' is the break point of the overshadowing class $\mathbf{E}'$, then $a' > a_\infty$ because $\mathbf{E}'$ is right-overshadowing, in other words it is obstructive for $z > a_\infty$ to the right of a_∞; see Definition 29. Further, we must have $\ell(a') < \ell(z)$ for all $z \in [a_\infty, a')$ where $a_\infty \approx 5.86\ldots$ is the accumulation point of the staircase. Therefore we cannot have

$a' > 6$. Since

$$[5; 2] < [5; 1, 6] < a_\infty = [5; 1, 6, \{5, 1\}^\infty] < a',$$

it follows from Remark 76 that the possibilities for a' in decreasing order are:

$$6 > [5; 1, 7] > [5; 1, 6, 4] > [5; 1, 6, 5] > [5; 1, 6, 5, 2] > \dots$$

Notice for example that Lemma 18 implies that we cannot have $a' = [5; 1, 6, 3]$ or $[5; 1, 6, 2]$ because, with $z = [5; 1, 6, 5]$, $\mathbf{m}'$ would differ from $\mathbf{w}(z)$ in at least two places of the last block.

Suppose first that $(d', m') = (11, 4)$. If $a' = 6$ then, by Lemmas 14 (iii) and 18, $\mathbf{m}' = (k^{\times 5}, k + \varepsilon)$ where $\varepsilon \in \{-1, 0, 1\}$. Therefore the linear Diophantine equation gives

$$6k + \varepsilon = 3d' - m' - 1 = 28,$$

which has no solution. On the other hand, if $a' = [5; 1, 7]$ the tuple $\mathbf{m}' = (4^{\times 5}, 1^{\times 8})$ does satisfy this equation. However, it does not satisfy the quadratic identity $(d')^2 - (m')^2 = \sum m_i^2 - 1$. A good way to think of this quadratic identity is as follows. An exceptional curve in class $\mathbf{E}'$ is obtained by blowing up the singular points of a rational curve C of degree d' with branch points of orders $m', m'_1, \dots, m'_n$, where $\mathbf{m}' = (m'_1, \dots, m'_n)$ and the order of a branch point is the number of branches through that point. Thus a generic perturbation of a branch point of order k has $\frac{k(k-1)}{2}$ nearby simple double points. The quadratic identity can be reformulated into the statement that, when C is perturbed so that it only has simple double points, then the total number of these double points is $\frac{(d'-1)(d'-2)}{2}$. Since a curve of degree 11 has 45 double points, and the coefficient $m' = 4$ accounts for 6 of them, the tuple $\mathbf{m}'$ must therefore account for 39 double points. But the tuple $(4^{\times 5}, 1^{\times 8})$ only accounts for $5 \times 6 = 30$ double points. Further, if $\ell(a')$ were longer any possible choices for $\mathbf{m}'$ would have even fewer double points. Indeed, when $\mathbf{m}'$ is subject to the linear Diophantine constraint, $\sum (m'_i)^2$ is at most $\frac{(3d'-m'-1)^2}{\ell(\mathbf{m}')} < (d')^2 - (m')^2 + 1$ whenever $\ell(\mathbf{m}') \geq 8$. Therefore there are no solutions for $(d', m') = (11, 4)$.

Similar arguments rule out the cases $(8, 3)$ and $(4, 2)$, while the degree of $(2, 1)$ is simply too small for the break point to be a'. If $(d', m') = (5, 2)$, there is an exceptional divisor $(5, 2; 2^{\times 5}, 1^{\times 2})$, but its break point is $[5; 2] < a_\infty$ which is too small. Hence there are no overshadowing classes, so that the pre-staircase $\mathcal{S}^E_{u,0}$ satisfies condition (i) in Theorem 51. Hence it is a staircase as claimed in Theorem 58. ◊

Lemma 78 *The pre-staircases $\mathcal{S}^E_{u,n}$, $n \geq 0$, satisfy condition (i) in Theorem 51.*

Proof Since the case $n = 0$ is treated in Example 77, we will suppose that $n \geq 1$. One can calculate that

$$(d_0, m_0) = \left(10n^2 + 55n + 73,\ 2n^2 + 13n + 20\right)$$

$$(d_1, m_1) = \left(40n^4 + 340n^3 + 1022n^2 + 1261n + 524,\ 8n^4 + 76n^3 + 250n^2 + 331n + 145\right)$$

Hence

$$m_1 d_0 - m_0 d_1 = 24n^2 + 106n + 105 = (2n + 3)(12n + 35) > 0.$$

Therefore, because $\sqrt{\sigma_n} = 2n + c$ where $2 < c < 3$, we have

$$\frac{m_1 d_0 - m_0 d_1}{\sqrt{\sigma_n}(2n + 3)} = \frac{12n + 35}{\sqrt{\sigma_n}} = 6 + \frac{c'}{2n + c} \quad \text{for some } c' > 0.$$

Thus this function decreases, with limit 6. Further, it is approximately 10.25 when $n = 1$, is < 8.8 when $n = 2$, < 8.1 when $n = 3$, and < 7.2 when $n = 6$, and < 7 when $n \geq 8$.

One can calculate that with $r/s = 1/3$ we have

$$\frac{3d_0 - m_0}{d_0 - 3m_0} = \frac{28n^2 + 152n + 199}{4n^2 + 16n + 13} = 7 + \frac{40n + 108}{4n^2 + 16n + 13}$$

Therefore, because the b_k increase, it follows from Remark 68 (ii) that condition (ii) in Lemma 67 will hold with $r/s = 1/3$ provided that

$$\frac{12n + 35}{\sqrt{\sigma_n}} < 7 + \frac{40n + 108}{4n^2 + 16n + 13}, \qquad \text{for all } n \geq 1.$$

But this is easy to check from the estimates given above for $(12n+35)/\sqrt{\sigma_n}$. Indeed, this is immediate for $n \geq 8$ since we saw above that $\frac{12n+35}{\sqrt{\sigma_n}} < 7$ in this case. If $n = 1$ then we saw that $RHS \approx 10.25$ while $LHS = 7 + 148/33 > 11$. The other cases follow by a similar computation.

Hence it remains to check that there are no overshadowing classes with $m'/d' > 1/3$ and suitably small degree. When $n = 1$, we have $b_{1,\infty} \approx 0.25416$, so that $3/(1 - 3b_{1,\infty}) < 13$. Therefore we must rule out overshadowing classes of degrees $d' \leq 12$, $m'/d' > 1/3$, and $|m' - b_{1,\infty} d'| < 1$. Thus, the possibilities for (d', m') are:

$$(d', m') = (8, 3),\quad (5, 2),\quad (4, 2),\quad (2, 1),\quad (1, 1). \tag{3.3.11}$$

Because the steps of $\mathcal{S}^E_{u,1}$ have centers $[5; 1, 8, \{7, 3\}^k, 6]$, it follows as in Example 77 that the possibilities for the break point a' of $\mathbf{E}'$ are

$$6 > [5; 1, 9] > [5; 1, 8, 6] > [5; 1, 8, 7, 4] > \ldots.$$

But it is straightforward to check that, just as in Example 77, there are no suitable classes.

When $n > 1$ the proof is very similar. Since $b_{n,\infty} \le b_{1,\infty}$, the only possible (d', m') are among those listed in (3.3.11), while the possible values of a' are $6 > [5; 1, 7+2n] > \dots$ and so are longer than those for $n = 1$. Therefore as n increases the argument gets easier, because in the case of $(d', m') = (8, 3), (5, 2)$, or $(4, 2)$, we have $\frac{(3d'-m'-1)^2}{\ell(\mathbf{m}')} < (d')^2 - (m')^2 + 1$ whenever $\ell(\mathbf{m}') \ge 7$. Because we have $\ell([5; 1, 7 + 2n]) = 12 + 2n$, the break point of $\mathbf{E}'$ has to be 6, and we showed above that there is no corresponding overshadowing class. As before, we cannot have $(d', m') = (2, 1)$ or $(1, 1)$. Therefore, there are no overshadowing classes for any n, and the proof is complete.

Proof of Theorem 58 This holds by Lemmas 75, 78 and Proposition 86.

3.4 Cremona Reduction

This subsection is devoted to the proof of the following result.

Proposition 79 *Each pre-staircase*

$$\mathcal{S}^L_{\ell,n}, \mathcal{S}^U_{u,n}, \mathcal{S}^E_{u,n}, \quad n \ge 0, \qquad \mathcal{S}^L_{u,n}, \mathcal{S}^U_{\ell,n}, \mathcal{S}^E_{\ell,n}, \quad n \ge 1$$

mentioned in Sect. 3.1 is perfect, that is, it consists of exceptional classes.

Proof This is contained in Remark 81 (ii) (for the case $n = 0$), Corollary 84 for the pre-staircases labelled U, and Proposition 86 for the other cases.

We will use the following recognition principle, which is explained in [17, Prop.1.2.12], for example.

Lemma 80 *An integral class* $\mathbf{E} := dL - \sum_{i=1}^N n_i E_i$ *in the* N *fold blowup* $\mathbb{CP}^2 \# N\overline{\mathbb{CP}}^2$ *represents an exceptional divisor if and only if it may be reduced to* E_1 *by repeated application of Cremona transformations.*

Here, a **Cremona transformation** is the composition of the transformation

$$\big(d, n_1, n_2, n_3, \dots\big) \mapsto \big(2d - (n_1 + n_2 + n_3), d - (n_2 + n_3), d - (n_1 + n_3), d - (n_1 + n_2), \dots\big)$$

with a reordering operation that is usually taken to put the entries in nonincreasing order (but does not have to be this). If this reordering does put the entries in decreasing order, we call it a **standard Cremona move**.

Note that because Cremona transformations preserve the first Chern class and the self-intersection number, once we have proved that $\mathbf{E}_k$ reduces to E_1 there is no need to make an independent check that $\mathbf{E}_k \cdot \mathbf{E}_k = -1$. Further, because Cremona moves are reversible (indeed if one omits the reordering step they have order 2), they

generate an equivalence relation on the set of solutions to the Diophantine equations. Hence any two classes with a common reduction are themselves equivalent.

Remark 81

(i) **(On notation)** There are two useful notations for an exceptional class, namely $\mathbf{E} := (d, m; \mathbf{m})$, and $(d, m_1, \dots, m_N)$ with $m_1 \geq m_2 \geq \cdots \geq m_N$. We use the former (with the semicolon) when we are interested in the geometry of the class as in Sect. 2, while we use the latter (with no semicolon) when, as in Example 85, we are interested only in its numerics and therefore include m among all the other coefficients.

(ii) **(The case $n = 0$)** The three pre-staircases $\mathcal{S}^L_{\ell,0}, \mathcal{S}^U_{u,0}, \mathcal{S}^E_{u,0}$ are rather special since they have no twin. Indeed $\mathcal{S}^L_{\ell,0}$ is not even defined in Theorem 54, although as we saw in Remark 55 it can be identified with the Fibonacci staircase. We will prove in Proposition 86 that the classes in these three staircases are all equivalent. Therefore, because the Fibonacci staircase is known to consist of exceptional classes by McDuff and Schlenk [17], these three pre-staircases are all perfect. Hence the other arguments below assume that $n \geq 1$.

(iii) As we will see, the reduction processes for the different pre-staircases are closely related. This appears most clearly in Proposition 86. However, we do not attempt here to provide a unified proof valid for all pre-staircases, but instead will consider each pre-staircase family separately. This point will be discussed more fully in our next paper. Note also that, because our arguments concern the linear relations among the coefficients at the beginning of the tuple, for all the arguments except those in Corollary 84 the choice of end_n is irrelevant. ◊

We will first discuss the classes $\mathbf{E}^U_{\ell,n,k}, \mathbf{E}^U_{u,n,k}$, in the pre-staircases $\mathcal{S}^U_{\ell,n}, \mathcal{S}^U_{u,n}$ with centers

$$\frac{p_{n,k}}{q_{n,k}} = \begin{cases} [\{2n+5, 2n+1\}^k, \mathrm{end}_n], & \text{for } \mathcal{S}^U_{\ell,n},\ n \geq 1,\ k \geq 0 \\ [2n+7; \{2n+5, 2n+1\}^k, \mathrm{end}_n], & \text{for } \mathcal{S}^U_{u,n},\ n \geq 1,\ k \geq 0, \end{cases}$$

and linear relations

$$\begin{aligned} (2n+3)d_{n,k} &= (n+1)p_{n,k} + (n+2)q_{n,k} \quad \text{for } \mathcal{S}^U_{\ell,n} \\ (2n+3)d_{n,k} &= (n+2)p_{n,k} - (n+4)q_{n,k} \quad \text{for } \mathcal{S}^U_{u,n}. \end{aligned} \tag{3.4.1}$$

As in Lemma 63 (i), the (integral) weight expansion of $[\{2n+5, 2n+1\}^k, 2n+4]$ for $k \geq 1$ is denoted

$$\left(a_{n,k}^{\times(2n+5)}, b_{n,k}^{\times(2n+1)}, (a_{n,k} - (2n+1)b_{n,k})^{\times(2n+5)}, \dots, 1^{\times(2n+4)}\right). \tag{3.4.2}$$

while that for $[\{2n+5, 2n+1\}^k, 2n+5, 2n+2]$ for $k \geq 1$ is denoted

$$\big(a'_{n,k}{}^{\times(2n+5)}, b'_{n,k}{}^{\times(2n+1)}, (a'_{n,k} - (2n+1)b'_{n,k})^{\times(2n+5)}, \dots, 1^{\times(2n+4)}\big). \tag{3.4.3}$$

Further, in the pre-staircase $\mathcal{S}^U_{\ell,n,k}$ we have

$$p_{n,k-1} = \begin{cases} b_{n,k}, & \text{if } \mathrm{end}_n = 2n+4, \\ b'_{n,k}, & \text{if } \mathrm{end}_n = 2n+5, 2n+2. \end{cases} \tag{3.4.4}$$

Here are our two main results about these pre-staircases.

Lemma 82 *For each $n \geq 1, k \geq 0$ there is a tuple $\mathbf{E}'_{n,k}$ that reduces to $\mathbf{E}^U_{\ell,n,k}$ in $n+2$ standard Cremona moves and is such that $\mathbf{E}^U_{u,n,k}$ reduces to it in $n+3$ standard Cremona moves.*

Lemma 83 *For each $n \geq 1, k \geq 0$ the class $\mathbf{E}^U_{\ell,n,k+1}$ reduces to $\mathbf{E}'_{k,n}$ in $3n+2$ moves all of which are standard except for the last one.*

Corollary 84 *The classes $\mathbf{E}^U_{\ell,n,k}$ and $\mathbf{E}^U_{u,n,k}$ are perfect for all $n \geq 1, k \geq 0$ and both choices of* end_n.

Proof By Lemmas 82 and 83 it suffices to check that the classes $\mathbf{E}^U_{\ell,n,0}$, $n \geq 1$, are perfect. But the center of this class is either $2n+4$ or $[2n+5; 2n+2]$ and the relation is $(2n+3)d = (n+1)p + (n+2)q$. Therefore with $\mathrm{end}_n = 2n+4$, the class is $(n+2, n+1; \mathbf{w}(2n+4)) = B^U_{n-1}$, which is easily seen to be perfect. Further, with $\mathrm{end}_n = (2n+5, 2n+2)$ the class is

$$\begin{aligned} &\big((n+1)(2n+5), (n+1)(2n+3) - 1; (2n+2)^{\times(2n+5)}, 1^{\times(2n+2)}\big) \\ &\quad = \big((n+2)(2n+1) + (2n+3), (n+2)(2n+1), (2n+2)^{\times(2n+5)}, 1^{\times(2n+2)}\big) \\ &\quad = \big((n+3)s + 2, (n+2)s, (s+1)^{\times(2n+5)}, 1^{\times(2n+2)}\big), \quad s := 2n+1. \end{aligned}$$

Each standard Cremona move takes the tuple $\big(d, m, (s+1)^{\times i}, 1^{\times j}\big)$ to $\big(d-s, m-s, (s+1)^{\times(i-2)}, 1^{\times(j+2)}\big)$. Hence after $(n+2)$ such moves, we obtain the tuple $\big(2n+3, 0, 2n+2, 1^{\times(4n+6)}\big)$ that corresponds to the perfect class

$$\big(2n+3, 2n+2; 1^{\times(4n+6)}\big) = \mathbf{B}^U_{2n}.$$

This completes the proof.

Proof of Lemma 82 The class $\mathbf{E}'_{n,k}$ in Lemma 82 has the following form:

$$\Big(\ell, s_0, s_1^{\times(2n+4)}, q^{\times(2n+5)}, r^{\times(2n+1)} \dots\Big), \quad \text{where } s_0 + s_1 = \ell, \quad s_1 := \frac{\ell - q}{2n+3} \tag{3.4.5}$$

$$\text{and } (\ell, q, r) = \begin{cases} (b_{n,k+1}, a_{n,k}, b_{n,k}) & \text{if end}_n = (2n+4) \\ (b'_{n,k+1}, a'_{n,k}, b'_{n,k}) & \text{if end}_n = (2n+5, 2n+2), \end{cases}$$

and where[14] $(q^{\times(2n+5)}, r^{\times(2n+1)} \dots)$ is the integral weight expansion of the center $p^U_{\ell,n,k}/q^U_{\ell,n,k}$ of $\mathbf{E}^U_{\ell,n,k}$.

We first prove that $\mathbf{E}'_{n,k}$ reduces to $\mathbf{E}^U_{\ell,n,k}$. To this end, note that one standard Cremona move gives

$$(\ell, s_0, s_1^{\times(2n+4)}, q^{\times(2n+5)}, \dots) \to (\ell - s_1, \ell - 2s_1, s_1^{\times(2n+2)}, q^{\times(2n+5)}, \dots),$$

which still has the property that the first entry equals the sum of the second two. Hence a total of $(n+2)$ such moves gives

$$\big(\ell - (n+2)s_1, \ell - (n+3)s_1, q^{\times(2n+5)}, \dots\big).$$

To finish the proof that $\mathbf{E}'_{n,k}$ reduces to $\mathbf{E}^U_{\ell,n,k}$ as claimed, we must show that the above tuple is precisely $\mathbf{E}^U_{\ell,n,k}$.

But by (3.4.5), the center of $\mathbf{E}^U_{\ell,n,k}$ is precisely $\ell/q = ((2n+5)q + r)/q$. Then, by (3.4.5) we have $s_1 = q + \frac{q+r}{2n+3}$ and we must show that

$$d := d^U_{\ell,n,k} = \ell - (n+2)\big(q + \frac{q+r}{2n+3}\big),$$
$$m := m^U_{\ell,n,k} = \ell - (n+3)\big(q + \frac{q+r}{2n+3}\big).$$

But the linear relation in $\mathcal{S}^U_{\ell,n,k}$ gives

$$(2n+3)d = (n+1)\ell + (n+2)q = (n+1)\big((2n+5)q + r\big) + (n+2)q.$$

It is now easy to check that the two expressions for d are the same.

It remains to prove that the class $\mathbf{E}^U_{u,n,k}$ reduces to $\mathbf{E}'_{n,k}$. As noted in the proof of Lemma 66, we now have

$$q^U_{u,n,k} = \ell, \quad p^U_{u,n,k} = (2n+7)\ell + q,$$

where ℓ, q are as above, and d, m are redefined as follows:

$$\begin{aligned}(2n+3)d := (2n+3)\, d^U_{u,n,k} &= (n+2)\big((2n+7)\ell + q\big) - (n+4)\ell \\ &= (2n^2 + 10n + 10)\ell + (n+2)q, \end{aligned} \tag{3.4.6}$$

[14] We assume here that $k > 0$—the case $k = 0$ can safely be left to the reader.

$$(2n+3)m := (2n+3)\, m^U_{u,n,k} = (2n^2+8n+6)\ell + (n+3)q.$$

Thus

$$\mathbf{E}^U_{u,n,k} \;=\; \big(d, m, \ell^{\times(2n+7)}, q^{\times(2n+5)}, \dots\big),$$

and we aim to reduce it to

$$\mathbf{E}'_{n,k} = \big(\ell, s_0, s_1^{\times(2n+4)}, q^{\times(2n+5)}, \dots\big), \quad s_0 + s_1 = \ell, \;\; s_1 := \frac{\ell - q}{2n+3}$$

We will do this by $n+3$ standard Cremona moves, where all the moves until the last replace 2 copies of ℓ by 2 copies of s_1, and the last removes the final 3 copies of ℓ and shrinks length by 3. The first move uses the identity

$$d - m - \ell = s_1 = \frac{\ell - q}{2n+3},$$

which follows easily from (3.4.6). These moves have the following form (where ... now includes the terms $q^{\times(2n+5)}$)

$\big(d, m, \ell^{\times(2n+7)}, \dots\big)$

$\to \;\; \big(2d - m - 2\ell,\; d - 2\ell,\; \ell^{\times(2n+5)}, \dots, (d-m-\ell)^{\times 2} = s_1^{\times 2}, \dots\big)$ after 1 move

$\to \;\; \big(3d - 2m - 4\ell,\; 2d - m - 4\ell,\; \ell^{\times(2n+3)}, s_1^{\times 4}, \dots\big)$ after 2 moves

$\to \;\; \big((n+3)d - (n+2)m - 2(n+2)\ell = 2\ell,\; \ell^{\times 3},$

$\qquad\qquad (n+2)d - (n+1)m - 2(n+2)\ell = s_0, s_1^{\times(2n+4)}, \dots\big)$ after $n+2$ moves

$\to \;\; \big(\ell, s_0, s_1^{\times(2n+4)}, \dots\big),$

where we have used the identity

$$(n+3)d - (n+2)m - 2(n+2)\ell = 2\ell, \tag{3.4.7}$$

(which follows from (3.4.6)) and its corollary

$$(n+2)d - (n+1)m - 2(n+2)\ell = 2\ell - d + m = \ell - s_1 =: s_0.$$

This completes the proof.

Example 85 Since the proof of Lemma 83 is somewhat more complicated, we begin with an example. Since $\mathbf{E}^U_{\ell,1,2} = \big(1538, 987; \mathbf{w}(3191/436)\big)$, we have

$$\mathbf{E}^U_{\ell,1,2} = \big(1538, 987, 436^{\times 7}, 139^{\times 3}, 19^{\times 7}, 6^{\times 3}, 1^{\times 6}\big)$$

which reduces to $\mathbf{E}'_{1,1}$ as follows:

0	1538	987	$436^{\times 7}$	$139^{\times 3}$	$19^{\times 7}$	...		
1	1217	666	$436^{\times 5}$	$139^{\times 3}$	$115^{\times 2}$	$19^{\times 7}$	...	
2	896	$436^{\times 3}$	345	$139^{\times 3}$	$115^{\times 4}$	$19^{\times 7}$	...	
3	484	345	$139^{\times 3}$	$115^{\times 4}$	$24^{\times 3}$	$19^{\times 7}$	...	
4	345	206	139	$115^{\times 4}$	$24^{\times 3}$	$19^{\times 7}$	...	-2
5	230	$115^{\times 3}$	91	$24^{\times 4}$	$19^{\times 7}$	...		-1
6	139	115	$24^{\times 6}$	$19^{\times 7}$	...			-1

All the above Cremona moves are standard, except for the last one where we reduce on the terms $115^{\times 2}$, 91 instead of the three largest. Further, the last column indicates the number of terms that have been reduced to zero; thus altogether the length of the class is reduced by 4. Notice also that these transformations only affect the beginning terms: nothing happens to any term of size 19 or less. Finally, one can check that the last row is precisely $\mathbf{E}'_{1,1}$.

This reduction has several stages. In the notation used in Lemma 83 below, the first three moves reduce the multiplicity of $436 =: q$ from 7 to 0; the next move reduces the multiplicity of $139 =: r$ from 3 to its final value 1, and finally we reduce the multiplicity of the new term $115 = s_0$ from 4 to 1, whilst increasing that of $24 =: s_1$.

Finally notice that although the last move was nonstandard, if we replace it by a standard move we would get $(115, 91, 24^{\times 4}, \dots)$ which is precisely what we get from $\mathbf{E}'_{1,1}$ by doing one standard move. Hence one can compare the reductions we do in this paper, with those provided by a computer program. We chose here to reduce to $\mathbf{E}'_{n,k}$ because it has a nice formula in terms of the basic parameters of the problem. ◊

Proof of Lemma 83 We must show that for each $n \geq 1, k \geq 1$ the class $\mathbf{E}^U_{\ell,n,k+1}$ may be reduced to $\mathbf{E}'_{n,k}$ by Cremona moves. The class $\mathbf{E}^U_{\ell,n,k+1}$ for $k > 0$ starts with the entries[15]

$$\left(d_{n,k+1}, m_{0,n,k+1}, a_{n,k+1}^{\times(2n+5)}, b_{n,k+1}^{\times(2n+1)}, \dots\right) =: \left(d, m, q^{\times(2n+5)}, r^{\times(2n+1)}, \dots\right), \tag{3.4.8}$$

and we aim to reduce it to

$$\mathbf{E}'_{n,k} := \left(r, r - s_1, s_1^{\times(2n+4)}, v\mathbf{w}(r/v)\right), \qquad v := a_{n,k} = q - (2n+1)r, \quad s_1 := \frac{r - v}{2n+3}. \tag{3.4.9}$$

[15] Note that the values of d, m, q, r here are different from those in Lemma 82; for example now $q = a_{n,k+1}$ rather than $b_{n,k+1}$. However, their geometric meaning is the same.

The first part of the reduction process is formally the same as that in the reduction of $\mathbf{E}^U_{u,n,k}$ to $\mathbf{E}'_{n,k}$ in Lemma 82: we reduce the multiplicity of the term in q from $2n+5$ to 3 by $(n+1)$ standard moves, and then reduce it to zero by one further standard move:

$$\begin{aligned}
&\big(d, m, q^{\times(2n+5)}, \dots\big) \\
&\quad\to\ \big((n+2)d-(n+1)m-2(n+1)q,\ q^{\times 3}, (n+1)d-nm-2(n+1)q, \dots \\
&\qquad\qquad (d-m-q)^{\times(2n+2)}, \dots\big) \quad \text{after } n+1 \text{ moves} \\
&\quad\to\ \big(2(n+2)d-2(n+1)m-(4n+5)q,\ (n+1)d-nm-2(n+1)q, \dots \\
&\qquad (d-m-q)^{\times(2n+2)}, ((n+2)d-(n+1)m-2(n+2)q)^{\times 3}\dots\big) \quad \text{after } n+2 \text{ moves.}
\end{aligned}$$

We claim that

$$s_0 := d-m-q, \quad s_1 := (n+2)d-(n+1)m-2(n+2)q$$

satisfy the identities in (3.4.9), namely

$$s_0+s_1 = r, \qquad (2n+3)s_1 = r-v = (2n+2)r-q.$$

These are two linear identities in the quantities d, m, q, r that determine $\mathbf{E}^U_{\ell,n,k}$. The easiest way to check them is to verify that they hold when $k=0, 1$, and then use the recursion. The values for d, m, q, r at $k=0$ can be obtained from Lemma 63; note in particular that when $k=0$ and $\mathrm{end}_n = 2n+4$ we have $(d, m, q, r) = (n+2, n+1, 1, -1)$. The values for $k=1$ and $\mathrm{end}_n = 2n+4$ are in (3.2.7). The details of the other case are left to the reader.

We can write the other entries in the final tuple above as

$$2(n+2)d-2(n+1)m-(4n+5)q = 2s_1+q, \quad (n+1)d-nm-2(n+1)q = q+s_1-s_0.$$

Thus, after $n+2$ moves, we have obtained the tuple

$$\big(2s_1+q,\ q+s_1-s_0, r^{\times(2n+1)}, s_0^{\times(2n+2)}, s_1^{\times 3}, \dots\big),$$

and the next task is to perform n moves on the first three terms to reduce the multiplicity of the term $r^{\times(2n+1)}$ to 1. Because $2s_1+q-(q+s_1-s_0)-r = s_0+s_1-r = 0$, at each step we reduce the length of the tuple by 2. Moreover, $q+s_1-s_0 = 2s_1+q-r$ so that under these moves the first two terms change as follows:

$$\big(2s_1+q,\ 2s_1+q-r\big) \to \big(2s_1+q-r,\ 2s_1+q-2r\big) \to \cdots \to \big(2s_1+q-nr,\ 2s_1+q-(n+1)r\big).$$

Thus we have

$$\big(2s_1+q-nr,\ 2s_1+q-(n+1)r, r, s_0^{\times(2n+2)}, s_1^{\times 3}, \dots\big).$$

The next reduction step gives

$$\begin{aligned}&\big(2s_1+q-nr-s_0,\ 2s_1+q-(n+1)r-s_0, r-s_0, s_0^{\times(2n+1)}, s_1^{\times 3}, \dots\big)\\ &\quad=\big((n+1)s_0-(n-1)s_1,\ ns_0-ns_1, s_0^{\times(2n+1)}, s_1^{\times 4}, \dots\big)\end{aligned}$$

Finally we do n moves that use two copies of s_0 at each step to reduce the multiplicity of s_0 to 1, while building up the multiplicity of s_1 to $2n+4$ and reducing each of the first two terms by $s_0 - s_1$ at each step. Note that throughout this reduction the terms designated by $\dots$ in (3.4.8) remain unchanged, and hence start with $v^{\times(2n+5)}$. Thus finally we have

$$\big(s_0+s_1=r, s_0, s_1^{\times(2n+4)}, v^{\times(2n+5)}, \dots\big) =: \mathbf{E}'_{n,k}.$$

This completes the proof of Lemma 83.

By Corollary 84, we have now completed the proof that the pre-staircases $\mathcal{S}^U_{\bullet,\bullet}$ are perfect. It remains to consider the other pre-staircase families. For ease of reference, here are the steps and relations for the relevant classes:

$$\begin{aligned}
\mathbf{E}^U_{\ell,n,k+1}\quad & p'/q' := [2n+5; 2n+1, \{2n+5, 2n+1\}^k, \mathrm{end}_n],\\
&\qquad (2n+3)d' = (n+1)p' + (n+2)q', \quad n\ge 1,\\
\mathbf{E}^L_{u,n,k}\quad & p/q := [6; 2n-1, 2n+1, \{2n+5, 2n+1\}^k, \mathrm{end}_n],\\
&\qquad (2n+3)d = -(n-1)p + (11n+2)q, \quad n\ge 1,\\
\mathbf{E}^E_{\ell,n,k}\quad & p/q := [5; 1, 2n+4, 2n+1, \{2n+5, 2n+1\}^k, \mathrm{end}_n],\\
&\qquad (2n+3)d = (n+2)p - (n+4)q, \quad n\ge 1,\\
\mathbf{E}^U_{u,n,k}\quad & p'/q' := [2n+7; \{2n+5, 2n+1\}^k, \mathrm{end}_n],\\
&\qquad (2n+3)d' = (n+2)p' - (n+4)q'. \quad n\ge 0,\\
\mathbf{E}^L_{\ell,n,k}\quad & p/q := [6; 2n+1, \{2n+5, 2n+1\}^k, \mathrm{end}_n],\\
&\qquad (2n+3)d = -(n+1)p + (n+2)q, \quad n\ge 0,\\
\mathbf{E}^E_{u,n,k}\quad & p/q := [5; 1, 2n+6, \{2n+5, 2n+1\}^k, \mathrm{end}_n],\\
&\qquad (2n+3)d = -(n+4)p + (11n+31)q, \quad n\ge 0.
\end{aligned}$$

Note that we now use $'$ to distinguish the entries in $\mathcal{S}^U_{\bullet,n}$ from those in the other pre-staircases.

Proposition 86

(i) *For $n \ge 1$, the tuple obtained from $\mathbf{E}^L_{u,n,k}$ by 2 standard moves that get rid of the term $q^{\times 6}$ equals the one obtained from $\mathbf{E}^U_{\ell,n,k+1}$ by 3 moves that reduce the multiplicity of q' from $2n+5$ to $2n-1$, where the first two are standard, each creating two copies of a new term c', and the third uses two copies of q' and one of c'.*

(ii) *For $n \geq 0$, the tuple obtained from $\mathbf{E}^L_{\ell,n,k}$ by 3 standard moves, that first get rid of the term $q^{\times 6}$ and then get rid of the new term $2d-2q$ of multiplicity 3, equals that obtained from $\mathbf{E}^U_{u,n,k}$ by 3 standard moves that reduce the multiplicity of q' from $2n+7$ to $2n+1$.*

(iii) *For $n \geq 1$, the tuple obtained from $\mathbf{E}^E_{\ell,n,k}$ by 5 standard moves equals that obtained from $\mathbf{E}^U_{\ell,n,k+1}$ by 1 standard move.*

(iv) *For $n \geq 0$, the tuple obtained from $\mathbf{E}^E_{u,n,k}$ by 5 standard moves equals that obtained from $\mathbf{E}^U_{u,n,k}$ by 1 standard move.*

In particular, the classes in all the above pre-staircases are perfect.

Proof We prove (i). We have

$$\mathbf{E}^L_{u,n,k} = \big(d, m; q^{\times 6}, r^{\times(2n-1)}, v^{\times(2n+1)}, \dots\big), \qquad \text{where}$$
$$(2n+3)d = (5n+8)q - (n-1)r, \ \ (2n+3)m = (n+3)q - 5nr.$$

Thus $d = 2q + c$, where $c := \frac{(n+2)q-(n-1)r}{2n+3}$; and, if we reorder $\mathbf{E}^L_{u,n,k}$ so that m comes after $q^{\times 6}$ and then do two moves to get rid of the terms $q^{\times 6}$, we obtain the tuple

$$\big(4c - q, c^{\times 3}, (2c-q)^{\times 3}, r^{\times(2n-1)}, m, \dots\big). \tag{3.4.10}$$

On the other hand,

$$\mathbf{E}^U_{\ell,n,k} = \big(d', m', (q')^{\times(2n+5)}, (r')^{\times(2n+1)}, \dots\big), \qquad \text{where} \quad q' = r, r' = v$$
$$d' = (n+2)q' + \frac{n+1}{2n+3}(q'+r'), \quad m' = (n+1)q' + \frac{n}{2n+3}(q'+r').$$

Hence $d' - m' - q' = \frac{q'+r'}{2n+3} =: c'$. Now reduce three times (using one copy of c' and two of q' at the third move) to obtain

$$\begin{aligned}\big(d', m', (q')^{\times(2n+5)}, \dots\big) &\to \big(2d' - m' - 2q', d - 2q', (q')^{\times(2n+3)}, (c')^{\times 2}, \dots\big)\\ &\to \big(3d' - 2m' - 4q', 2d' - m' - 4q', (q')^{\times(2n+1)}, (c')^{\times 4}, \dots\big)\\ &\to \big(6d' - 4m' - 10q' - c', (3d' - 2m' - 5q' - c')^{\times 2}, 3d' - 2m' - 6q',\\ &\qquad (q')^{\times(2n-1)}, 2d' - m' - 4q', (c')^{\times 3}, \dots\big)\end{aligned}$$

We now claim that this tuple is precisely the same as that in (3.4.10). To see this, express all the quantities in terms of the variables q', r', recalling that

$$r = q', \quad v = r', \quad q = (2n-1)r + v = (2n-1)q' + r'.$$

Further, by definition of c' we have $3d' - 2m' - 5q' - c' = 2d' - m' - 4q'$ so that this term occurs with multiplicity 3, and one can check that $c = 2d' - m' - 4q'$. One can then check that the other terms agree:

$$4c - q = 6d' - 4m' - 10q' - c', \quad 2c - q = c', \quad m = 3d' - 2m' - 6q'.$$

This completes the proof of (i).

The proof of (ii) is very similar, and can be done as above. Details are left to the reader. Alternatively, since reduction is a linear transformation it suffices to prove this for $k = 0, 1$ and arbitrary n; the result for $k > 1$ then follows by recursion. One can either do this by calculating the relevant transformations explicitly, or by noting that when $k = 0, 1$, the entries $d', m', \dots$ in the tuples $\mathbf{E}^L_{\ell,n,k}, \mathbf{E}^U_{u,n,k}$ are polynomials in n of degree at most 5 (the explicit formulas for $\mathrm{end}_n = 2n + 4$ are given in (3.2.8) and (3.3.4)). Hence we need to prove some polynomial identities of degree at most 5. But these hold for all n if and only if they hold for 5 different values of n. Thus, without even explicitly calculating these polynomials, we know that the claim must hold for all n, k provided that it holds for $k = 0, 1$ and $0 \le n \le 5$, something that is very easy to check by computer.

We illustrate (iii), (iv) by an example. We have

$$\mathbf{E}^U_{\ell,1,1} = \left(67, 43, 19^{\times 7}, \dots\right) \to \left(53, 29, 19^{\times 5}, \dots, 5^{\times 2} \dots\right).$$

On the other hand, 5 standard moves on $\mathbf{E}^E_{\ell,1,0}$ give:

$$\begin{aligned}\mathbf{E}^E_{\ell,1,0} &= \left(350, 139^{\times 5}, 120, 96, 19^{\times 6}, \dots\right) \to \left(283, 139^{\times 2}, 120, 96, 72^{\times 3}, 19^{\times 6}, \dots\right) \\ &\to \left(168, 96, 72^{\times 3}, 24^{\times 2}, 19^{\times 6}, \dots, 5, \dots\right) \overset{-2}{\to} \left(96, 72, 24^{\times 3}, 19^{\times 6}, \dots, 5, \dots\right) \\ &\overset{-2}{\to} \left(72, 48, 24, 19^{\times 6}, \dots, 5, \dots\right) \overset{-1}{\to} \left(53, 29, 19^{\times 5}, \dots, 5^{\times 2}, \dots\right),\end{aligned}$$

where the superscripts over the arrows denote the number of entries that go to zero at that step. Thus overall 5 terms go to zero, which is the difference in length between the center $[5; 1, 2n + 4, 2n + 1, \{2n + 5, 2n + 1\}^k, \mathrm{end}_n]$ of $\mathbf{E}^E_{\ell,n,k}$ and the corresponding center $[2n + 5; 2n + 1, \{2n + 5, 2n + 1\}^k, \mathrm{end}_n]$ of $\mathbf{E}^U_{\ell,n,k+1}$. A formal proof of the general case can be written out by one of the methods explained above, and is left to the interested reader.

The final claim that all the pre-staircases are perfect holds when $n = 0$ because the Fibonacci staircase is perfect (by McDuff and Schlenk [17]) and holds when $n > 0$ by Corollary 84.

4 Obstructions from ECH Capacities

An alternative way to characterize the ellipsoid embedding function for the target X involves the ECH capacities of the ellipsoid and of X. When we use the computer to plot an approximation of the graph of c_X, we use that alternative characterization of c_X:

$$c_X(z) = \sup_k \frac{c_k(E(1,z))}{c_k(X)}, \tag{4.0.1}$$

so long as X is "convex" (see Definition 87); the equality (4.0.1) is a consequence of [3, Theorem 1.2]. First defined in [9], ECH capacities are invariants of a symplectic 4-manifold (X, ω) that obstruct symplectic embeddings in the following sense:

$$(X,\omega) \hookrightarrow (Y,\omega') \implies \forall_{k\in\mathbb{N}_0}\ c_k(X,\omega) \le c_k(Y,\omega'), \tag{4.0.2}$$

where the c_k form a non-decreasing sequence

$$0 = c_0(X,\omega) < c_1(X,\omega) \le c_2(X,\omega) \le \ldots \le \infty.$$

4.1 Toric Domains

To compute the ECH capacities of the Hirzebruch surfaces H_b, we use the more general theory of ECH capacities of toric domains.

Definition 87 A **toric domain** is the region X_Ω in $\mathbb{C}^2$ determined by

$$X_\Omega := \left\{(z_1,z_2) \in \mathbb{C}^2 \middle| \left(\pi|z_1|^2, \pi|z_2|^2\right) \in \Omega\right\}$$

where $\Omega \subset \mathbb{R}^2$. A toric domain is **concave** if Ω is a closed region in the first quadrant under the graph of a convex function with both axis intercepts nonnegative. Following [11], a toric domain is **convex** if Ω is a closed region in the first quadrant under the graph of a nonincreasing concave function with both axis intercepts nonnegative.

When X is a toric domain we use the standard symplectic form $\omega = \sum_{i=1}^2 dx_i \wedge dy_i$, where $z_i = x_i + \sqrt{-1}y_i$, and we drop ω from the notation for the ECH capacities. Note that ellipsoids are both concave and convex toric domains.

Let X_b denote the convex toric domain X_{Ω_b}, where Ω_b is the Delzant polytope of H_b. By Cristofaro-Gardiner et al. [7, Theorem 1.2], embeddings of ellipsoids into H_b and X_b are equivalent, and therefore $c_{X_b} = c_{H_b}$.

Remark 88 When X is a concave toric domain and Y is a convex toric domain, (4.0.2) is in fact an equivalence [3, Theorem 1.2]. This is the case when we have an ellipsoid embedding into X_b, which justifies the alternative definition (4.0.1). ◊

The way that we use (4.0.1) to plot approximations to the graph is by taking the maximum over a large but finite number of ratios of ECH capacities, rather than the supremum. In order to do this, we need to compute the ECH capacities of an ellipsoid and of X_b.

Remark 89 The sequence of ECH capacities for an ellipsoid $E(a, b)$ is the sequence $N(a, b)$, where for $k \geq 0$, the term $N(a, b)_k$ is the $(k+1)^{\text{st}}$ smallest entry in the array $(am + bn)_{m,n\in\mathbb{N}_0}$, counted with repetitions [16, §1].

To compute the ECH sequence of X_b we use a different method, based on [3, Appendix A]. The definitions and theorem below are based on [3, Definitions A.6, A.7, A.8] and can be found there in more detail. ◊

Definition 90 A **convex lattice path** is a piecewise linear path $\Lambda : [0, c] \to \mathbb{R}^2$ such that all its vertices, including the first $(0, x(\Lambda))$ and last $(y(\Lambda), 0)$, are lattice points and the region enclosed by Λ and the axes is convex. An **edge** of Λ is a vector v from one vertex of Λ to the next. The **lattice point counting function** $\mathcal{L}(\Lambda)$ counts the number of lattice points in the region bounded by a convex lattice path Λ and the axes, including those on the boundary.

Let $\Omega \subset \mathbb{R}^2_{\geq 0}$ be a convex region in the first quadrant. The **Ω-length** of a convex lattice path Λ is defined as

$$\ell_\Omega(\Lambda) = \sum_{v\in \text{Edges}(\Lambda)} \det\left[v \;\; p_{\Omega,v}\right] \tag{4.1.1}$$

where for each edge v we pick an auxiliary point $p_{\Omega,v}$ on the boundary of Ω such that Ω lies entirely "to the right" of the line through $p_{\Omega,v}$ and direction v.

Theorem 91 ([3, Corollary A.5]) *Let X be the toric domain corresponding to the region Ω. Then its kth ECH capacity is given by:*

$$c_k(X) = min\left\{\ell_\Omega(\Lambda) : \Lambda \text{ is a convex lattice path with } \mathcal{L}(\Lambda) = k+1\right\}. \tag{4.1.2}$$

In fact, the lattice path Λ that realizes the minimum in (4.1.2) is shaped roughly like the boundary of the region Ω, see [10, Ex. 4.16(a)]. In particular, the slopes of the edges of the minimizing lattice path must also be slopes of edges of Ω, and we use this fact to compute the capacities of X_b: see Lemma 100.

4.2 ECH Capacities and Exceptional Classes

In this subsection we prove a relationship between the obstructions from exceptional classes of Sect. 3 and those from ECH capacities. This relationship underlies our use of ECH capacities to identify live perfect classes contributing to c_{H_b}, as explained in Sect. 5.4. Let $\Lambda_{d,m}$ denote the lattice path from $(0, d-m)$ to $(m, d-m)$ to $(d, 0)$, which encloses $\mathcal{L}(\Lambda_{d,m})$ lattice points as in Definition 90.

Lemma 92 *If* $\mathbf{E} = (d, m; \mathbf{m})$ *is an exceptional class with* $\mathbf{m} = q\mathbf{w}(p/q)$*, then there is an interval containing the center* p/q *of* $\mathbf{E}$ *on which the obstructions from* $\mathbf{E}$ *and* $c_{\mathcal{L}(\Lambda_{d,m})-1}(X_b)$ *satisfy*

$$\mu_{\mathbf{E},b}(z) \leq \frac{c_{\mathcal{L}(\Lambda_{d,m})-1}(E(1,z))}{c_{\mathcal{L}(\Lambda_{d,m})-1}(X_b)} \tag{4.2.1}$$

Proof We will show that $c_{\mathcal{L}(\Lambda_{d,m})-1}(X_b) \leq d - mb$ and that $c_{\mathcal{L}(\Lambda_{d,m})-1}(E(1,z)) = \mathbf{m} \cdot \mathbf{w}(z)$ for z sufficiently close to p/q. For the former, we have by (4.1.2) and (4.1.1):

$$c_{\mathcal{L}(\Lambda_{d,m})-1}(X_b) \leq \ell_{\Omega_b}(\Lambda_{d,m}) = d - bm.$$

For the latter, we first analyze $c_{\mathcal{L}(\Lambda_{d,m})-1}(E(1,z))$. Note that

$$qN\left(1, \frac{p}{q}\right)_k = N(q,p)_k. \tag{4.2.2}$$

Let $K(p,q) = \frac{(p+1)(q+1)}{2} - 1$. We claim that

$$N(q,p)_{K(p,q)} = N(q,p)_{K(p,q)+1} = pq$$

and these are the only two values of k for which $N(q,p)_k = pq$. If the point in $\mathbb{Z}^2_{\geq 0}$ with coordinates (n_1, n_2) is labeled $n_1 q + n_2 p$ then $N(q,p)$ is the list of such labels ordered as follows. Take the line of slope $-p/q$ and move it from left to right across the plane. Then $N(q,p)$ is the kth label whose lattice point this line crosses, starting from zero. If points are colinear then their indices start from the number of lattice points between the line of slope $-p/q$ through $n_1 q + n_2 p$ and the axes and increase from there. Because p and q are coprime,

$$n_1 q + n_2 p = pq \Rightarrow n_2 p = (p - n_1)q$$

implies that both $n_2 p$ and $(p - n_1)q$ must equal an integer multiple of pq, and we conclude that either $n_1 = 0$ or $n_1 = p$. Therefore there are only two values of k for which $N(q,p)_k = pq$, and the values of k start from the number of lattice points

between the line of slope $-p/q$ through $(pq, 0)$ and $(0, pq)$ and the axes. There are exactly $K(p, q)$ such lattice points.

Next we show that $\mathcal{L}(\Lambda_{d,m}) = K(p, q)+1$. Firstly, by counting the lattice points in the triangle between $(0, 0)$, $(0, d)$, and $(d, 0)$, then removing those in the triangle between $(0, d - m)$, $(0, d)$, and $(m, d - m)$, we obtain

$$\begin{aligned}\mathcal{L}(\Lambda_{d,m}) &= \frac{1}{2}(d(d+3) - m(m+1)) + 1 \qquad (4.2.3)\\ &= \frac{1}{2}(d^2 - m^2 + 1 + 3d - m + 1)\\ &= \frac{1}{2}(pq + p + q + 1) \qquad \text{by (2.1.2) and (2.1.5)}\\ &= K(p, q) + 1\end{aligned}$$

By (4.2.3) and (4.2.2), we have $c_{\mathcal{L}(\Lambda_{d,m})-1}(E(1, p/q)) = N(1, p/q)_{K(p,q)} = p$, and by (2.1.5), we have $\mathbf{m} \cdot \mathbf{w}(p/q) = p$.

We will show that there is an interval I containing p/q for which, when $z \in I, z \geq p/q$,

$$N(1, z)_{K(p,q)} = p = \mathbf{m} \cdot \mathbf{w}(z) \qquad (4.2.4)$$

and when $z \in I, z \leq p/q$,

$$N(1, z)_{K(p,q)} = z = \mathbf{m} \cdot \mathbf{w}(z) \qquad (4.2.5)$$

The first equalities in (4.2.4) and (4.2.5) both follow from considering the lattice $\mathbb{Z}^2_{\geq 0}$ labeled with n_1q+n_2zq defining $N(q, zq)_k$. As z increases from p/q, the slope of the line defining $N(q, zq)_k$ decreases from $-p/q$ to $-z$. When moving this line from left to right, it will now reach $(pq, 0)$ slightly before it reaches $(0, zq)$, but if z is not increased too much, the triangle between the line through $(pq, 0)$ and the axes will still contain $K(p, q)$ points, and there will be no points the line passes strictly between $(pq, 0)$ and $(0, zq)$, so $N(q, zq)_{K(p,q)} = pq$, showing the first equality of (4.2.4).

Similarly, as z decreases, the slope of the line increases from $-p/q$ to $-z$. Now the triangle between the line through $(0, zq)$ of slope $-z$ and the axes will contain $K(p, q)$ points. There will be no points between the line through $(0, zq)$ of slope $-z$ and the line through $(pq, 0)$ of slope $-z$ if z is close enough to p/q, so $N(q, zq)_{K(p,q)} = z$, showing the first equality of (4.2.5).

The second equalities in both (4.2.4) and (4.2.5) follow from Lemma 16.

Remark 93 In all cases we have checked, the inequality (4.2.1) is an equality. Note that if **E** is live, then we must have an equality. However, it seems to be true more generally that there are pairs of lattice paths and exceptional classes for which (4.2.1) is an equality on an interval, even when **E** is not live (see Fig. 10), or is

not quasi-perfect (see Fig. 5). In each of Figs. 10 and 5, there is an interval on which the obstruction coming from the ECH capacity (in bright blue) coincides with the obstruction coming from the exceptional class (in brown), and this interval includes the break point of the exceptional class obstruction. Both figures were plotted using `PlotSingleCapacityObstructionFxn` (see Sect. 5.2) for the bright blue curve and `PlotMu` (see Sect. 5.3) for the brown curve.

Specifically, every constraint $\mu_{\mathbf{E},b}$ has a break point a' in the sense of Lemma 14 even if $\mathbf{E}$ is not perfect, and formulas analogous to those in Lemma 16 do always hold. Exactly what these formulas are depends on the relation between $\mathbf{w}(a')$ and $\mathbf{m}$. One might be able construct a general proof of this fact using the approach to ECH capacities taken in [5, 6] where these capacities are interpreted in terms of the properties of the ECH cobordism map given by a symplectic embedding $E(1,a) \overset{s}{\hookrightarrow} H_b$; see in particular the discussion in [6, §3.1]. However, to prove this here would take us too far afield.

Instead, we will occasionally discuss $\mathbf{E}$ which "correspond" to an ECH capacity c_k in the sense that (4.2.1) is an equality on an interval of interest on which $\mu_{\mathbf{E},b}$ is nontrivial (if $\mathbf{E}$ is quasi-perfect, this interval will always contain its center). ◊

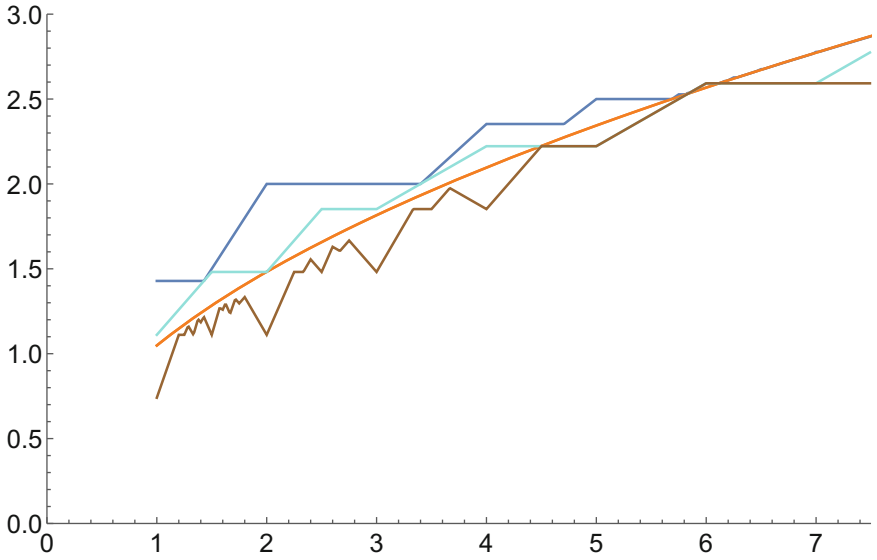


Fig. 5 Comparison between the ECH and exceptional class obstructions for $b = 0.3$. Here a lower bound $c^{\leq}_{H_{0.3}}$ for $c_{H_{0.3}}$ is in dark blue (see Sect. 5), the volume obstruction is in orange, the obstruction from the 8th capacity $c_8(X_{0.3})$ is in bright blue, and the obstruction from the class $\mathbf{E} = \left(3, 1; 2, 1^{\times 5}\right)$ is in brown. The plot of $\mu_{\mathbf{E},0.3}$ is atop that of the obstruction from $c_8(X_{0.3})$ where they agree

4.3 There Is No Infinite Staircase for $b = 1/5$

Let

$$X = 5H_{\frac{1}{5}} = \mathbb{C}P^2_5 \# \overline{\mathbb{C}P}^2_1.$$

The embedding capacity function of X is a scaling of that of $H_{\frac{1}{5}}$, with $c_X(z) = \frac{1}{5} c_{\frac{1}{5}}(z)$. Thus, if X has an infinite staircase then it accumulates at $z = \mathrm{acc}(1/5) = 6$.

Theorem 94 *Let $X = 5H_{\frac{1}{5}}$. For sufficiently small ε the ellipsoid embedding function of X is given by*

$$c_X(z) = \begin{cases} \frac{1}{2} & \text{for } z \in (6-\varepsilon, 6] \\ \frac{z+6}{24} & \text{for } z \in [6, 6+\varepsilon) \end{cases}.$$

Thus, there is no infinite staircase for X, or equivalently for the Hirzebruch surface $H_{\frac{1}{5}}$.

The proof follows closely the proof that the ellipsoid $E(3,4)$ does not have an infinite staircase, as presented in [4, Section 2.5].

Proof The main part of the proof is showing the following claim, whose proof we postpone by a few paragraphs.

Claim 95 For $z \in [6, 6+\varepsilon)$, we have $c_X(z) \leq (z+6)/24$.

Assuming this claim, the rest of the proof goes as follows: the claim gives us an upper bound for $c_X(z)$ for $z \in [6, 6+\varepsilon)$. To get a lower bound, either see Example 34 or use ECH capacities and (4.0.1): the 19th ECH capacity of X is 24 (see how to compute it in Sect. 5.1), whereas the 19th ECH capacity of $E(1,z)$ for this range of z is $z+6$, see Remark 89. Then we have $c_X(z) \geq (z+6)/24$, and in fact $c_X(z) = (z+6)/24$.

For the range $z \in (6-\varepsilon, 6]$, we obtain a lower bound either via Example 22 or again using ECH capacities and (4.0.1): the 5th ECH capacity of X is 10 (see how to compute it in Sect. 5.1), whereas the 5th ECH capacity of $E(1,z)$ for this range of z is 5, see Remark 89. Therefore $c_X(z) \geq 1/2$. To get the upper bound $c_X(z) \leq 1/2$ we first observe that using Claim 95 and the volume constraint at $z = 6$, we obtain $c_X(6) = 1/2$. Since $c_X(z)$ is nondecreasing, we have $c_X(z) \leq c_X(6) = 1/2$ for $z \leq 6$, and therefore in fact $c_X(z) = 1/2$ for $z \in (6-\varepsilon, 6]$.

We now proceed to proving Claim 95. Note that since $c_X(z)$ is continuous we can assume that z is irrational; this will be convenient in some of the arguments that follow. We begin by observing that

$$c_X(z) \leq \frac{z+6}{24} \iff E(1,z) \hookrightarrow \frac{z+6}{24} X$$

$$\begin{aligned} &\iff E(\frac{24}{z+6}, \frac{24z}{z+6}) \hookrightarrow X \\ &\iff \mathrm{ehr}_{\Delta_{\frac{z+6}{24}, \frac{z+6}{24z}}}(t) \geq \mathrm{cap}_X(t), \qquad \forall t \in \mathbb{Z}_{\geq 0}, \end{aligned}$$

where the cap function counts the number of ECH capacities up to t

$$\mathrm{cap}_X(t) = \#\{k | c_k(X) \leq t\}$$

and the Ehrhart function, which is the cap function for an ellipsoid, counts the number of lattice points inside a scaling of a triangle

$$\mathrm{ehr}_{\Delta_{u,v}}(t) = \#\left\{\mathbb{Z}^2 \cap t\Delta_{u,v}\right\},$$

with $\Delta_{u,v}$ the triangle with vertices $(0,0)$, $(u,0)$ and $(0,v)$.

The last equivalence then follows from (4.0.2), Remark 88, and the fact that the ECH capacities of X are all integers because the corresponding region $\Omega \subset \mathbb{R}^2$ is a lattice polygon, see (4.1.1) and (4.1.2).

Because it is hard to count lattice points in a generic triangle, we compare $\mathrm{ehr}_{\Delta_{\frac{z+6}{24}, \frac{z+6}{24z}}}(t)$ with $\mathrm{ehr}_{\Delta_{\frac{z+6}{24}, \frac{z+6}{24z}}}(t)|_{z=6} = \mathrm{ehr}_{\Delta_{\frac{1}{2}, \frac{1}{12}}}(t)$:

$$\mathrm{ehr}_{\Delta_{\frac{z+6}{24}, \frac{z+6}{24z}}}(t) = \mathrm{ehr}_{\Delta_{\frac{1}{2}, \frac{1}{12}}}(t) + D - U - d \tag{4.3.1}$$

where $D = D(t)$ and $U = U(t)$ are respectively the number of lattice points in the closed regions R_U and R_D as in Fig. 6, and $d = d(t)$ is the number of lattice points on the segment that is the left boundary of R_D (but excluding the potential lattice point $(t/4, t/24)$).

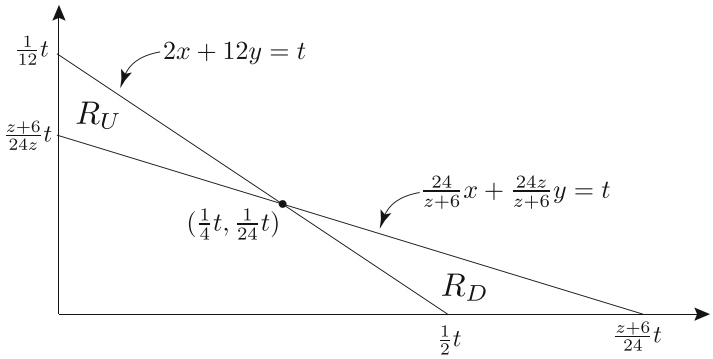


Fig. 6 The triangles $\Delta_{\frac{z+6}{24}t, \frac{z+6}{24z}t}$ and $\Delta_{\frac{1}{2}t, \frac{1}{12}t}$ delimit the regions R_U and R_D. The corresponding lattice point count is given in (4.3.1)

We will then also compare the cap function of X with the Ehrhart function of the same triangle:

$$\mathrm{cap}_X(t) = \mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t) - \tilde{d},$$

where $\tilde{d} = \tilde{d}(t) := \mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t) - \mathrm{cap}_X(t)$.

The proof of Claim 95 will then be complete if we show the following:

Claim 96 For all t we have $U \leq D$, and furthermore for $t \equiv 10 \pmod{24}$ we have $U \leq D - 1$.

Claim 97 For $t \not\equiv 10 \pmod{24}$ we have $\tilde{d} \geq d$ and for $t \equiv 10 \pmod{24}$ we have $\tilde{d} = d - 1$.

Proof of Claim 96 Recall that we are assuming that $z \notin \mathbb{Q}$.

We begin by proving the more general inequality $U \leq D$, in the following way: for each $y_0 \in [t/24, t/12] \cap \mathbb{Z}$ we define a corresponding $y_1 := \lfloor t/12 \rfloor - y_0 \in [0, t/24] \cap \mathbb{Z}$. The values of y_0 range through all possible heights of lattice points in R_U, and for each height we will show that the number of lattice points on that slice of R_U is no greater than the number of lattice points in the slice of R_D at height y_1, which implies that $U \leq D$:

$$\#\{R_U \cap \{y = y_0\}\} \leq \#\{R_D \cap \{y = y_1\}\}. \tag{4.3.3}$$

We can rewrite (4.3.3) as

$$\lfloor x_2 \rfloor - \lceil \max\{0, x_1\} \rceil + 1 \leq \lfloor x_4 \rfloor - \lceil x_3 \rceil + 1, \tag{4.3.4}$$

where

$$x_1 = \frac{t(z+6) - 24zy_0}{24} \qquad x_2 = \frac{t - 12y_0}{2}$$
$$x_3 = \frac{t - 12y_1}{2} \qquad x_4 = \frac{t(z+6) - 24zy_1}{24}.$$

Now, since $\lfloor x_2 \rfloor - \lceil \max\{0, x_1\} \rceil + 1 \leq \lfloor x_2 \rfloor - \lfloor x_1 \rfloor$, we have that (4.3.4) follows from proving that

$$\lfloor x_2 \rfloor - \lfloor x_1 \rfloor \leq \lfloor x_4 \rfloor - \lfloor x_3 \rfloor \qquad \text{whenever } x_3 \notin \mathbb{Z} \tag{4.3.5}$$
$$\lfloor x_2 \rfloor - \lfloor x_1 \rfloor \leq \lfloor x_4 \rfloor - \lfloor x_3 \rfloor + 1 \qquad \text{whenever } x_3 \in \mathbb{Z}. \tag{4.3.6}$$

We first write

$$\lfloor x_2 \rfloor - \lfloor x_1 \rfloor = \frac{t}{2} - \left\{\frac{t}{2}\right\} - 6y_0 - \frac{t(z+6)}{24} + zy_0 + \left\{\frac{t(z+6)}{24} - zy_0\right\}$$

$$\lfloor x_4 \rfloor - \lfloor x_3 \rfloor = \frac{t(z+6)}{24} - zy_1 - \left\{ \frac{t(z+6)}{24} - zy_1 \right\} - \frac{t}{2} + \left\{ \frac{t}{2} \right\} + 6y_1,$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x. We start with the case when $x_3 \notin \mathbb{Z}$, or equivalently, when t is odd. The inequality (4.3.5) can be rewritten as

$$-2\left\{ \frac{t}{2} \right\} - (z-6)\left\{ \frac{t}{12} \right\} + \delta \leq 0 \tag{4.3.7}$$

where

$$\delta = \left\{ \frac{t(z+6)}{24} - zy_0 \right\} + \left\{ \frac{t(z+6)}{24} + zy_0 - z\lfloor \frac{t}{12} \rfloor \right\}.$$

Since $\{m\} + \{n\} \leq \{m+n\} + 1$, and furthermore we can assume for this simplification that $t \in \{0, 1, 2, \ldots, 11\}$, we have

$$\delta \leq \left\{ \frac{t}{2} + z\left\{ \frac{t}{12} \right\} \right\} + 1 \leq (z-6)\left\{ \frac{t}{12} \right\} + 1.$$

Thus we see that (4.3.7) holds when t is odd, since $-2\{t/2\} + 1 = 0$.

We now move on to the case when $x_3 \in \mathbb{Z}$, or equivalently, when t is even. The argument follows exactly as above, except that the right hand side of (4.3.7) is 1 and $-2\{t/2\} + 1 = 1$.

This concludes the proof that $U \leq D$ for general t, we now focus on the case when $t \equiv 10 \pmod{24}$ and will show that $U \leq D - 1$.

We show that in this case, as y_0 ranges over the integers $[t/24, t/12] \cap \mathbb{Z}$, the corresponding y_1 is never equal to the height of the lattice point

$$\left(\frac{t - 12y'}{2}, y' \right) \in \mathbb{Z}^2 \cap R_D,$$

where $y' := \lfloor t/24 \rfloor$. This means that that lattice point is not accounted for in the proof given above for $U \leq D$, and implies that in fact $U \leq D - 1$. Indeed, y_1 is maximized for $y_0 = \lceil t/24 \rceil$ which makes $y_1 = \lfloor t/12 \rfloor - \lceil t/24 \rceil$. When $t \equiv 10 \pmod{24}$, this becomes $y_1 = \lfloor t/24 \rfloor - 1 < y'$. This concludes the proof of Claim 96.

Proof of Claim 97 To prove Claim 97, we must compute d and

$$\tilde{d} = \mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t) - \mathrm{cap}_X(t).$$

Recall that $d = d(t)$ is the number of lattice points on the segment that is the left boundary of R_D (but excluding the potential lattice point $(t/4, t/24)$). Any such lattice point (m, n) must satisfy $2m + 12n = t$, which implies that t is even.

Therefore, $d(t) = 0$ for t odd. Conversely, if t is even and (x, y) is on the left boundary of R_D then $x = \frac{t-12y}{2}$ and $y < t/24$. If furthermore $y \in \mathbb{Z}$ then also $x \in \mathbb{Z}$. It follows that $d(t) = \lceil t/24 \rceil$ for t even.

To compute the function $\mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t)$ we first use [1, Exercise 2.34] to obtain the quadratic and linear terms. To obtain the constant terms, we then use the fact that by [1, Theorem 3.23] the function is quasipolynomial with period at most $lcm(2, 12) = 12$ and compute the necessary ECH capacities of the ellipsoid $E(2, 12)$ using the method described in Remark 89:

$$\mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t) = \frac{t^2}{48} + \begin{cases} \frac{t}{3} & t \text{ even} \\ \frac{7t}{24} & t \text{ odd} \end{cases} + \begin{cases} 1 & t \equiv 0, 8 \\ \frac{11}{16} & t \equiv 1, 9 \\ \frac{5}{4} & t \equiv 2, 6 \\ \frac{15}{16} & t \equiv 3, 7 \end{cases} \quad \begin{cases} \frac{4}{3} & t \equiv 4 \\ \frac{49}{48} & t \equiv 5 \\ \frac{7}{12} & t \equiv 10 \\ \frac{13}{48} & t \equiv 11 \end{cases} \pmod{12}$$

Alternately, it is possible to compute $\mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t)$ using Pick's theorem.

To compute the function $\mathrm{cap}_X(t)$ we use [20, Thm 5.11], with Ω being the region in $\mathbb{R}^2$ corresponding to X, and therefore with area and Ω-perimeter both equal to 24 and affine perimeter equal to 14. We then have that there exists $t_0 \in \mathbb{N}$ such that for $t > t_0$ this function is a quasipolynomial of the form

$$\mathrm{cap}_X(t) = \mathrm{ehr}_\Omega(\lfloor t/48 \rfloor) + \gamma_i,$$

with period 24 and $\gamma_i \in \mathbb{Q}$ for $i = 0, \ldots, 23$. We find that $t_0 = 42$ by checking the conditions in [20, Thm 5.11] and computing enough values of $\mathrm{cap}_X(t)$ using the methods in Sect. 5.1:

(i) $t_0 \geq 2 \times 24 - 14 = 34$;
(ii) $\mathrm{cap}_X(t_0 - 1) < \mathrm{cap}_X(t_0) < \mathrm{cap}_X(t_0 + 1) < \ldots < \mathrm{cap}_X(t_0 + 47)$.

Computing further values of $\mathrm{cap}_X(t)$ we obtain the general formula for $t > 42$:

$$\mathrm{cap}_X(t) = \frac{t^2}{48} + \frac{7t}{24} + \begin{cases} 1 \; t \equiv 0, 10 \\ \frac{11}{16} \; t \equiv 1, 9 \\ -\frac{2}{3} \; t \equiv 2, 8 \\ -\frac{17}{16} \; t \equiv 3, 7 \\ \frac{1}{2} \; t \equiv 4, 6 \\ \frac{49}{48} \; t \equiv 5 \end{cases} \quad \begin{cases} \frac{13}{48} \; t \equiv 11, 23 \\ -\frac{3}{2} \; t \equiv 12, 22 \\ -\frac{5}{16} \; t \equiv 13, 21 \\ \frac{5}{6} \; t \equiv 14, 20 \\ \frac{15}{16} \; t \equiv 15, 19 \\ 0 \; t \equiv 16, 18 \\ -\frac{95}{48} \; t \equiv 17. \end{cases} \pmod{24}.$$

Alternatively, we could compute the function $\mathrm{cap}_X(t)$ by using the fact that the sequence of ECH capacities of X is given by the sequence subtraction $N(5) - N(1)$,

where $N(j)$ is the sequence of ECH capacities of the ball $E(j, j)$, see [7, Definition 2.4 and equation (2.6)].

Still for $t \geq 42$ we now compute $\tilde{d} = \mathrm{ehr}_{\Delta_{\frac{1}{2},\frac{1}{12}}}(t) - \mathrm{cap}_X(t)$ and compare it with d:

$$\tilde{d} = \begin{cases} d & t \equiv 0, 1, 4, 5, 6, 9, 11, 14, 15, 19, 20, 23 \\ d+1 & t \equiv 2, 8, 13, 16, 18, 21 \\ d+2 & t \equiv 3, 7, 12, 22 \\ d+3 & t \equiv 17 \\ d-1 & t \equiv 10. \end{cases} \pmod{24}$$

For $t \leq 42$ we instead use the computer (see Sect. 5.1) to obtain $\mathrm{cap}_X(t)$ and conclude that Claim 97 holds also in this range.

Since Claims 96 and 97 are now proven, so is Claim 95, and thus the proof of Theorem 94 is complete.

5 Mathematica Code

In this section we explain how we experimentally identified the infinite staircases of Theorems 2, 54, and 56, and the properties of the blocking classes discussed in Sect. 2.3, particularly Proposition 42.

We first had to produce code which approximated c_{H_b} in a reasonable amount of time and identified the obstructions from different exceptional classes. In Sects. 5.1–5.3 we explain how to plot a lower bound $c^{\leq}_{H_b}$ to c_{H_b}, the obstruction from a single ECH capacity $c_k(X_b)$, and the obstruction $\mu_{\mathbf{E},b}$ from a single exceptional class $\mathbf{E}$ using Mathematica. In Sect. 5.1 we explain the theoretical framework of our code, while Sects. 5.2 and 5.3 primarily contain the code itself. Our methods allow us to use 25,000 ECH capacities to plot $c^{\leq}_{H_b}$ in a reasonable amount of time (e.g., the algorithm producing the kth ECH capacity is $O(k)$). They also allow us to visually compare the obstructions arising from ECH capacities with those from exceptional classes, as explained in Sect. 4.2.

Very similar methods allow for the computation of the capacities and ellipsoid embedding functions for polydisks, with small changes only to the formulas in the functions `LatticePts` and `Action`. These changes are due to the fact that the paths Λ and the region Ω determining the polydisk as a toric domain will have sides of different slopes than those of the analogous paths and regions in the case of H_b. The obstructions from the exceptional classes of polydisks require changes in the formulas for exceptional classes and obstructions: see [19, §1.1] for the case of polydisks and [7, §2.2] for the case of general rational toric domains. Note that our code is optimized for cases where Ω has two sides in addition to the sides on the axes, as the Diophantine equations solved by the function `Index` become much

more computationally expensive as the number of sides of Ω increases. This makes the code from [7, Appendix C] more useful when Ω has more than two sides in addition to the sides on the axes.

In Sect. 5.4 we explain our strategy for identifying blocking classes, approximating $Block$, and identifying elements of $Stair$ experimentally. While some of our methods are generalizations of the analysis of blocking classes from Sect. 2.3, we explain how we apply these ideas using a computer. Finally, Sect. 5.5 contains illustrations of examples discussed throughout the paper.

5.1 Computing Many ECH Capacities of X_b Quickly

Implementing Theorem 91 directly in practice is rather slow. The number of lattice points $\mathcal{L}(\Lambda)$ bounded by the lattice path Λ is a quadratic polynomial in the number of sides of Λ. Enumerating all convex lattice paths Λ with $\mathcal{L}(\Lambda) = k + 1$ requires solving the equation $\mathcal{L}(\Lambda) = k + 1$ with the path Λ having a number of sides that ranges between one and some upper bound depending on k. Once k is large enough that Λ could have three or more sides, this procedure slows dramatically.

In order to shorten our search, we constrain the slopes of the edges of the lattice paths to lie within a certain range. We then use Lemma 100, inspired by Hutchings [10, Exercise 4.16(a)], to restrict the set of lattice paths over which the minimum of (4.1.2) is taken to those with sides parallel to the vectors $(1, 0)$ and $(1, -1)$, at the expense of allowing a broader range of values of $\mathcal{L}(\Lambda)$.

Definition 98 For any nonzero vector $v \in \mathbb{R}^2$, define $\theta(v) \in (-2\pi, 0]$ to be the angle for which v is a positive multiple of $(\cos\theta(v), \sin\theta(v))$.

Note that this is not the same as the θ of in [3, Definition A.6].

The proof of the following Lemma was pointed out to us by Michael Hutchings.

Lemma 99 *The minimum in (4.1.2) can be taken over convex lattice paths Λ with $\mathcal{L}(\Lambda) = k + 1$ and $\theta(\Lambda'(t)) \in [-\pi/4, 0]$ for all t for which $\Lambda(t)$ is not a vertex of Λ.*

Proof The argument is a repeat of the first proof of Proposition 5.6 in [11].

Lemma 100 *When $\Omega = \Omega_b$, the Delzant polytope of H_b, the minimum on the right hand side of (4.1.2) is the same if it is taken over convex lattice paths Λ whose edges are parallel to the edges of Ω and for which*

$$k + 1 \leq \mathcal{L}(\Lambda) \leq 2k + 1. \tag{5.1.1}$$

Remark 101 The first conclusion of Lemma 100 (i.e., without the bound (5.1.1)) is inspired by Hutchings [10, Exercise 4.16(a)], which holds for convex domains in T^*T^2.

Proof By Lemma 100, if $\Lambda(t)$ is not a vertex and $\theta(\Lambda'(t)) \notin \{0, -\pi/4\}$ then $\theta(\Lambda'(t)) \in (0, -\pi/4)$. Let (t_0, t_1) denote the interval on which $\theta(\Lambda'(t)) \in (0, -\pi/4)$, assuming it is nonempty. Let $\Lambda(t_0) = (x_0, y_0)$ and $\Lambda(t_1) = (x_1, y_1)$.

Denote by Λ_Ω the new convex lattice path obtained by replacing $\Lambda|_{[t_0,t_1]}$ with the convex path consisting of two edges, one parallel to $(1, 0)$ followed by one parallel to $(1, -1)$. This means the only vertex of Λ_Ω with both coordinates greater than zero is the point $(x_1 - (y_0 - y_1), y_0)$; it has positive coordinates because $\theta(\Lambda')|_{(t_0,t_1)} \in (0, -\pi/4)$, therefore $x_1 - x_0 > y_0 - y_1$ and $y_0 > y_1 \geq 0$.

The new path Λ_Ω is a lattice path, because $x_1 - x_0$ and $y_0 - y_1$ are integers. While the argument until this point could apply to more general Ω, it would not always guarantee that Λ_Ω has vertices at lattice points.

The number of lattice points between Λ_Ω and the axes is at least $\mathcal{L}(\Lambda)$. It remains to show that it is at most $2\mathcal{L}(\Lambda) - 1$. It suffices to consider the case $x_0 = 0$, $y_1 = 0$. We have

$$\mathcal{L}(\Lambda_\Omega) = (x_1 - y_0 + 1)(y_0 + 1) + \frac{y_0(y_0 + 1)}{2}$$

while for $\mathcal{L}(\Lambda)$, we have is at least the number of lattice points between the line from $(0, y_0)$ to $(x_1, 0)$ and the axes. That is,

$$\begin{aligned}\mathcal{L}(\Lambda) &\geq \#\{\text{lattice points between the line from } (0, y_0) \text{ to } (x_1, 0) \text{ and the axes}\} \\ &\geq \frac{(x_1 + 1)(y_0 + 1)}{2}\end{aligned}$$

Therefore if $y_0 \geq 1$,

$$\mathcal{L}(\Lambda_\Omega) \leq (x_1 + 1)(y_0 + 1) - 1 \leq 2\mathcal{L}(\Lambda) - 1$$

If $y_0 = 0$, then $\Lambda_\Omega = \Lambda$ because Λ consists of a single edge from $(0, 0)$ to $(x_1, 0)$.

5.2 Obstructions from Single ECH Capacities and a Lower Bound for c_{H_b}

Recall that

$$Stair := \{b \mid H_b \text{ has a staircase}\} \subset [0, 1), \quad \text{and}$$
$$Block := \bigcup \left\{ J_{\mathbf{B}} \;\middle|\; \mathbf{B} \text{ is a blocking class} \right\} \subset [0, 1).$$

It its not possible for a computer to compute the right hand side of (4.0.1) for general b. However, understanding a close lower bound for c_{H_b} allowed us to approximate the set $Block$ and to identify many of the blocking classes contributing

to $Block$. Define

$$c^{\leq}_{H_b}(z) := \max_{k=1,\ldots,25000} \frac{c_k(E(1,z))}{c_k(X_b)}$$

Then we have

$$Block^{\leq} := \left\{ b \in [0,1) \middle| c^{\leq}_{H_b}(\mathrm{acc}(b)) > V_b(\mathrm{acc}(b)) \right\} \subset Block$$

because for all z,

$$c^{\leq}_{H_b}(z) \leq c_{H_b}(z)$$

In this subsection we explain our code to plot $c^{\leq}_{H_b}$ and the obstructions $\frac{c_k(E(1,z))}{c_k(X_b)}$ from a single ECH capacity.

Let Λ be a convex lattice path with edges parallel to the vectors $(1,0)$ and $(1,-1)$, i.e., the edges of Ω_b. The set of all such lattice paths is in bijection with the first quadrant, inclusive of the axes, via the Cartesian coordinates (x,y) of the vertex where the edge parallel to $(1,0)$ ends and the edge parallel to $(1,-1)$ begins. In these coordinates, if Λ corresponds to (x,y), we compute $\mathcal{L}(\Lambda)$ using

```
LatticePts[x_,y_]:=(x+1)(y+1)+(y(y+1))/2
```

We compute the set of lattice paths Λ with $\mathcal{L}(\Lambda) = k+1$ using

```
Index[k_]:=Values[Solve[LatticePts[x,y] == k+1 && 0≤x && 0≤y,
{x, y},
  Integers]]
```

We compute the set of lattice paths Λ with sides parallel to $(1,0)$ and $(1,-1)$ and $\mathcal{L}(\Lambda)$ at most 50,001, and give it the name `LatticePaths50000`. By Lemma 100, this will allow us to compute the first 25,000 ECH capacities of H_b.

```
LatticePaths50000=Table[{k,Index[k]},{k, 0, 50000}]
```

Running `LatticePaths50000` takes approximately 20 min on a personal laptop. However, it's a one-time cost: plotting each c_{H_b} takes far less time.

To compute the action of a lattice path Λ corresponding to (x,y), we compute its Ω_b-length. When $\Omega = \Omega_b$, the definition (4.1.1) is equivalent to

$$\ell_{\Omega_b}(x,y) = x(1-b) + y$$

therefore to compute the action of Λ we use

```
Action[b_]:=Function[{x,y},x(1-b)+y]
```

Instead of computing the actions of all the generators from `LatticePaths50000` and minimizing in one step, it is much faster to break the process up into the following three functions:

```
ActionList50000[b_]:=Table[Table[Action[b]
  @@LatticePaths50000[[k+1]][[2]][[i]],
  {i,Length[LatticePaths50000[[k+1]][[2]]]},{k,0,50000}]
MinActionList50000[ALb_]:=Array[Min[ALb[[#]]]&,50001]
```

```
CapacitiesList[MALb_]:=Join[{0.},Table[Min[Array[MALb[[#]]&,k,
k + 1]], {k,1,25000}]]
```

Because the index origin is set to be $k+1$, `CapacitiesList` minimizes over the entries $k+1, \ldots, 2k+1 = 2(k+1)-1$ of `MALb`, which Lemma 100 guarantees are all we need to consider. In order to compute the capacities of X_b, it is fastest to run each command in succession, applying `MinActionList50000` to the output of `ActionList50000`, then `CapacitiesList` to the output of `MinActionList50000`, instead of running their composition. It is also much faster to treat b as a decimal number rather than an exact number. Note that this is the point in the process where it is possible to introduce imprecision. If the plot of $c^{\leq}_{H_b}$ looks close to but cannot for some reason be less than or equal to the actual plot of c_{H_b}, the rounding inherent in treating b as a decimal is probably to blame. This does happen for known infinite staircases at extremely small scales.

To compute the capacities of the ellipsoid $E(1,z)$, we use

```
EllipsoidCap:=Compile[{{z,_Real},{K,_Integer}},Take[Sort[Flatten
[Table[ x+z*y,{x,0,Ceiling[z((4/z)(-z-1/2+Sqrt[(z+1/2)^2-z+z*K/2]
))]}, {y,0,Ceiling[(4/z)(-z-1/2+Sqrt[(z+1/2)^2-z+z*K/2])]}]],Less]
,K+1]]
```

The bounds on x and y in `EllipsoidCap` come from the relationship between the sequence $N(1,z)$ and triangles in the lattice.

In order to compute the obstruction $\frac{c_k(E(1,z))}{c_k(X_b)}$ from the kth ECH capacity, use the function`SingleCapacityObstructionFxn` with `EllipsoidCap[z,K]` (for $K \geq k+1$) as `ECzk`, and the list of 25,000 capacities previously computed for X_b as `CLbK`:

```
SingleCapacityObstructionFxn[ECzK_,CLbK
_,k_]:=ECzK[[k+1]]/CLbK[[k+1]]
```

To plot `SingleCapacityObstructionFxn` over an interval from `zmin` to `zmax` as a line plot with step `zstep`, plug the list of capacities for X_b in as `CLbK` with `k` at most the length of `CLbK` into

```
PlotSingleCapacityObstructionFxn[zmin_,zmax_,zstep_,CLbK_,k_,b_]
:=Show[ ListLinePlot[Array[{(#-1)*zstep+zmin,SingleCapacityObstr
uctionFxn[ EllipsoidCap[(#-1)*zstep+zmin,k+5],CLbK,k]}&,
Floor[(zmax-zmin)/zstep]], PlotStyle→Cyan]]
```

To compute $c^{\leq}_{H_b}(z)$, find the maximum over the first 25,000 capacities of $\frac{c_k(E(1,z))}{h}$, where h is the $k+1^{\text{st}}$ element of the output of `CapacitiesList` (remember the first element will be $c_0(X_b)$). This is done by plugging `EllipsoidCap[z,25,000]` as `ECzk`, the list of 25,000 capacities previously computed for X_b as `CLbK`, and $K = 25{,}000$ into

```
CapacityFxn[ECzK_,CLbK_,K_]:=Max[Array[ECzK[[#+1]]/CLbK[[#+1]]&,
K]]
```

In order to plot `CapacityFxn` on [zmin,zmax] as a line plot with step `zstep`, plug your list of capacities for X_b as `CLbK` with $K = 25{,}000$ into

```
PlotCapacityFxn[zmin_,zmax_,zstep_,CLbK_,K_,b_]:=
  Show[Plot[Sqrt[(z/2)/(.5(1-b^2))],
```

```
{z,zmin,zmax},PlotStyle→Orange],ListLinePlot[
Array[{(#-1)*zstep+zmin,CapacityFxn[EllipsoidCap[(#-1)*zstep+
zmin,K],
CLbK,K]}&,Floor[(zmax-zmin)/zstep]]]]
```

Starting with `zstep` at 0.01 is a good choice if the interval is $[1, 10]$; when analyzing the behavior near the accumulation point, with the interval much smaller, it is possible to make `zstep` much smaller. We frequently use a `zstep` of 0.000001.

5.3 Obstructions from Exceptional Classes

We now discuss our code computing $\mu_{\mathbf{E},b}$ for exceptional classes $\mathbf{E}$.

It is often useful to understand obstructions which do not correspond in the sense of Remark 93 to any of the first 25,000 ECH capacities. As in the case of identifying lattice paths, finding all relevant exceptional classes takes a lot of time and effort, whether done by a computer or a human. But once $\mathbf{E}$ is identified, computing $\mu_{\mathbf{E},b}(z)$ can be done quickly. We explain our code for this here.

Plotting the obstruction $\mu_{\mathbf{E},b}$ from a single exceptional class $\mathbf{E}$ first requires computing the weight expansion of a number a, using

```
WeightDecomp[p_,q_]:=Module[{list={},aux,bux,k},If[p>q,{aux,bux}=
{p,q}, {aux,bux}={q,p}];While[Chop[bux]≠0,k=Floor[aux/bux];
AppendTo[list, Table[bux,k]];{aux,bux}={bux,aux-k*bux}];
Flatten[list]]
```

If $z = p/q$ is rational, then `WeightDecomp[p,q]` computes $q\mathbf{w}(z)$, while `WeightDecomp[z,1]` computes $\mathbf{w}(z)$.

We take the dot product of vectors of unequal length using

```
ExtendedDot[M_,w_]:=Take[M,Min[Length[M],Length[w]]].Take[w,Min[
Length[M],Length[w]]]
```

For $\mathbf{E} = (d, m, \mathbf{m})$ we compute $\mu_{\mathbf{E},b}(z)$ by setting `d` $= d$, `m` $= m$, `M` $= \mathbf{m}$, and `z` $= z$ in

```
Mu[b_,d_,m_,M_,z_]:=ExtendedDot[M,WeightDecomp[a,1]]/(d-b*m)
```

We plot `Mu` for $z \in$ [`zmin`, `zmax`] with step `zstep` using

```
PlotMu[zmin_,zmax_,zstep_,d_,m_,M_,b_]:=Show[ListLinePlot[Array[
{(#-1)*zstep+zmin,Mu[b,d,m,M,(#-1)*zstep+zmin]}&,Floor[(zmax-
zmin)/ zstep]],PlotStyle→Brown]]
```

In Fig. 7 we have used `PlotMu` to depict the obstructions from the exceptional classes whose obstructions underlay the steps centered at 1219/208, 3194/545, and 8363/1427 in the infinite staircase for $b = b_0^E$ of Theorem 58. Using the procedure `PlotMu` instead of the procedure `PlotSingleCapacityObstructionFxn` allows us to plot these stairs, as the corresponding ECH capacities are the 127,489th, 872,234th, and 5,971,895th, respectively, which do not contribute to $c^{\leq}_{H_{b_0^E}}(z)$. In order not to overload the figure, we have not drawn the obstruction coming from the classes centered at 170/29, 463/79, whose centers are greater than

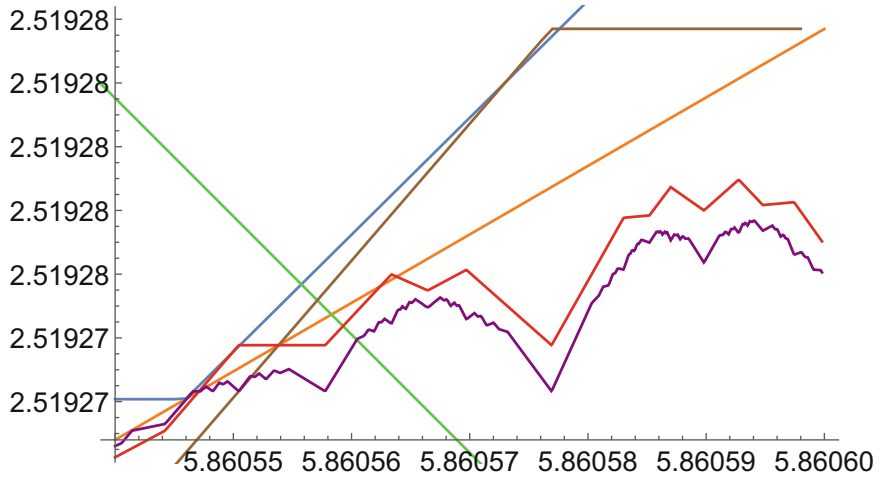


Fig. 7 This is a plot of the capacity and volume functions for $b = b_0^E = \mathrm{acc}_L^{-1}([5;1,6,\{5,1\}^\infty])$. The function $c^{\leq}_{H_{b_0^E}}(z)$ is in dark blue, and in this case it is not a good approximation of the true $c_{H_{b_0^E}}(z)$. The volume obstruction is in orange. The green curve is parameterized by $b \mapsto \left(\mathrm{acc}(b), \sqrt{\frac{\mathrm{acc}(b)}{1-b^2}}\right)$, therefore the green curve intersects the orange when $z = \mathrm{acc}(b_0^E)$. Using $\mathbf{E}_{u,0,k}$ to denote the exceptional class with center $[5;1,6,\{5,1\}^k,4]$ and $\mathbf{E}'_{u,0,k}$ to denote the exceptional class with center $[5;1,6,\{5,1\}^k,5,2]$, the obstruction $\mu_{\mathbf{E}_{u,0,1},b_0^E}(z)$ is in brown, the obstruction $\mu_{\mathbf{E}'_{u,0,1},b_0^E}(z)$ is in red, and the obstruction $\mu_{\mathbf{E}_{u,0,2},b_0^E}(z)$ is in purple. Because these correspond, in the sense of Remark 93, to the obstructions from the 127,489th, 872,234th, and 5,971,895th ECH capacities, respectively, they are not captured by our program computing $c^{\leq}_{H_{b_0^E}}$: in fact, on this interval $c^{\leq}_{H_{b_0^E}}$ is given instead by the obstruction from $c_8(H_{b_0^E})$

the values of z plotted, or the obstructions from the class centered at 21895/3736, or any obstructions from Theorem 58 with $k \geq 3$, which Mathematica cannot plot at any scale with any definition.

In general, once we have an exceptional class in mind, it is computationally faster to run `PlotMu` instead of `PlotSingleCapacityObstructionFxn`, because in order to run the latter we first need to compute all the ECH capacities up to c_k, where $k = \frac{1}{1}(d(d+3)-m(m+1))$. If we are interested in an exact irrational value of b, there is a great difference, although we almost never do this. However, if we want to overlay our plots on a graph of $c^{\leq}_{H_b}$, there is no discernible difference, because once we have computed the first 25,000 capacities of X_b necessary to graph $c^{\leq}_{H_b}$, any difference in speed between `PlotMu` and `PlotSingleCapacityObstructionFxn` is not noticeable.

5.4 Strategy for Finding Staircases

In this section we explain our experimental strategy for identifying infinite staircases. We first explain our method for approximating $Block^{\leq}$, then our two methods for finding staircases once we understood much of $Block^{\leq}$.

We investigate $Block^{\leq}$ by identifying the exceptional class realizing the values of k for which $c_k(X_b)$ determines $c^{\leq}_{H_b}(\mathrm{acc}(b))$. At first, we simply chose values of b at random and identified the smallest k for which the plot from running `PlotSingleCapacityObstructionFxn` agreed with the plot from `PlotCapacityFxn` for values of z near $\mathrm{acc}(b)$. We encode $\mathrm{acc}(b)$ using
`AccPt[b_]:=(1/2)((3-b)^2/(1-b^2)-2+Sqrt[((3-b)^2/(1-b^2)-2)^2-4])`

We identified the appropriate values of k by visually inspecting the plots of the obstruction from c_k overlaid upon the plots of $c^{\leq}_{H_b}$. See Fig. 8.

Inspired by our computations, when $c_{H_b}(\mathrm{acc}(b)) > \sqrt{\frac{\mathrm{acc}(b)}{1-b^2}}$ and k satisfies

$$\frac{c_k(E(1,\mathrm{acc}(b)))}{c_k(X_b)} = c_{H_b}(\mathrm{acc}(b))$$

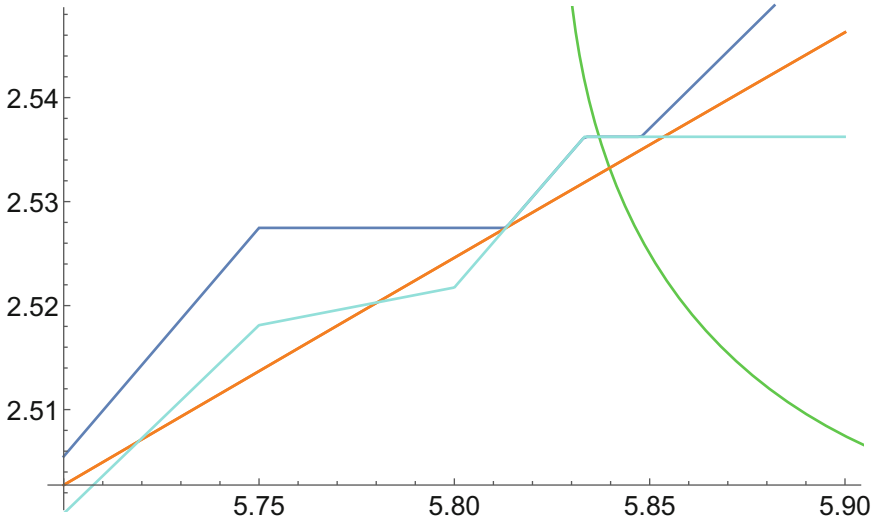


Fig. 8 Here, $b = 0.3$. The function $c^{\leq}_{H_{0.3}}(z)$ is in dark blue, the volume obstruction is in orange, and the obstruction from the 125th ECH capacity is in bright blue. The green curve is parameterized by the potential accumulation point, therefore the orange and green curves intersect when $z = \mathrm{acc}(0.3)$, and we can see that the dark blue and the light blue plots coincide in a neighbourhood of that point. Indeed, 125 is the smallest k for which c_k obstructs an infinite staircase at $b = 0.3$. Note that $c^{\leq}_{H_{0.3}}(z)$ has a corner point at $z = 35/6$, and near this value is given by the class $\mathbf{E} = \big(15, 4; 6\mathbf{w}(35/6)\big)$ as described in Remark 19 (ii)

we say c_k **obstructs an infinite staircase at** b.

For values of b very close to one for which c_{H_b} has an infinite staircase, the smallest k for which c_k obstructs an infinite staircase at b becomes very high, so it grows infeasible to check every possible value of k by inspection of plots such as the one depicted in Fig. 8. We introduced a new function, `IndexAtAccPt[b_,CLbK_,K_]`. Again, the list `CLbK` is the list of the first 25,000 capacities of H_b. `IndexAtAccPt` outputs the values of $k \in \{0, \ldots, K\}$ for which $\frac{c_k(E(1,\mathrm{acc}(b)))}{c_k(X_b)}$ is maximized. The reason we want to be able to use values of $K < 25{,}000$ is because when $K = 25{,}000$, `IndexAtAccPt` is fairly slow.

```
IndexAtAccPt[b_,CLbK_,K_]:=Block[{list=Array[EllipsoidCap[
  AccPt[b],K][[#+1]]/CLbK[[#+1]]&,K]},Position[list,Max[list]]]
```

Given k and b such that c_k obstructs an infinite staircase at b, we also want to identify any quasi-perfect exceptional classes giving the same obstruction, that is, those "corresponding" exceptional classes in the sense of Remark 93. For low values of k, this can be done by hand, see Remark 19 (ii). For large values of k, we do this by applying the function `FindPQFromDM` to the values of the output of `FindDMFromK`, where

```
FindDMFromK[k_]:=Solve[.5(d(d+3)-m(m+1))==k&&d≥0&&m≥0,{d,m}
,Integers]
FindPQFromDM[d_,m_]:=Solve[d^2+1-m^2==p*q&&3d-m==p+q&&p>q&&q≥0&&
GCD[p,q]==1,{p,q},Integers]
```

Through the application of these two functions we have occasionally found multiple candidate quasi-perfect classes, but it has always turned out that if there are any solutions, then exactly one makes sense for the value of b under consideration.

We expect that for many k, the ECH capacity c_k obstructs an infinite staircase for an interval of values of b. In our computations, for each k where c_k obstructs an infinite staircase, we have been able to find a perfect center-blocking class $\mathbf{B}$ corresponding to c_k as in Remark 93. Therefore, for every pair $b_1 \leq b_2$ we have found whose infinite staircases are obstructed by the same c_k, we know by Lemma 38 that

$$[b_1, b_2] \subset J_{\mathbf{B}},$$

and therefore

$$[b_1, b_2] \subset Block^{\leq} \subset Block.$$

By identifying many such pairs b_1 and b_2, we built up a good approximation to $Block^{\leq}$. Next we searched for values of b whose embedding function c_{H_b} contained infinite staircases outside our approximation to $Block^{\leq}$. We explain our two methods for doing so.

Checking Rational Points in $[0, 1) \setminus$ ***Block***$^{\leq}$ Suppose we know that c_k and $c_{k'}$ obstruct infinite staircases at b, b' respectively, with $b < b'$, and that we have checked enough values $b'' \in (b, b')$ at small enough resolution to convince ourselves

that there is no $k \in \{1, \ldots, 25{,}000\}$ for which c_k obstructs an infinite staircase at any $b'' \in (b, b')$. In terms of our computations, this last condition means that for all $b'' \in (b, b')$ we have checked, we have $c^{\leq}_{H_{b''}}(\text{acc}(b'')) \leq V_{b''}(\text{acc}(b''))$.

We would like to search through all rational numbers $p/q \in (\text{acc}(b), \text{acc}(b'))$ (if $b, b' \geq 1/3$; if $b, b' \leq 1/3$ then we check the interval $(\text{acc}(b'), \text{acc}(b))$) to see if there are any d, m for which $(d, m; q\mathbf{w}(p/q))$ solves the Diophantine equations (2.1.2) and whose continued fractions contain repeated patterns. We do this using the function `ActualClassesWithCF` applied to $(\text{acc}(b), \text{acc}(b'), q1, q2)$, which searches through all such rational numbers with $q1 \leq q \leq q2$ (when $b, b' \leq 1/3$, we switch the places of $\text{acc}(b)$ and $\text{acc}(b')$ accordingly). To define this function we need the preliminary function `NonemptyClassesWithCF`:

```
NonemptyClassesWithCF[LB_,UB_,q1_,q2_]:=Select[Table[Table[
{Flatten[{ Values[Solve[d^2+1-m^2==pq&&3d-m==p+q&&d≥0&&m≥0,{d,m},
Integers]],p,q}], ContinuedFraction[p/q]},{p,Ceiling[LBq],
Floor[UBq]}],{q,q1,q2}],#≠ {}&]
ActualClassesWithCF[LB_,UB_,q1_,q2_]:=Select[Flatten[
  NonemptyClassesWithCF[LB,UB,q1,q2],1],Length[#[[1]]]==4&]
```

`ActualClassesWithCF` returns lists with terms of the form $((d, m, p, q), CF(p/q))$ where $(d, m, q\mathbf{w}(p/q))$ solves the Diophantine equations (2.1.2), the rational $p/q \in (\text{acc}(LB), \text{acc}(UB))$ has $q1 \leq q \leq q2$, and $CF(p/q)$ is the continued fraction of p/q. If we see any terms whose $CF(p/q)$ part contains a periodic piece, we check whether there are any solutions to (2.1.2) for the obvious extension of the periodic pattern, and if so, whether the associated exceptional classes reduce properly under Cremona transformations.

This method works well when the periodic pieces are short. To handle the more complex continued fractions investigated in our next paper, we developed our second method, explained below, which also illuminates the relationship between infinite staircases and blocking classes.

Finding the Staircases from a Blocking Class B Our second method provides insight into the relevance of periodic continued fractions. Often the denominators of the rationals p/q that we need to check in order to find a pattern are extremely large, making our first method quite slow. However, the first method has the advantage that we also learn the ends end_n differentiating between the continued fractions of the centers of the classes contributing to the individual stairs and the continued fraction of the accumulation point. In contrast, with our second method we can solve for the accumulation point precisely but need to search or guess to find the individual stairs.

Assume $\mathbf{B} = (d, m; q\mathbf{w}(p/q))$ blocks an interval $J_{\mathbf{B}}$ with $I_{\mathbf{B}} = (\alpha_{\mathbf{B},\ell}, \alpha_{\mathbf{B},u})$. First we identify $\alpha_{\mathbf{B},\ell}$ and $\alpha_{\mathbf{B},u}$—these will be the accumulation points of the two infinite staircases determined by $\mathbf{B}$ as in Conjecture 4. Our procedure relies on the computer so that it will work for complicated $\mathbf{B}$; note however that for simpler $\mathbf{B}$, it may be possible to find $J_{\mathbf{B}}$ by hand using Lemma 44.

To determine $\alpha_{\mathbf{B},\ell}$, let $a_\ell(b)$ be the nonzero solution in z to

$$\frac{qz}{d-bm} = \sqrt{\frac{z}{1-b^2}} = V_b(z). \tag{5.4.1}$$

Note that the left hand side of (5.4.1) is the value of $\mu_{\mathbf{B},b}(z)$ for z close to and smaller than p/q, by Lemma 16. Then let b_ℓ be the solution in b to

$$\mathrm{acc}(b) = a_\ell(b)$$

with $0 \le b < 1$ and in the same component of $[0,1) - \{1/3\}$ as $J_{\mathbf{B}}$. Then $\alpha_{\mathbf{B},\ell} = \mathrm{acc}(b_\ell)$. Because $\alpha_{\mathbf{B},\ell}$ is a quadratic irrationality, its continued fraction is periodic.

With $\alpha_{\mathbf{B},\ell}$ in hand, we can now attempt to guess the endings of the continued fractions of the centers of the classes contributing to the staircase $\mathcal{S}_{\beta_{\mathbf{B},\ell}}$ (if $J_{\mathbf{B}} \subset (1/3,1)$) or $\mathcal{S}_{\beta_{\mathbf{B},u}}$ (if $J_{\mathbf{B}} \subset [0,1/3)$). We do this using the function `SearchingForEnding`, which uses the function `FindDMFromCF`:

```
FindDMFromCF[CF_]:=Solve[d^2-m^2+1==Numerator
  FromContinuedFraction[CF]]*
  Denominator[FromContinuedFraction[CF]]&&3d-m==Numerator[
  FromContinuedFraction[CF]]+Denominator
  [FromContinuedFraction[CF]]&&
  d≥0&&m≥0,{d,m},Integers]
SearchingForEnding[CF_,end_]:=FindDMFromCF[Join[CF,end]]
```

By constructing tables consisting of `SearchingForEnding[CF,end]` with `end` varying over judiciously chosen tuples of length up to four, we were able to identify the necessary ends. This method always contains some element of guess-and-check, but our general rule of thumb is that if $CF(\alpha_{\mathbf{B},\ell}) = [\mathbf{X}, \{\mathbf{Y}\}^\infty]$ then there are ways to write $\mathbf{Y} = (\mathbf{y_1}, \mathbf{y_2})$ and $\mathbf{Y} = (\mathbf{y'_1}, \mathbf{y'_2})$ so that `SearchingForEnding[CF,end]` returns nonempty pairs (d, m) for

$$\mathtt{CF} = \{\mathbf{X}, \{\mathbf{Y}\}^k\}, \qquad \mathtt{end} = \text{exactly one of } \{\mathbf{y_1}, 2n+2\}, \{\mathbf{y_1}, 2n+4\}, \{\mathbf{y_1}, 2n+6\}$$

and for

$$\mathtt{CF} = \{\mathbf{X}, \{\mathbf{Y}\}^k\}, \qquad \mathtt{end} = \text{exactly one of } \{\mathbf{y'_1}, 2n+2\}, \{\mathbf{y'_1}, 2n+4\}, \{\mathbf{y'_1}, 2n+6\}.$$

Identifying the "opposite" staircase $\mathcal{S}_{\beta_{\mathbf{B},\ell}}$ (if $J_{\mathbf{B}} \subset [0,1/3)$ or $\mathcal{S}_{\beta_{\mathbf{B},u}}$ (if $J_{\mathbf{B}} \subset (1/3,1)$) is done in a similar manner, but using $\alpha_{\mathbf{B},u}$, which is identified as follows. Let $a_u(b)$ be the solution in z to

$$\frac{p}{d-bm} = \sqrt{\frac{z}{1-b^2}} = V_b(z). \tag{5.4.2}$$

Note that the left hand side of (5.4.2) is the value of $\mu_{\mathbf{B},b}(z)$ for z close to and larger than p/q. Then let b_u be the solution in b to

$$\mathrm{acc}(b) = a_u(b)$$

which is in the same component of $[0, 1) - \{1/3\}$ as $J_{\mathbf{B}}$. Then $\alpha_{\mathbf{B},u} = \mathrm{acc}(b_u)$ and is again a quadratic irrationality, with a periodic continued fraction. We identify endings as for $\alpha_{\mathbf{B},\ell}$ above.

With either method we could only identify staircases for single values of n, but after seeing enough staircases we could usually guess fairly easily how they ought to generalize to higher n. We were able to discover all the two-periodic staircases discussed in this paper using the methods of Sect. 5.4. The methods of Sect. 5.4. became necessary for all but the simplest $2k$-periodic staircases of our next paper with $k > 1$.

5.5 *Plots of Ellipsoid Embedding Functions*

In this section we give figures illustrating some of the phenomena mentioned throughout the paper. In the following, all plots show $\mu_{\mathbf{E},b}$ overlaid atop $\frac{c_k(E(1,z))}{c_k(X_b)}$ on the interval where they agree.

Perfect Versus Quasi-Perfect Figures 9 and 10 illustrate Lemma 15. Figure 9 shows that the perfect class $\mathbf{E}' = (73, 20; 29\mathbf{w}\,(170/29))$ of Example 28(iii) is live at $170/29$, as guaranteed by Proposition 21 (i).

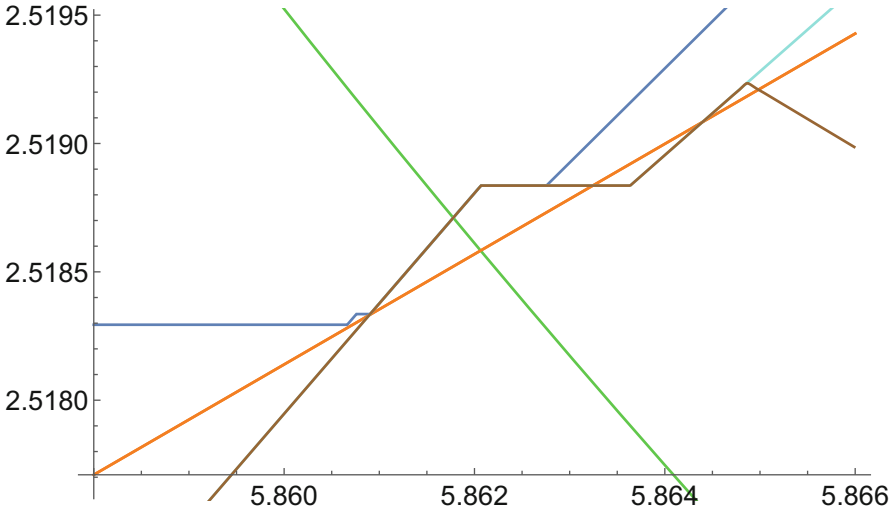


Fig. 9 Here $b = \mathrm{acc}_L^{-1}(170/29) \approx 5.862$. The function $c^{\leq}_{H_b}$ is in dark blue, the volume obstruction is in orange, the obstruction from the 2564th ECH capacity is in bright blue, and the obstruction from the perfect class $\mathbf{E}' = (73, 20; 29\mathbf{w}\,(170/29))$ of Example 28(ii) is in brown; notice these last two agree on an interval including the center $170/29 \approx 5.86207$ of $\mathbf{E}'$. The green curve is as before, crossing the orange curve with z-coordinate $170/29$. This diagram also exhibits $\mathbf{E}'$ (which is a step in the staircase $\mathcal{S}^L_{u,0}$) as a blocking class, illustrating Proposition 49

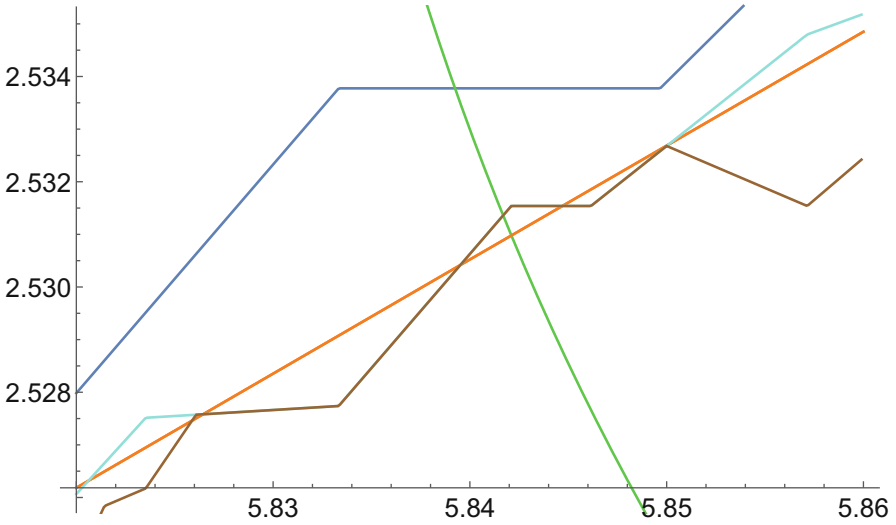


Fig. 10 Here $b = \mathrm{acc}_L^{-1}(111/19)$. The function $c_{H_b}^{\leq}$ is in dark blue, the volume obstruction is in orange, the obstruction from the 1119th ECH capacity is in bright blue, and the obstruction from the quasi-perfect class $\mathbf{E}'' = (48, 14; 19\mathbf{w}(111/19))$ of Example 28(ii) is in brown; notice these last two agree on an interval including the center $111/19$ of $\mathbf{E}''$. The green curve is as before, crossing the orange curve with z-coordinate $111/19$. For this value of b the obstruction from the 125th ECH capacity equals that from the perfect class $\mathbf{E}' = (15, 4; 6\mathbf{w}(35/6))$, and both obstructions are stronger than those from the 1119th ECH capacity and $\mathbf{E}''$ near $\frac{111}{19}$

In Fig. 10, the point $b = \mathrm{acc}_L^{-1}(111/19)$ is obstructed by the 125th capacity and the perfect class $\mathbf{E}' = (15, 4; 6\mathbf{w}(35/6)) = \mathbf{B}_0^E$ of Example 28(ii). The 1119th capacity and the quasi-perfect class $\mathbf{E}'' = (48, 14; 19\mathbf{w}(111/19))$ of Example 28(ii), while still providing an obstruction, is strictly less than $c_{H_b}^{\leq} \leq c_{H_b}$. Therefore Fig. 10 illustrates the conclusion of Lemma 15 (iii) that, for this b, the obstruction $\mu_{\mathbf{E}'',b}$ is nontrivial at $111/19$. However, there is no reason why $\mu_{\mathbf{E}'',b}$ should be live at $111/19$ since $\mathbf{E}''$ does not satisfy the reduction criterion in Lemma 63 and so is not perfect. Thus Proposition 21 does not apply; see Remark 23.

Typical Behavior of a Blocking Class In this subsection we analyze the effect of the blocking class $\mathbf{B}_0^U = (3, 2; \mathbf{w}(6))$ on c_{H_b} as b varies. Firstly, by the method outlined in Sect. 5.4, we find

$$J_{\mathbf{B}_0^U} = (\beta_{\mathbf{B}_0^U,\ell}, \beta_{\mathbf{B}_0^U,u}) = \left(\frac{3-\sqrt{5}}{2}, \frac{3(7+\sqrt{5})}{44}\right) \approx (0.381966, 0.629732),$$

and by applying acc, we find

$$I_{\mathbf{B}_0^U} = (\alpha_{\mathbf{B}_0^U,\ell}, \alpha_{\mathbf{B}_0^U,u}) \approx (5.8541, 7.17082).$$

Because $\mathbf{w}(6) = (1^{\times 6})$, the obstruction $\mu_{\mathbf{B}_0^U,b}$ is nontrivial at $z \geq 6$ if

$$\frac{6}{3-2b} > \sqrt{\frac{z}{1-b^2}} \Leftrightarrow 6 \leq z < \frac{36(1-b^2)}{(3-2b)^2},$$

and $\mu_{\mathbf{B}_0^U,b}$ is nontrivial at $5 \leq z < 6$ if

$$\frac{5+(z-5)}{3-2b} > \sqrt{\frac{z}{1-b^2}} \Leftrightarrow \frac{(3-2b)^2}{1-b^2} < z < 6.$$

While it is possible for $\mu_{\mathbf{B}_0^U,b}$ to be nontrivial for $z < 5$, we have $\mu_{\mathbf{B}_0^U,b}(5) \leq V_b(5)$ for all b, meaning that the maximal z-interval containing the center $z = 6$ of $\mathbf{B}_0^U$ on which $\mu_{\mathbf{B}_0^U,b}$ is nontrivial is $\left(\frac{(3-2b)^2}{1-b^2}, \frac{36(1-b^2)}{(3-2b)^2}\right)$.[16] This interval is nonempty so long as

$$b \in \left(\frac{6-\sqrt{6}}{10}, \frac{6+\sqrt{6}}{10}\right) \approx (0.355051, 0.844949),$$

and if nonempty it contains $z = 6$. Thus when $b = \frac{6-\sqrt{6}}{10}, \frac{6+\sqrt{6}}{10}$ we have

$$\frac{(3-2b)^2}{1-b^2} = \frac{36(1-b^2)}{(3-2b)^2} = 6.$$

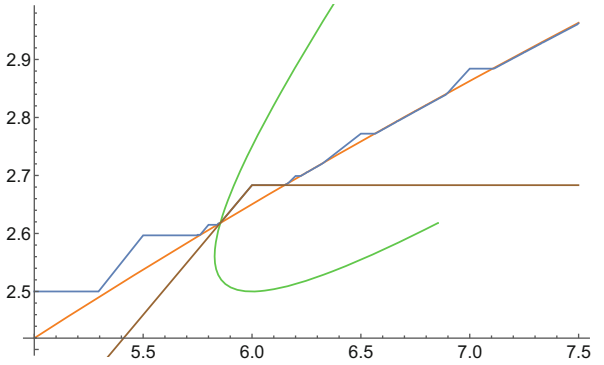
The interval $\left(\frac{(3-2b)^2}{1-b^2}, \frac{36(1-b^2)}{(3-2b)^2}\right)$ is longest when $b = 2/3 = m/d$, when it is the interval $(5, 36/5)$ and $\frac{(3-2b)^2}{1-b^2}$ reaches a minimum when $\frac{36(1-b^2)}{(3-2b)^2}$ reaches a maximum; notice that $I_{\mathbf{B}_0^U} \subset (5, 36/5)$. However, $\mathbf{B}_0^U$ does not block $b = m/d$, because $\text{acc}\,(2/3) \approx 7.66962 > 36/5$, as explained in Remark 39.

Figure 11 depicts c_{H_b} for

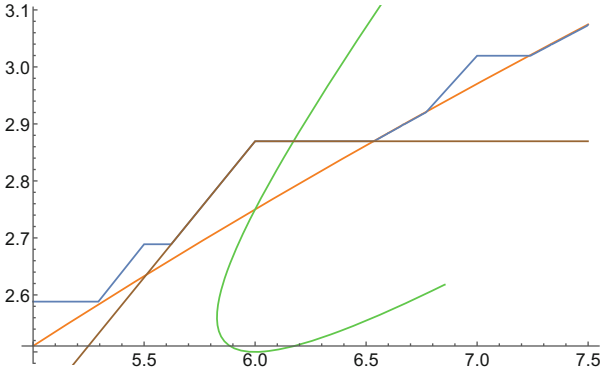
$$b = \frac{3-\sqrt{5}}{2} = \beta_{\mathbf{B}_{\ell,0}^U}, \quad b = 5/11 = \text{acc}_U^{-1}(6), \quad b = \frac{3(7+\sqrt{5})}{44} = \beta_{\mathbf{B}_{u,0}^U}$$

highlighting the obstruction $\mu_{\mathbf{B}_0^U,b}$. Notice that in (b),(c) there is a new obstruction coming in from the left that dominates $\mu_{\mathbf{B}_0^U,b}$ near the left end point of the interval where $\mu_{\mathbf{B}_0^U,b}$ is obstructive. This illustrates the point that for general b it is often the case that $\mu_{\mathbf{B}_0^U,b}$ is not live over the entire z-interval on which it is nontrivial.

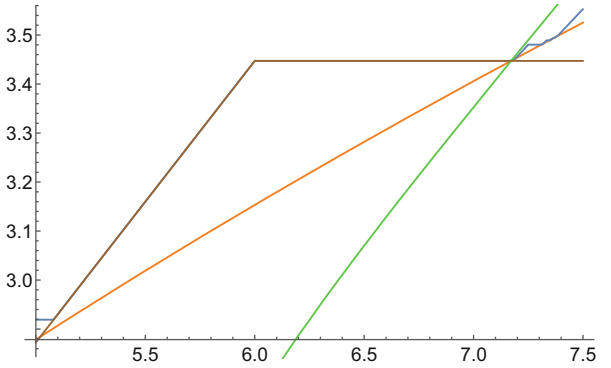
[16] Note also that by Lemma 14 (ii), if $\mathbf{B}_0^U$ is nontrivial on an interval containing points less than 5, then that interval cannot include 6, since $\ell(5) < \ell(6)$, but 6 would have to be the break point.



(a)

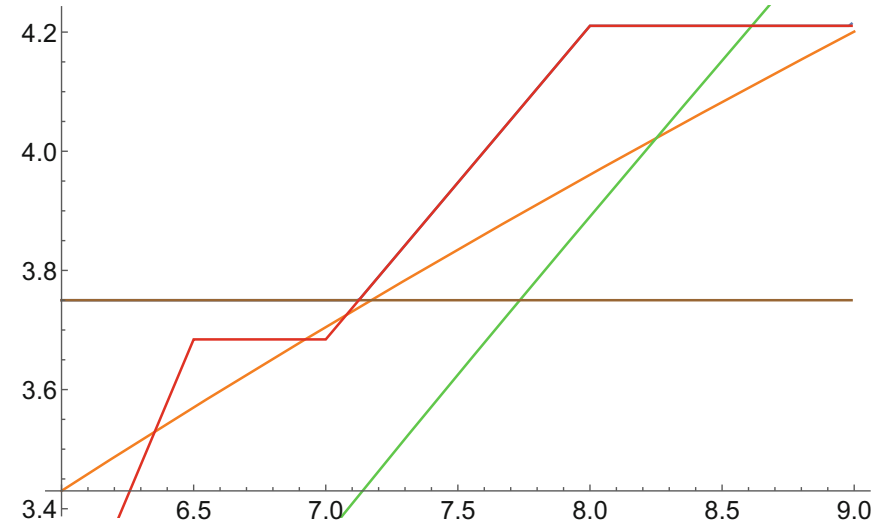


(b)

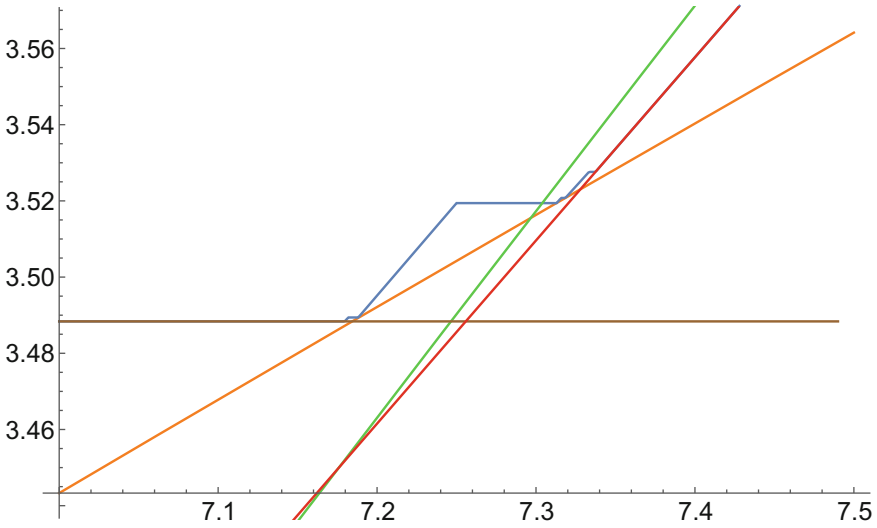


(c)

Fig. 11 As b increases from $\beta_{\mathbf{B}_0^U,\ell}$ to the center of $\mathbf{B}_0^U$ to $\beta_{\mathbf{B}_0^U,u}$, the plot of c_{H_b} varies from the infinite staircase $\mathcal{S}^U_{\ell,0}$ depicted in (**a**), through blocked staircases such as that of $c_{H_{5/11}}$ depicted in (**b**), to the infinite staircase $\mathcal{S}^U_{0,u}$ depicted in (**c**). In all plots the orange curve indicates the volume obstruction, the function $c^{\leq}_{H_b}$ is in dark blue, the obstruction $\mu_{\mathbf{B}_0^U,b}$ is in brown, and the curve $(\mathrm{acc}(b), V_b(\mathrm{acc}(b)))$ is in bright green. (**a**) $b = \beta_{\mathbf{B}_0^U,\ell} = \frac{3-\sqrt{5}}{2}$. (**b**) $b = \mathrm{acc}_U^{-1}(6) = \frac{5}{11}$. (**c**) $b = \beta_{\mathbf{B}_0^U,u} = \frac{3(7+\sqrt{5})}{44}$



(a)



(b)

Fig. 12 In both plots the orange curve indicates the volume obstruction, the function $c^{\leq}_{H_b}$ is in dark blue when visible, the obstruction $\mu_{\mathbf{B}^U_0,b}$ is in brown, the obstruction $\mu_{\mathbf{B}^U_1,b}$ is in red, and the curve $(\mathrm{acc}(b), V_b(\mathrm{acc}(b)))$ is in bright green. Note the differing scales; the break point 8 of the red curve is visible in (**a**) but not in (**b**). (**a**) $b = 0.7$. (**b**) $b = 0.64$

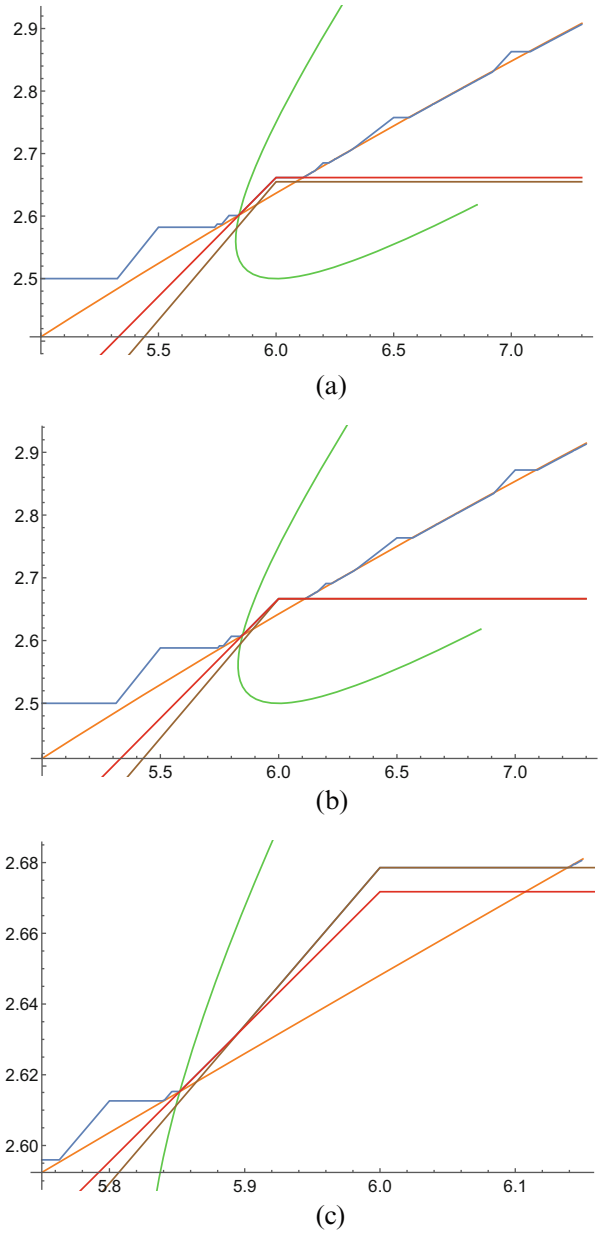


Fig. 13 In all plots the orange curve indicates the volume obstruction, the function $c^{<}_{H_b}$ is in dark blue, the obstruction $\mu_{\mathbf{B}^U_0,b}$ is in brown, the obstruction from $(3, 1; 2, 1^{\times 5})$ is in red, and the curve $(\mathrm{acc}(b), V_b(\mathrm{acc}(b)))$ is in bright green. Note the differing scales; in all diagrams, both the brown and red curves break at $z = 6$. (**a**) $b = 0.37$. (**b**) $b = \frac{3}{8} = 0.375$. (**c**) $b = 0.38$

When $b \in \left(\frac{6-\sqrt{6}}{10}, \frac{6+\sqrt{6}}{10}\right) - \left(\frac{3-\sqrt{5}}{2}, \frac{3(7+\sqrt{5})}{44}\right)$—that is, when $\mu_{\mathbf{B}_0^U,b}$ is nontrivial but $\mathbf{B}_0^U$ is not blocking—this phenomenon persists. Figure 12 illustrates the situation when $b \in \left[\frac{3(7+\sqrt{5})}{44}, \frac{6+\sqrt{6}}{10}\right) \approx (0.629732, 0.844949)$, showing the interaction between $\mathbf{B}_0^U$ and $\mathbf{B}_1^U$. When $b = 0.7$, the obstruction $\mu_{\mathbf{B}_0^U,b}$ (in brown), though nontrivial, is not live on the interval $(7.125, 7.171875)$; instead $\mu_{\mathbf{B}_1^U,b}$, where $\mathbf{B}_1^U = (4, 3; \mathbf{w}(8))$, (in red) is live. As b decreases towards $\frac{3(7+\sqrt{5})}{44}$, the intervals on which $\mu_{\mathbf{B}_0^U,b}$ (in brown) and $\mu_{\mathbf{B}_1^U,b}$ (in red) are nontrivial become disjoint, as predicted by Proposition 42, and one begins to see some staircase classes. Indeed, by Theorem 56, when $b = \beta_{\mathbf{B}_0^U,u}$, H_b admits a descending staircase $\mathcal{S}_{u,0}^U$ whose $k = 0$ step with $\text{end}_0 = 4$ has center $29/4 = [7; 4]$, the largest center amongst steps in $\mathcal{S}_{u,0}^U$. Meanwhile, the step with the second-lowest center in the ascending staircase $\mathcal{S}_{\ell,1}^U$ (the $k = 0$ step with $\text{end}_1 = (7, 4)$) has the same center; it is visible as the large dark blue step in Fig. 12b.

Figure 13 illustrates the situation when $b \in \left(\frac{6-\sqrt{6}}{10}, \frac{3-\sqrt{5}}{2}\right]$. When $b \in \left[\frac{3}{8}, \frac{3-\sqrt{5}}{2}\right)$, as in (c), there is still a z-interval on which $\mu_{\mathbf{B}_0^U,b}$ (in brown) is live, but when $b \in \left(\frac{6-\sqrt{6}}{10}, \frac{3}{8}\right)$, as in (a), the obstruction from the class $(3, 1; 2, 1^{\times 5})$ (in red) overwhelms $\mu_{\mathbf{B}_0^U,b}$. However, note that $(3, 1; 2, 1^{\times 5})$ is never blocking in this interval, because the value of its obstruction function at $\text{acc}(b) < 6$ is always $V_b(\text{acc}(b))$, as we saw in (2.2.5). In particular, in each of (a), (b), (c) the orange, green and red curves have a common point of intersection.

Acknowledgments We are indebted to the organizers (Bahar Acu, Catherine Cannizzo, Dusa McDuff, Ziva Myer, Yu Pan, and Lisa Traynor) of the Women in Symplectic and Contact Geometry and Topology (WiSCon) network for bringing us together at a research collaboration workshop in July 2019. ICERM was a gracious host for our week in Providence. We would also like to thank Dan Cristofaro-Gardiner, Michael Hutchings, and Ben Wormleighton for helpful conversations and lemmata.

MB was supported by the grant DFG-CRC/TRR 191 "Symplectic Structures in Geometry, Algebra and Dynamics".

TSH was supported by NSF grant DMS 1711317 and is grateful for the hospitality of Clare Hall during 2019–2020.

EM was supported by the EPSRC grant Centre for Doctoral Training, *London School of Geometry and Number Theory*.

GTM was supported by Fundação para a Ciência e a Tecnologia (FCT) through the grant PD/BD/128420/2017.

MW was supported by NSF grant RTG–1745670.

References

1. M. Beck and S. Robins, *Computing the continuous discretely: integer point enumeration in polyhedra*, Springer, Berlin, 2007.

2. R. Casals and R. Vianna, "Sharp Ellipsoid Embeddings and Toric Mutations", arxiv:2004.13232.
3. D. Cristofaro-Gardiner, "Symplectic embeddings from concave toric domains into convex ones", *J. Diff. Geom.*, 112(2):199–232, 2019.
4. D. Cristofaro-Gardiner, "Special eccentricities of rational four-dimensional ellipsoids", arXiv:2004.13647.
5. D. Cristofaro-Gardiner and R. Hind, "Symplectic embeddings of products", *Comm. Math. Helv.*, 93(1):1–32, 2018.
6. D. Cristofaro-Gardiner, R. Hind and D. McDuff, "The ghost stairs stabilize to sharp symplectic embedding obstructions", *J. Topol.*, 11(2):309–378, 2018.
7. Cristofaro-Gardiner, Holm, Mandini, Pires, "On infinite staircases in toric symplectic four-manifolds", arXiv:2004.13062.
8. D. Frenkel, D. Muller, "Symplectic embeddings of 4-dimensional ellipsoids into cubes", *J. Symp. Geom.* 13(4):765–847, 2015.
9. M. Hutchings, "Quantitative embedded contact homology", *J. Diff. Geom.* 88(2):231–266, 2011.
10. M. Hutchings, *Lecture notes on embedded contact homology*, in *Contact and symplectic topology*, 389–484, Bolyai Soc. Math. Stud. **26**, *János Bolyai Math. Soc.*, Budapest, 2014.
11. M. Hutchings, "Beyond ECH capacities", *Geom. Topo.* 20(2):1085–1126, 2016.
12. B-H Li and T-J Li, "Symplectic genus, minimal genus and diffeomorphisms", *Asian J. Math.* **6** (2002)123–144.
13. T-J Li, A-K Liu, "Uniqueness of symplectic canonical class, surface cone and symplectic cone of 4-manifolds with $b^+ = 1$", *J. Differential Geom*. **58** (2001) 331–370.
14. M. Landry, M. McMillan, and E. Tsukerman, "On symplectic capacities of toric domains." *Involve* **8** (2015), no. 4, 665–676.
15. D. McDuff, "Symplectic embeddings of 4-dimensional ellipsoids." *J. Topol.* **2** (2009), no. 1, 1–22. *Corrigendum*: *J. Topol.* **8** (2015) no 4, 1119–1122.
16. D. McDuff, "The Hofer conjecture on embedding symplectic ellipsoids", *J. Differential Geom.* 88(3):519–532, 2011.
17. D. McDuff and F. Schlenk, "The embedding capacity of 4-dimensional symplectic ellipsoids", *Ann. Math* (2) 175 (2012), no. 3, 1191–1282.
18. D. McDuff and D. Salamon, *Introduction to Symplectic Topology*, 3rd edition, Oxford University Press 2017.
19. M. Usher, "Infinite staircases in the symplectic embedding problem for four-dimensional ellipsoids into polydisks", *Algebr. Geom. Topol.*, Volume 19, Number 4 (2019), 1935-/2022.
20. B. Wormleighton, "Algebraic Capacities", arXiv:2006.13296

Action-Angle and Complex Coordinates on Toric Manifolds

Haniya Azam, Catherine Cannizzo, and Heather Lee

1 Introduction

We start off with an overview of the two sets of coordinates, action-angle and complex, on compact toric manifolds.

Compact symplectic toric manifolds (M^{2n}, ω) modulo $\mathbb{T}^n \cong (S^1)^n$ equivariant symplectomorphisms are in one-to-one correspondence with a class of compact convex polytopes known as Delzant polytopes, modulo translation [11]. One direction of this correspondence is Atiyah [4] and Guillemin-Sternberg's [18] convexity theorem that given a symplectic toric manifold (M^{2n}, ω, μ), where $\mu : M^{2n} \to \mathbb{R}^n$ is the moment map of the Hamiltonian toric action, the image $\mu(M^{2n})$ is a convex polytope. Conversely, given a Delzant polytope Δ^n, Delzant's construction produces a compact symplectic toric manifold M^{2n}_Δ such that the image of its moment map $\mu : M^{2n}_\Delta \to \mathbb{R}^n$ is Δ^n. The toric manifold M^{2n}_Δ is obtained from a symplectic reduction of $\mathbb{C}^d$ (where d is the number of facets of Δ^n) with respect to the action of a $(d-n)$-dimensional subtorus $N \subset \mathbb{T}^d$. Denote by μ_N the moment map of this action by N. The reduced space, $M^{2n}_\Delta = \mathbb{C}^d /\!\!/ N = \mu_N^{-1}(a)/N$ (for regular values a), carries a Hamiltonian $\mathbb{T}^n \cong \mathbb{T}^d/N$-action and a canonical $\mathbb{T}^n$-invariant symplectic

H. Azam
Lahore University of Management Sciences, Lahore, Pakistan
e-mail: haniya.azam@lums.edu.pk

C. Cannizzo
Simons Center for Geometry and Physics, State University of New York, Stony Brook, NY, USA
e-mail: ccannizzo@scgp.stonybrook.edu
https://sites.google.com/view/ccannizzo/

H. Lee (✉)
Department of Mathematics, University of Washington, Seattle, WA, USA
e-mail: heathml@uw.edu

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_3

form ω. Each $\mathbb{T}^n$-orbit is a fiber of the moment map μ, i.e. the preimage of a point in Δ^n under μ. This $\mathbb{T}^n$-action is free on the preimage $\mu^{-1}(\mathring{\Delta})$ of the interior $\mathring{\Delta}$ of Δ^n, and it is degenerate on the complement, which is the preimage $\mu^{-1}(\partial\Delta^n)$ of the boundary. Therefore, $\mu^{-1}(\mathring{\Delta})$ is diffeomorphic to $\mathring{\Delta} \times \mathbb{T}^n$.

Since $\mathbb{C}^d$ carries a $\mathbb{T}^d$-invariant Kähler structure, it induces an ω-compatible $\mathbb{T}^n$-invariant Kähler structure (g, J) on the reduced space M^{2n}_Δ. The complex manifold (M^{2n}_Δ, J) is then a toric variety with the complexified torus $\mathbb{T}^n_\mathbb{C}$-action. This complex torus $\mathbb{T}^n_\mathbb{C} \cong (\mathbb{C}^*)^n$ acts freely and transitively on $\mu^{-1}(\mathring{\Delta})$, hence $\mu^{-1}(\mathring{\Delta})$ is diffeomorphic to $\mathbb{T}^n_\mathbb{C}$. We now have two natural sets of coordinates on $\mu^{-1}(\mathring{\Delta}) \subset M^{2n}_\Delta$ coming from the action-angle coordinates on $\mathring{\Delta} \times \mathbb{T}^n$ and the complex toric coordinates on $\mathbb{T}^n_\mathbb{C}$ via the identification

$$\mathring{\Delta} \times \mathbb{T}^n \cong \mu^{-1}(\mathring{\Delta}) \cong \mathbb{T}^n_\mathbb{C}. \tag{1}$$

These two sets of coordinates are related by Legendre transform; however, explicit formulas for switching between them might be complicated or sometimes impossible to obtain, as we will illustrate with examples (most elaborately in Sect. 3.5). Understanding the relation between these two sets of coordinates was the initial motivation for this paper (see Sect. 5.3) and is one of the main topics explored.

For the Kähler form ω on M, Guillemin [16] extended this story further by showing that the ω possesses a $\mathbb{T}^n$-invariant Kähler potential $F : \mathbb{T}^n_\mathbb{C} \to \mathbb{R}$ such that $\omega = 2i\partial\overline{\partial}F$ and it only depends on the real part of the Lie algebra $\mathrm{Lie}(\mathbb{T}^n_\mathbb{C}) \cong \mathbb{C}^n$ coordinates, so it is a function $F : \mathbb{R}^n \to \mathbb{R}$. Guillemin also provided a dual potential $G : \mathring{\Delta} \to \mathbb{R}$, which is a function of the moment map coordinates and it is a Legendre transform of F. The formulas for the potential functions are given by (up to adding constants) the combinatorial structure of Δ^n.

Even though the theories above are all done for compact manifolds, the idea behind describing symplectic toric manifolds via symplectic reduction carries over to noncompact manifolds as well. However, for noncompact symplectic toric manifolds, the moment map image might not be convex and the fibers of the moment map might not be connected; see [24] for a classification result for noncompact symplectic manifolds. In this paper, for noncompact examples, we will only focus on those where the theories for compact toric manifolds carry over, i.e. those with convex moment map images and connected moment map fibers.

In fact, when it comes to noncompact examples, we will almost exclusively focus on the total space of the canonical line bundle $\pi : K_M \to M$ of a compact toric manifold. Because the tangent bundle $T_{K_M} \cong \pi^* T_M \oplus K_M$ and $c_1(K_M) = -c_1(T_M)$, we get that $c_1(T_{K_M}) = 0$, so K_M is a Calabi-Yau (CY) toric manifold. In connection to mirror symmetry, the singular symplectic fibration $W : K_M \to \mathbb{C}$, where W is the product of the homogeneous coordinates, gives an example of a Landau-Ginzburg model (K_M, W). Furthermore, when M is in addition Fano, the critical locus of the fibration $W : K_M \to \mathbb{C}$ provides another source of Calabi-Yau manifolds that are of interest to many. This is behind our motivation for understanding the symplectic structure on K_M, and more generally,

(K_M, W) above serves as a prototypical class of examples of a toric Landau-Ginzburg model (Y, W), where Y is a CY toric manifold.

Here is an outline of the paper. In Sect. 2, we summarize the construction of compact toric manifolds M^{2n} as symplectic quotients. In Sect. 3, we discuss the $\mathbb{T}^n$-action on M^{2n} and its moment map. In Sect. 4, we present Guillemin's [16] Kähler potentials. We use many examples to illustrate the theories, with the complex projective space $\mathbb{P}^n$ being the main running example throughout Sects. 2–4. Another running example throughout this paper is the aforementioned noncompact example of K_M. (In Sects. 2 and 3, the K_M example is contained in Sects. 2.3, 3.2, 3.4, 3.5.) Finally in Sect. 5, we explain some applications to mirror symmetry. In particular, we mention how to compute the monodromy of a fiber around the singularity of a symplectic fibration given by a superpotential on K_M, where the horizontal distribution is the ω-orthogonal complement to the tangent space of each fiber; for this we use the symplectic form ω obtained via symplectic reduction. Section 6 lists notation used in the text.

2 Toric Manifolds as Symplectic Quotients

In this section, we describe the compact symplectic toric manifold M of complex dimension n, but we will defer the discussion about the toric $\mathbb{T}^n$-action on M to Sect. 3 and the Kähler structure to Sect. 4. The description of M doesn't start until Sect. 2.2. Below we give a brief summary of the content of each subsection.

Section 2.1 is entirely complex geometric, in which we set up some notation and describe a proper and holomorphic action of a complex $(d-n)$-dimensional torus $N_{\mathbb{C}} = (\mathbb{C}^*)^{d-n}$ on $\mathbb{C}^d$. We then introduce an open set U where $N_{\mathbb{C}}$ acts freely, and construct the quotient $U/N_{\mathbb{C}}$ with the induced complex manifold structure. We take this complex geometric starting point because many examples we would like to consider, such as complex projective space, are most naturally described in this way.

Section 2.2 gives the description of M as a symplectic quotient of $\mathbb{C}^d$ (with the standard symplectic structure) with respect to a Hamiltonian N-action, where $N = U(1)^{d-n}$ is a real $(d-n)$-dimensional torus. More explicitly, M can be described as a quotient $\mu_N^{-1}(a)/N$ where a is a regular value of the moment map μ_N of the N-action. This symplectic quotient $\mu_N^{-1}(a)/N$ can be identified with a complex geometric quotient $U_a/N_{\mathbb{C}}$ for a certain Zariski open set $U_a \subset \mathbb{C}^d$ on which $N_{\mathbb{C}}$ acts freely. Now $U = U_a$ depends on which chamber a is in, where going from one chamber to another amounts to crossing a wall of values where μ_N is not regular.

We use the complex project space $\mathbb{P}^n$ as a running example to illustrate the set-up and theories in Sects. 2.1 and 2.2 (see Examples 2, 3, and 5). In Sect. 2.3, we provide an example of a noncompact toric manifold K_M, which is the canonical line bundle of a compact toric manifold M.

2.1 Complex Geometric Quotients

Consider the action of the complex torus $\mathbb{T}^d_{\mathbb{C}} = (\mathbb{C}^*)^d$ on $\mathbb{C}^d$ given by

$$\begin{aligned} \mathbb{T}^d_{\mathbb{C}} \times \mathbb{C}^d &\to \mathbb{C}^d \\ (\widetilde{t}, z) &\mapsto \widetilde{t} \cdot z = (\widetilde{t}_1 z_1, \ldots, \widetilde{t}_d z_d), \end{aligned} \tag{2}$$

where $\widetilde{t} = (\widetilde{t}_1, \ldots, \widetilde{t}_d) \in \mathbb{T}^d_{\mathbb{C}}$ (to distinguish it from coordinates t_i on the complex algebraic torus $\mathbb{T}^{n=\dim_{\mathbb{C}} M}_{\mathbb{C}}$ in Chap. 3) and $z = (z_1, \ldots, z_d) \in \mathbb{C}^d$.

Let $N_{\mathbb{C}} := (\mathbb{C}^*)^{d-n}$ and fix any injective group homomorphism

$$\rho_{\mathbb{C}} : N_{\mathbb{C}} = (\mathbb{C}^*)^{d-n} \longrightarrow \mathbb{T}^d_{\mathbb{C}} = (\mathbb{C}^*)^d. \tag{3}$$

Then the image $\rho_{\mathbb{C}}(N_{\mathbb{C}})$ is a $(d-n)$-dimensional subtorus in $\mathbb{T}^d_{\mathbb{C}}$. We then have a short exact sequence of abelian groups

$$1 \to N_{\mathbb{C}} \xrightarrow{\rho_{\mathbb{C}}} \mathbb{T}^d_{\mathbb{C}} \xrightarrow{\beta_{\mathbb{C}}} \mathbb{T}^n_{\mathbb{C}} \to 1, \tag{4}$$

where $\mathbb{T}^n_{\mathbb{C}} := \operatorname{coker}\rho_{\mathbb{C}} = \mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}} \cong (\mathbb{C}^*)^n$ and $\beta_{\mathbb{C}}$ is any group homomorphism that makes the above sequence exact. Let's fix notations for the maps $\rho_{\mathbb{C}}$ and $\beta_{\mathbb{C}}$ below.

1. The map $\rho_{\mathbb{C}}$ is of the form

$$\rho_{\mathbb{C}}(\lambda_1, \ldots, \lambda_{d-n}) = \left(\prod_{\ell=1}^{d-n} \lambda_\ell^{Q_1^\ell}, \ldots, \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_d^\ell} \right) \tag{5}$$

   where $Q_k^\ell \in \mathbb{Z}$, $1 \le \ell \le d-n$, $1 \le k \le d$.
2. The map $\beta_{\mathbb{C}}$ is of the form

$$\beta_{\mathbb{C}}(\widetilde{t}_1, \ldots, \widetilde{t}_d) = \left(\prod_{k=1}^{d} \widetilde{t}_k^{v_1^k}, \ldots, \prod_{k=1}^{d} \widetilde{t}_k^{v_n^k} \right) \tag{6}$$

   where $v_m^k \in \mathbb{Z}$, $1 \le k \le d$, $1 \le m \le n$.
3. The exactness of (4) implies the exponent on each λ_ℓ in the mth coordinate of $\beta_{\mathbb{C}} \circ \rho_{\mathbb{C}}$ is 0, namely

$$\sum_{k=1}^{d} v_m^k Q_k^\ell = 0, \quad 1 \le m \le n, \quad 1 \le \ell \le d-n. \tag{7}$$

The choice of $\beta_{\mathbb{C}}$ delimited by Eq. (7) is not unique, and one can see this more easily by considering the linearized maps in matrix form. Let $\mathfrak{n}_{\mathbb{C}} \cong \mathbb{C}^{d-n}$, $\mathfrak{t}^d_{\mathbb{C}} \cong \mathbb{C}^d$, and $\mathfrak{t}^n_{\mathbb{C}} \cong \mathbb{C}^n$ be the Lie algebras of $N_{\mathbb{C}}$, $\mathbb{T}^d_{\mathbb{C}}$, and $\mathbb{T}^n_{\mathbb{C}}$, respectively. Taking the differentials of the two group homomorphisms in (4) at the identity, we obtain the following short exact sequence of complex vector spaces:

$$0 \to \mathfrak{n}_{\mathbb{C}} \xrightarrow{(d\rho_{\mathbb{C}})_1} \mathfrak{t}^d_{\mathbb{C}} \xrightarrow{(d\beta_{\mathbb{C}})_1} \mathfrak{t}^n_{\mathbb{C}} \to 0. \tag{8}$$

The linear maps $(d\rho_{\mathbb{C}})_1$ and $(d\beta_{\mathbb{C}})_1$ are given by matrices Q and B respectively, where

$$Q_{d\times(d-n)} = \begin{bmatrix} Q_1^1 & \cdots & Q_1^{d-n} \\ \vdots & & \vdots \\ Q_d^1 & \cdots & Q_d^{d-n} \end{bmatrix}, \quad B_{n\times d} = \begin{bmatrix} v_1^1 & \cdots & v_1^d \\ \vdots & & \vdots \\ v_n^1 & \cdots & v_n^d \end{bmatrix}. \tag{9}$$

Equation (7) is then equivalent to $BQ = 0$. Given Q, the relation $BQ = 0$ does not uniquely determine B since it only requires that the row vectors of B generate $Q^{\perp}$. Given Q, to write $B : \mathbb{Z}^d \to \mathbb{Z}^d / \langle \mathrm{col}(Q) \rangle \cong \mathbb{Z}^n$ in matrix form we need to make an identification of $\mathbb{Z}^d / \langle \mathrm{col}(Q) \rangle$ with $\mathbb{Z}^n$, where $\langle \mathrm{col}(Q) \rangle$ denotes the column space of Q. Also because B is surjective, $\langle \mathrm{col}(B) \rangle = \mathbb{Z}^n$ and the column vectors of B are primitive. Other choices for the matrix B will hence differ by left multiplication of elements in $\mathrm{GL}(n, \mathbb{Z})$ with determinant ± 1. (On the other hand, if we fix B, then the choices for Q would differ by right multiplication of elements in $\mathrm{GL}(d-n, \mathbb{Z})$ so as to keep $\ker B$ fixed).

Example 1 ($\mathbb{P}(1, 2, 3)$) As an example, take $Q = [1, 2, 3]^T$, $d = 3$, and $n = 2$. First we make an identification of the quotient with $\mathbb{Z}^{n=2}$. Extend the column vectors of Q to a $\mathbb{Z}$-basis for $\mathbb{Z}^3$, given by

$$\{f_1, f_2, f_3\} := \{(1, 2, 3)^T, (0, 1, 0)^T, (0, 0, 1)^T\},$$

so that

$$\frac{\mathbb{Z}^3}{\langle \mathrm{col}(Q) \rangle} = \frac{\mathbb{Z}f_1 \oplus \mathbb{Z}f_2 \oplus \mathbb{Z}f_3}{\mathbb{Z}f_1} = \mathbb{Z}f_2 \oplus \mathbb{Z}f_3.$$

Since $\ker B = \langle \mathrm{col}(Q) \rangle$, let B be the linear map $f_1 \mapsto (0, 0)$, $f_2 \mapsto (1, 0)$, and $f_3 \mapsto (0, 1)$. Then using that

$$(1, 0, 0) = f_1 - 2f_2 - 3f_3 \implies B : (1, 0, 0) \mapsto (0, 0) - 2(1, 0) - 3(0, 1) = (-2, -3),$$

with respect to the standard bases on $\mathbb{Z}^{d=3}$ on $\mathbb{Z}^{n=2}$ we see that $B : \mathbb{Z}^d \to \mathbb{Z}^n$ is

$$B = \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}.$$

(This example leads to the weighted projective space $\mathbb{P}(1, 2, 3)$, cf. the complex projective space $\mathbb{P}^2 = \mathbb{P}(1, 1, 1)$ in Example 2.) □

Let $N_{\mathbb{C}} = (\mathbb{C}^*)^{d-n}$ act on $\mathbb{C}^d$ by its image under $\rho_{\mathbb{C}} : N_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}}$, where the action of the subtorus $\rho_{\mathbb{C}}(N_{\mathbb{C}}) \subset \mathbb{T}^d_{\mathbb{C}}$ is inherited from the action of $\mathbb{T}^d_{\mathbb{C}}$. In other words, this action is

$$\begin{aligned} N_{\mathbb{C}} \times \mathbb{C}^d &\to \mathbb{C}^d \\ (\lambda, z) &\mapsto \lambda \cdot z := \rho_{\mathbb{C}}(\lambda) \cdot z \end{aligned} \tag{10}$$

where $\lambda \in N_{\mathbb{C}}$, $z \in \mathbb{C}^d$, and $\rho_{\mathbb{C}}(\lambda) \cdot z$ on the right hand side is given by componentwise multiplication as in Eq. (2).

This $N_{\mathbb{C}}$-action on $\mathbb{C}^d$ is holomorphic and proper, but it is not free everywhere, so simply taking the quotient $\mathbb{C}^d/N_{\mathbb{C}}$ will not give us a manifold structure. If U is a subset of $\mathbb{C}^d$ on which $N_{\mathbb{C}}$ acts freely, then the quotient $U/N_{\mathbb{C}}$ has a unique complex structure such that the quotient map $\pi : U \to U/N_{\mathbb{C}}$ is holomorphic. Many complex manifolds can be described in this way as a quotient, such as Example 2 for the complex projective space below.

Example 2 ($\mathbb{P}^n$) Consider $d = n + 1$ and $N_{\mathbb{C}} = (\mathbb{C}^*)^{d-n=1} = \mathbb{C}^*$. The image of the embedding

$$\rho_{\mathbb{C}} : N_{\mathbb{C}} = \mathbb{C}^* \to \mathbb{T}^d_{\mathbb{C}} = (\mathbb{C}^*)^{n+1}, \quad \rho_{\mathbb{C}}(\lambda_1) = (\lambda_1, \ldots, \lambda_1) \tag{11}$$

is a subtorus $\rho_{\mathbb{C}}(N_{\mathbb{C}}) \cong \mathbb{C}^*$ in $\mathbb{T}^d_{\mathbb{C}}$, which gives the following $N_{\mathbb{C}} = \mathbb{C}^*$-action on $\mathbb{C}^{d=n+1}$

$$\lambda_1 \cdot (z_1, \ldots, z_{n+1}) = (\lambda_1 z_1, \ldots, \lambda_1 z_{n+1}).$$

Consider $U = \mathbb{C}^{n+1} - \{0\}$; then the $N_{\mathbb{C}}$ action on U is free. We find that the quotient $U/N_{\mathbb{C}}$

$$\mathbb{P}^n = (\mathbb{C}^{n+1} - \{0\})/\mathbb{C}^*, \tag{12}$$

is the complex projective space consisting of complex lines through the origin in $\mathbb{C}^{n+1}$.

Note that so long as $\rho_{\mathbb{C}}$ is the one given above in Eq. (11), the resulting quotient is $\mathbb{P}^n$ as in Eq. (12), no matter what $\beta_{\mathbb{C}}$ is. The choice of $\beta_{\mathbb{C}}$ is not unique for a given $\rho_{\mathbb{C}}$. For the $\rho_{\mathbb{C}}$ given by Eq. (11), the linear map $(d\rho_{\mathbb{C}})_1$ has the matrix form

$$Q = [\, 1 \cdots 1 \,]^T. \tag{13}$$

One choice of B such that $BQ = 0$ is

$$B = \left[\begin{array}{c|c} & -1 \\ \mathbb{I}_n & \vdots \\ & -1 \end{array}\right], \tag{14}$$

where $\mathbb{I}_n$ is the $n \times n$ identity matrix. This matrix B corresponds to

$$\beta_{\mathbb{C}}(\widetilde{t}_1, \ldots, \widetilde{t}_{n+1}) = (\widetilde{t}_1 \widetilde{t}_{n+1}^{-1}, \ldots \widetilde{t}_n \widetilde{t}_{n+1}^{-1}). \tag{15}$$

□

2.2 Symplectic Quotients

Let $N = U(1)^{d-n}$ be the maximal compact subgroup of $N_{\mathbb{C}}$. The $N_{\mathbb{C}}$-action on $\mathbb{C}^d$ defined in Eq. (10) restricts to a Hamiltonian N-action on the symplectic manifold

$$\left(\mathbb{C}^d, \omega_0 = \frac{i}{2} \sum_{k=1}^{d} dz_k \wedge d\bar{z}_k\right). \tag{16}$$

In this subsection, we first consider the symplectic quotient $\mathbb{C}^d /\!\!/ N$ of $(\mathbb{C}^d, \omega_0)$ with respect to the action by N. The symplectic quotient provided in Theorem 1 is a symplectic toric manifold. After that, we will state in Theorem 2 that this symplectic quotient can be identified with a complex geometric quotient.

To describe the symplectic quotient $\mathbb{C}^d /\!\!/ N$, we first need to obtain the moment map of the N-action. Consider the maximal compact subgroup $\mathbb{T}^d = U(1)^d$ of $\mathbb{T}^d_{\mathbb{C}}$. Then the $\mathbb{T}^d_{\mathbb{C}}$-action on $\mathbb{C}^d$ defined in Eq. (2) restricts to a Hamiltonian $\mathbb{T}^d$-action on the symplectic manifold $(\mathbb{C}^d, \omega_0)$ given in Eq. (16) with a moment map (which is unique up to addition of a constant vector in $\mathbb{R}^d$ that plays a role in Eq. (45))

$$\mu_{\mathbb{T}^d} : \mathbb{C}^d \to \mathbb{R}^d, \quad \mu_{\mathbb{T}^d}(z_1, \ldots, z_d) = \frac{1}{2}(|z_1|^2, \ldots, |z_d|^2). \tag{17}$$

The image of N under $\rho_{\mathbb{C}}$ is a subtorus of $\mathbb{T}^d$, and one can equivalently think of the N-action on $\mathbb{C}^d$ as an action of its image $\rho_{\mathbb{C}}(N)$, which inherits the action of $\mathbb{T}^d$. Hence, the moment map (which is unique up to addition of a constant vector in $\mathbb{R}^{d-n}$) of the Hamiltonian N-action is

$$\begin{aligned} \mu_N &:= (d\rho_{\mathbb{C}})_1^* \circ \mu_{\mathbb{T}^d} : \mathbb{C}^d \to \mathbb{R}^{d-n}, \\ \mu_N(z_1, \ldots, z_d) &= \tfrac{1}{2}\left(\sum_{k=1}^{d} Q_k^1 |z_k|^2, \ldots, \sum_{k=1}^{d} Q_k^{d-n} |z_k|^2\right). \end{aligned} \tag{18}$$

Example 3 ($\mathbb{P}^n$) In our running example $\mathbb{P}^n$, the holomorphic $N_{\mathbb{C}} = \mathbb{C}^*$-action on $\mathbb{C}^{n+1}$ restricts to a Hamiltonian $N = U(1)$-action on the symplectic manifold $(\mathbb{C}^{n+1}, \omega_0)$. Its moment map can be calculated from the linear map $(d\rho_{\mathbb{C}})_1$, which in matrix form is given by Eq. (13), and $\mu_{\mathbb{T}^d}$ given in Eq. (17). Up to a constant, it is

$$\begin{aligned} &\mu_N : \mathbb{C}^{n+1} \to \mathbb{R}, \\ &\mu_N(z_1, \ldots, z_{n+1}) = (d\rho_{\mathbb{C}})_1^* \circ \mu_{\mathbb{T}^d}(z_1, \ldots, z_{n+1}) = \tfrac{1}{2} \sum_{k=1}^{n+1} |z_k|^2. \end{aligned} \tag{19}$$

□

For $a \in \mathbb{R}^{d-n}$, the following two statements are equivalent:

- a is a regular value of μ_N, meaning that $d\mu_N$ is surjective at each point in $\mu_N^{-1}(a)$, and hence $\mu_N^{-1}(a)$ is a smooth manifold of the expected real dimension $2d - (d - n) = d + n$.
- The N-action on $\mu_N^{-1}(a)$ is locally free.

This is because the surjectivity of $d\mu_N$ is equivalent to the linear independence of the Hamiltonian vector fields generated by the $d - n$ components of μ_N.

For the rest of this paper, we will only be considering the cases where a is a regular value of μ_N and that N acts freely on $\mu_N^{-1}(a)$, not just locally free. When the N doesn't act freely, the quotient might be an orbifold. We are now ready to describe the symplectic quotient $\mathbb{C}^d /\!\!/ N$ of $\mathbb{C}^d$ by the action of N.

Theorem 1 (Marsden-Weinstein [30] and Meyer [32] Symplectic Reduction) *Consider the Hamiltonian action on $(\mathbb{C}^d, \omega_0)$ by the compact group N with a moment map $\mu_N : \mathbb{C}^d \to \mathbb{R}^{d-n}$, as described above. If a is a regular value of μ_N and that N acts freely on $\mu_N^{-1}(a)$, then the symplectic quotient*

$$M := \mathbb{C}^d /\!\!/ N = \mu_N^{-1}(a)/N$$

admits a unique structure of a smooth manifold of real dimension $2n$ such that the projection $\pi_a : \mu_N^{-1}(a) \to \mu_N^{-1}(a)/N$ is a submersion, and it carries a canonical symplectic form ω_a such that

$$\pi_a^* \omega_a = \iota_a^* \omega_0,$$

where $\iota_a : \mu_N^{-1}(a) \hookrightarrow \mathbb{C}^d$ is the inclusion.

In Sect. 3.1, we will discuss that M carries an effective Hamiltonian $\mathbb{T}^n$-action, which makes it a toric symplectic manifold. Since $(\mathbb{C}^d, \omega_0)$ is Kähler, compatible with the standard complex structure, the symplectic reduction $\mu_N^{-1}(a)/N$ also has a natural Kähler structure. We will discuss the Kähler structure in more detail in Sect. 4.

This symplectic quotient $M = \mu_N^{-1}(a)/N$ can also be described as a complex geometric quotient by the following theorem.

Theorem 2 (Kempf-Ness [25], Kirwan [26], Audin [5], Guillemin [17]) *There is a Zariski open set $U_a \subset \mathbb{C}^d$ such that $\mu_N^{-1}(a) \subset U_a$, $N_{\mathbb{C}}$ acts freely on U_a, and*

$$M = \mu_N^{-1}(a)/N = U_a/N_{\mathbb{C}}.$$

The quotient M is a smooth manifold carrying the canonical symplectic structure ω_a as in Theorem 1 and it carries the unique complex structure such that $\widetilde{\pi}_a : U_a \to U_a/N_{\mathbb{C}}$ is holomorphic.

Here we decided to simply use the notation M instead of M_a since whenever we use M later on, it's always for a fixed a, so there's no confusion. The Kempf-Ness theorem [25] is a deep and more general theorem describing the equivalence between the geometric invariant theory (GIT) quotient of a smooth complex projective variety X by the linear action of a complex reductive group G_c and the symplectic quotient of X by the maximal compact subgroup G of G_c. We will not describe the geometric invariant theory in this article. Audin [5, Proposition 3.1.1] gave a proof of the above result in the specialized setting for toric manifolds that is similar to Kirwan's [26, Theorem 7.4] more general proof for quotients of Kähler manifolds. Guillemin [17, Section A1.2] gave a more elementary proof in the case of toric manifolds.

We do not explain the proof for Theorem 2 in this article; however, in Section "Description of U_a and Complex Toric Coordinates" we will provide an explicit description of U_a that can be found in [5, 17], corresponding to $\mu_N^{-1}(a)/N$, in this setting of toric manifolds. This theorem is also illustrated in our Examples 4, 5, and 7. Note that the proof of this theorem depends on the convexity property of the moment map of the $\mathbb{T}^n$-action on the symplectic toric manifold $\mu_N^{-1}(a)/N$, which is guaranteed for compact toric manifolds by Theorem 4 but not guaranteed in general for noncompact toric manifolds. However, for all of our noncompact examples, the image of the moment map is convex.

Example 4 ($\mathrm{Tot}(\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1)$*)* Consider $N_{\mathbb{C}} = \mathbb{C}^*$ acting on $\mathbb{C}^4$ by

$$\lambda_1 \cdot (z_1, z_2, z_3, z_4) = (\lambda_1 z_1, \lambda_1 z_2, \lambda_1^{-1} z_3, \lambda_1^{-1} z_4).$$

It restricts to a Hamiltonian $N = U(1)$-action on $(\mathbb{C}^4, \omega_0)$ with the moment map $\mu_N : \mathbb{C}^4 \to \mathbb{R}$ given by $\mu_N(z_1, z_2, z_3, z_4) = \frac{1}{2}(|z_1|^2 + |z_2|^2 - |z_3|^2 - |z_4|^2)$. Then $a \in \mathbb{R}$ is a regular value if and only if $a \neq 0$, and the level set is

$$\mu_N^{-1}(a) = \{|z_1|^2 + |z_2|^2 - |z_3|^2 - |z_4|^2 = 2a\} \subset \mathbb{C}^4.$$

We can see that for $z \in \mu_N^{-1}(a)$, if $a > 0$, then z_1 and z_2 cannot both be 0, and if $a < 0$, z_3 and z_4 cannot both be 0. If we take U_a to be

$$U_a = \begin{cases} (\mathbb{C}^2 - \{0\}) \times \mathbb{C}^2, & a > 0; \\ \mathbb{C}^2 \times (\mathbb{C}^2 - \{0\}), & a < 0, \end{cases}$$

then $\mu_N^{-1}(a) \subset U_a$ and

$$\mu_N^{-1}(a)/U(1) = U_a/\mathbb{C}^*,$$

giving the total space of $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1$. For $a > 0$, the zero section is given by

$$\{(z_1, z_2, 0, 0) \mid |z_1|^2 + |z_2|^2 = 2a\}/U(1) = S^3(\sqrt{2a})/U(1) = S^2(\sqrt{a/2}) \cong \mathbb{P}^1.$$

Similarly, for $a < 0$, the zero section is given by

$$\{(0, 0, z_3, z_4) \mid -|z_3|^2 - |z_4|^2 = 2a\}/U(1) = S^3(\sqrt{-2a})/U(1) = S^2(\sqrt{-a/2}) \cong \mathbb{P}^1.$$

□

Given the identification of orbit spaces $U_a/N_{\mathbb{C}} = \mu_N^{-1}(a)/N$, for any $z \in U_a$, it is in one of the $N_{\mathbb{C}}$-orbits, which corresponds to a unique N-orbit in $\mu_N^{-1}(a) \subset U_a$, so there exists $\widetilde{\lambda}_a(z) \in N_{\mathbb{C}}$ such that $\rho_{\mathbb{C}}(\widetilde{\lambda}_a(z)) \cdot z$ is in that N-orbit. Furthermore, because the $N_{\mathbb{C}}$-action is free, there is a unique $\lambda_a(z) \in N_{\mathbb{C}}/N = (\mathbb{R}_{>0})^{d-n}$ such that $\rho_{\mathbb{C}}(\lambda_a(z)) \cdot z \in \mu_N^{-1}(a)$. Therefore, we have the following deformation retraction.

Definition 1 (Deformation Retraction R_a) For any $z \in U_a$, take the unique $\lambda_a(z) \in N_{\mathbb{C}}/N = (\mathbb{R}_{>0})^{d-n}$ such that $\rho_{\mathbb{C}}(\lambda_a(z)) \cdot z \in \mu_N^{-1}(a)$. This defines a deformation retraction

$$R_a : U_a \to \mu_N^{-1}(a) \subseteq \mathbb{C}^d, \quad z \mapsto \lambda_a(z) \cdot z = \rho_{\mathbb{C}}(\lambda_a(z)) \cdot z. \tag{20}$$

We will see in the example provided in Sect. 3.5 that computing this deformation retraction is the main challenge one encounters when trying to express the moment map of the $\mathbb{T}^n$-action on M in terms of the homogeneous coordinates $z \in \mathbb{C}^d$.

Example 5 ($\mathbb{P}^n$) One can see from Eq. (19) that $a \in \mathbb{R}$ is a regular value of μ_N if and only if $a \neq 0$, and the level sets are

$$\mu_N^{-1}(a) = \begin{cases} \{z \in \mathbb{C}^{n+1} \mid \sum_{k=1}^{n+1} |z_k|^2 = 2a\} = S^{2n-1}(\sqrt{2a}), & a > 0; \\ \emptyset, & a < 0. \end{cases}$$

So if $z \in \mu_N^{-1}(a) \subset \mathbb{C}^{n+1}$ and $a > 0$, then $z \neq 0$. We see that if we take $U_a = \mathbb{C}^{n+1} - \{0\}$, then

$$\mu_N^{-1}(a)/U(1) = U_a/\mathbb{C}^* = (\mathbb{C}^{n+1} - \{0\})/\mathbb{C}^* = \mathbb{P}^n,$$

where $\mu_N^{-1}(a)$ inherits a symplectic structure while $(\mathbb{C}^{n+1} - \{0\})/\mathbb{C}^*$ inherits a complex structure.

For any $z \in \mathbb{C}^{n+1} - \{0\}$, by Eq. (19), the unique $\lambda_a(z) \in \mathbb{C}^*/U(1) \simeq \mathbb{R}_{>0}$ such that $\lambda_a(z) \cdot z \in \mu_N^{-1}(a)$ is $\lambda_a(z_1, \ldots, z_{n+1}) = \sqrt{\frac{2a}{\sum_{j=1}^{n+1} |z_j|^2}}$. Hence, with $\rho_{\mathbb{C}}$ given in Eq. (11), the deformation retraction is given by

$$R_a : \mathbb{C}^{n+1} - \{0\} \to \mu_N^{-1}(a), \qquad (z_1, \ldots, z_{n+1}) \mapsto \sqrt{\frac{2a}{\sum_{j=1}^{n+1} |z_j|^2}}(z_1, \ldots, z_{n+1}). \tag{21}$$

□

2.3 Canonical Line Bundle K_M of a Toric Manifold M

This is a noncompact example, though as mentioned in the introduction section, the theory for compact toric manifolds carries over to this example. Let M be a compact toric manifold obtained as a symplectic quotient of $(\mathbb{C}^d, \omega_0)$ with respect to the action of N as in the previous Sect. 2.2. In this subsection, we give a description of the total space of the canonical line bundle K_M over M as a symplectic reduction of

$$\left(\mathbb{C}^{d+1},\ \omega_0^+ = \frac{i}{2}\Big(\sum_{k=1}^{d} dz_k \wedge d\bar{z}_k + dp \wedge d\bar{p}\Big)\right), \tag{22}$$

also with respect to an action of N. So K_M is a toric manifold of complex dimension $n+1$, which is one complex dimension higher than M. We use $(z_1, \ldots z_d, p)$ to denote the coordinates on $\mathbb{C}^{d+1} = \mathbb{C}^d \times \mathbb{C}$, where $z \in \mathbb{C}^d$ are the homogeneous coordinates on M as in the previous section, and p is the additional coordinate in the fiber direction. Notation-wise in this subsection, we add a superscript $+$ to maps from the previous section to denote the extensions of those maps.

First we describe the complex geometric quotient. Let $\mathbb{T}_{\mathbb{C}}^{d+1} = (\mathbb{C}^*)^{d+1}$ act on $\mathbb{C}^{d+1}$ by

$$\begin{aligned} \mathbb{T}_{\mathbb{C}}^{d+1} \times (\mathbb{C}^d \times \mathbb{C}) &\to \mathbb{C}^d \times \mathbb{C} \\ (\widetilde{t}^+, z, p) &\mapsto \widetilde{t}^+ \cdot (z, p) := (\widetilde{t}_1 z_1, \ldots, \widetilde{t}_d z_d, \widetilde{t}_{d+1} p). \end{aligned}$$

Again the $\mathbb{T}_{\mathbb{C}}^{d+1}$-action on $\mathbb{C}^{d+1}$ restricts to a Hamiltonian $U(1)^{d+1}$-action on the symplectic manifold $(\mathbb{C}^{d+1}, \omega_0^+)$ with a moment map (which is unique up to addition of a constant vector in $\mathbb{R}^{d+1}$)

$$\mu_{\mathbb{T}^{d+1}} : \mathbb{C}^{d+1} \to \mathbb{R}^{d+1}, \quad \mu_{\mathbb{T}^{d+1}}(z_1, \ldots, z_d, p) = \frac{1}{2}(|z_1|^2, \ldots, |z_d|^2, |p|^2).$$

The $\rho_{\mathbb{C}}$ chosen above for M in Eq. (3) determines that for K_M, and the extended map $\rho_{\mathbb{C}}^+ : N_{\mathbb{C}} = (\mathbb{C}^*)^{d-n} \longrightarrow \mathbb{T}_{\mathbb{C}}^{d+1}$ needs to be

$$\rho_{\mathbb{C}}^+(\lambda_1, \ldots, \lambda_{d-n}) = \left(\prod_{\ell=1}^{d-n} \lambda_\ell^{Q_1^\ell}, \ldots, \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_d^\ell}, \prod_{\ell=1}^{d-n} \lambda_\ell^{-\sum_{k=1}^d Q_k^\ell} \right). \tag{23}$$

In Sect. 3.4, we will provide the justification that this extension $\rho_{\mathbb{C}}^+$ determines a N-action on $\mathbb{C}^{d+1}$ such that the symplectic reduction $\mathbb{C}^{d+1} /\!\!/ N$ is K_M, but for now we take this fact for granted.

For this given $\rho_{\mathbb{C}}^+$, the map $\beta_{\mathbb{C}}^+ : \mathbb{T}_{\mathbb{C}}^{d+1} \to \mathbb{T}_{\mathbb{C}}^{n+1}$ is of the form

$$\beta_{\mathbb{C}}^+(\widetilde{t}_1, \ldots, \widetilde{t}_{d+1}) = \left(\prod_{k=1}^{d} \widetilde{t}_k^{v_1^k}, \ldots, \prod_{k=1}^{d} \widetilde{t}_k^{v_n^k}, \prod_{k=1}^{d+1} \widetilde{t}_k \right)$$

and the corresponding matrices Q^+ and B^+ are given by

$$Q^+ = \begin{bmatrix} Q_1^1 & \cdots & Q_1^{d-n} \\ \vdots & & \vdots \\ Q_d^1 & \cdots & Q_d^{d-n} \\ -\sum_{k=1}^d Q_k^1 & \cdots & -\sum_{k=1}^d Q_k^{d-n} \end{bmatrix}, \quad B^+ = \begin{bmatrix} v_1^1 & \cdots & v_1^d & 0 \\ \vdots & & \vdots & \vdots \\ v_n^1 & \cdots & v_n^d & 0 \\ 1 & \cdots & 1 & 1 \end{bmatrix}. \tag{24}$$

Again by exactness of the analogue of Eq. (4), $0 \to \mathfrak{n}_{\mathbb{C}} \xrightarrow{(d\rho_{\mathbb{C}}^+)_1} \mathfrak{t}_{\mathbb{C}}^{d+1} \xrightarrow{(d\beta_{\mathbb{C}}^+)_1} \mathfrak{t}_{\mathbb{C}}^{n+1} \to 0$, we know that $B^+Q^+ = 0$.

Example 6 ($K_{\mathbb{P}^2}$) The Q and B for $\mathbb{P}^2$ from Example 2 extend to Q^+ and B^+ for $K_{\mathbb{P}^2}$ as

$$Q^+ = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \end{bmatrix}, \quad B^+ = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

where the corresponding maps are

$$\rho_{\mathbb{C}}^+(\lambda_1) = (\lambda_1, \lambda_1, \lambda_1, \lambda_1^{-3}), \quad \beta_{\mathbb{C}}^+(\widetilde{t}_1, \widetilde{t}_2, \widetilde{t}_3, \widetilde{t}_4) = (\widetilde{t}_1\widetilde{t}_3^{-1}, \widetilde{t}_2\widetilde{t}_3^{-1}, \widetilde{t}_1\widetilde{t}_2\widetilde{t}_3\widetilde{t}_4).$$

□

Now we describe the symplectic quotient. Let $N_{\mathbb{C}} \cong (\mathbb{C}^*)^{d-n}$ act on $\mathbb{C}^{d+1}$ by $\lambda \cdot (z, p) = \rho_{\mathbb{C}}^{+}(\lambda) \cdot (z, p)$, where $\lambda \in N_{\mathbb{C}}$, $z \in \mathbb{C}^d$, and $p \in \mathbb{C}$. The restriction of this action to the maximal compact subgroup $N = U(1)^{d-n}$ is a Hamiltonian N-action on the standard $\mathbb{C}^{d+1}$ of Eq. (22) with a moment map (unique up to addition of a constant vector in $\mathbb{R}^{d-n}$) given by

$$\begin{aligned} &\mu_N^+ : \mathbb{C}^{d+1} \to \mathbb{R}^{d-n} \\ &\mu_N^+(z_1, \ldots, z_d, p) = \frac{1}{2}\left(\sum_{k=1}^{d} Q_k^1(|z_k|^2 - |p|^2), \ldots, \sum_{k=1}^{d} Q_k^{d-n}(|z_k|^2 - |p|^2)\right). \end{aligned} \tag{25}$$

Let $a = (a_1, \ldots, a_{d-n}) \in \mathbb{R}^{d-n}$ be a regular value of μ_N^+. Then $(\mu_N^+)^{-1}(a)$ is a real submanifold of $\mathbb{C}^{d+1} = \mathbb{R}^{2d+2}$ of dimension $d + n + 2$ preserved by the N-action. The total space K_M of the canonical line bundle over M is the symplectic quotient

$$K_M = (\mu_N^+)^{-1}(a)/N = (U_a \times \mathbb{C})/N_{\mathbb{C}}. \tag{26}$$

Note that $U_a^+ = U_a \times \mathbb{C}$ because M sits in K_M as the zero section as discussed in Remark 1 below. Like Eq. (39), there is again a deformation retraction to go from $U_a \times \mathbb{C}$ to $(\mu_N^+)^{-1}(a)$; for any $(z, p) \in U_a \times \mathbb{C}$, there is a unique $\lambda_a(z, p) \in (\mathbb{R}_{>0})^{d-n} \subset (\mathbb{C}^*)^{d-n} = N_{\mathbb{C}}$ such that

$$\lambda_a(z, p) \cdot (z, p) \in (\mu_N^+)^{-1}(a),$$

from which we obtain the deformation retraction

$$R_a^+ : U_a \times \mathbb{C} \to (\mu_N^+)^{-1}(a), \quad (z, p) \mapsto \lambda_a(z, p) \cdot (z, p).$$

Then K_M carries a canonical symplectic form ω_a^+ satisfying $(\pi_a^+)^* \omega_a^+ = (\iota_a^+)^* \omega_0$, where

$$\iota_a^+ : (\mu_N^+)^{-1}(a) \hookrightarrow \mathbb{C}^{d+1}$$

is the inclusion and

$$\begin{cases} \pi_a^+ : (\mu_N^+)^{-1}(a) \to K_M = (\mu_N^+)^{-1}(a)/N, \\ \widetilde{\pi}_a^+ : U_a \times \mathbb{C} \to K_M = (U_a \times \mathbb{C})/N_{\mathbb{C}}, \end{cases} \tag{27}$$

are natural projections to the quotient K_M. We now illustrate with some examples.

Example 7 ($K_{\mathbb{P}^n}$) For $M = \mathbb{P}^n$, from Example 6, we see that to get $K_{\mathbb{P}^n}$, the $N_{\mathbb{C}} = \mathbb{C}^*$ action is

$$\lambda_1 \cdot (z_1, \ldots, z_{n+1}, p) = (\lambda_1 z_1, \ldots, \lambda_1 z_{n+1}, \lambda_1^{-n-1} p) \tag{28}$$

which restricts to a Hamiltonian $U(1)$-action with moment map

$$\mu_N^+(z_1, \ldots, z_{n+1}, p) = \frac{1}{2}\left(\sum_{k=1}^{n+1} |z_k|^2 - (n+1)|p|^2\right), \tag{29}$$

that is unique up to a constant. Then $a \in \mathbb{R}$ is a regular value of μ_N^+ if and only if $a \neq 0$. From the description of the regular level sets

$$(\mu_N^+)^{-1}(a) = \left\{\sum_{k=1}^{n+1} |z_k|^2 - (n+1)|p|^2 = 2a\right\}, \tag{30}$$

we see that if $(z, p) \in (\mu_N^+)^{-1}(a) \subset \mathbb{C}^{n+2}$ and $a > 0$, then $z \neq 0$, and if $a < 0$, then $p \neq 0$. Taking

$$U_a^+ = \begin{cases} (\mathbb{C}^{n+1} - \{0\}) \times \mathbb{C}, & a > 0; \\ \mathbb{C}^{n+1} \times (\mathbb{C} - \{0\}), & a < 0, \end{cases} \tag{31}$$

we see that

$$(\mu_N^+)^{-1}(a)/U(1) = \begin{cases} \left((\mathbb{C}^{n+1} - \{0\}) \times \mathbb{C}\right)/\mathbb{C}^* = K_{\mathbb{P}^n}, & a > 0; \\ \left(\mathbb{C}^{n+1} \times (\mathbb{C} - \{0\})\right)/\mathbb{C}^* = \mathbb{C}^{n+1}/\mathbb{Z}_{n+1}, & a < 0. \end{cases} \tag{32}$$

The justification for $\left((\mathbb{C}^{n+1} - \{0\}) \times \mathbb{C}\right)/\mathbb{C}^* = K_{\mathbb{P}^n}$ is provided in Sect. 3.4. Indeed, for $K_{\mathbb{P}^n}$, $U_a^+ = U_a \times \mathbb{C}$.

The deformation retraction $R_a^+ : U_a^+ \to (\mu_N^+)^{-1}(a)$ is given by

$$R_a^+(z_1, \ldots, z_{n+1}, p) = \lambda_a(z, p) \cdot (z, p) = (\lambda_a(z, p) z, \lambda_a(z, p)^{-n-1} p) \tag{33}$$

Computing $\lambda_a(z, p)$ will prove to be more challenging, which we explain in Sect. 3.5. □

Remark 1 (Connection Between M and K_M) In short, the connection is that the $p = 0$ level set is M. In more detail, let $a \in \mathbb{R}^{d-n}$ be a regular value, and define

$$h : (\mu_N^+)^{-1}(a) \to \mathbb{C}, \quad h(z_1, \ldots, z_d, p) = p.$$

Then from Eq. (25) for μ_N^+ and Eq. (18) for μ_N, we get

$$h^{-1}(p) = \left\{ (z_1, \ldots, z_d, p) \in \mathbb{C}^d \times \{p\} \;\middle|\; \sum_{k=1}^{d} Q_k^\ell |z_k|^2 = 2a_\ell + \sum_{k=1}^{d} Q_k^\ell |p|^2 \right\}$$
$$\cong \mu_N^{-1} \left(a_1 + \frac{1}{2} \sum_{k=1}^{d} Q_k^1 |p|^2, \; \ldots, \; a_{d-n} + \frac{1}{2} \sum_{k=1}^{d} Q_k^{d-n} |p|^2 \right).$$

When $p = 0$, we see that $h^{-1}(p) = \mu_N^{-1}(a)$. Between M and K_M there exist inclusion maps

$$\mu_N^{-1}(a) \times \{0\} \subset (\mu_N^+)^{-1}(a), \quad U_a \times \{0\} \subset U_a \times \mathbb{C}.$$

which descend to the following inclusion as the zero section:

$$i_0 : M = \mu_N^{-1}(a)/N = U_a/N_{\mathbb{C}} \hookrightarrow K_M = (\mu_N^+)^{-1}(a)/N = (U_a \times \mathbb{C})/N_{\mathbb{C}},$$

and $i_0^* \omega_a^+ = \omega_a$. □

Example 8 ($K_{\mathbb{P}^n}$) For the n-dimensional projective space,

$$h^{-1}(p) = \left\{ (z, p) \in \mathbb{C}^{n+1} \times \{p\} \mid |z_1|^2 + \cdots + |z_{n+1}|^2 = 2a + (n+1)|p|^2 \right\}$$
$$\cong \mu_N^{-1} \left(a + \tfrac{(n+1)|p|^2}{2} \right) = S^{2n-1} \left(\sqrt{2a + (n+1)|p|^2} \right)$$

where μ_N is the moment map given by Eq. (19). In particular, we can see that the level sets of h are compact spheres. □

3 Toric Actions and Moment Maps

We now discuss the toric geometry of M. That is, the geometry arising from the effective $\mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}} \cong \mathbb{T}^n_{\mathbb{C}}$-action on M.

3.1 Toric $\mathbb{T}^n$-action on M and Its Moment Map

For clarity, we have broken this section into 4 smaller subsections. Example 9 on $\mathbb{P}^n$ is provided towards the end to illustrate the theory discussed in this section. Since $\mathbb{T}^n_{\mathbb{C}} = \mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}}$ is a quotient, we start by describing orbits of the $\mathbb{T}^d_{\mathbb{C}}$-action on $\mathbb{C}^d$.

Orbits of the $\mathbb{T}^d_{\mathbb{C}}$-action on $\mathbb{C}^d$ There is a one-to-one correspondence between orbits of $\mathbb{T}^d_{\mathbb{C}}$ in $\mathbb{C}^d$ and multi-indices of the form

$$J = \emptyset \text{ or } J = (j_1, \ldots, j_r), \quad 1 \le j_1 < \cdots < j_r \le d. \tag{34}$$

This is because each orbit of $\mathbb{T}^d_{\mathbb{C}}$ is of the form

$$\mathbb{C}^d_J := \{(z_1, \ldots, z_d) \mid z_j = 0 \text{ iff. } j \in J\}. \tag{35}$$

which has complex dimension $d - r$. For $z \in \mathbb{C}^d_J$, the stabilizer of z is the subtorus

$$(\mathbb{T}^d_{\mathbb{C}})_J := \{(\tilde{t}_1, \ldots, \tilde{t}_d) \in \mathbb{T}^d_{\mathbb{C}} \mid \tilde{t}_j = 1 \text{ iff. } j \in \{1, \ldots, d\} \backslash J\}. \tag{36}$$

And note $\mathbb{C}^d_{\emptyset} = (\mathbb{C}^*)^d = \mathbb{T}^d_{\mathbb{C}}$ is the only orbit whose elements have trivial stabilizer.

Later in this section we will see that for the symplectic toric manifold $M = U_a/N_{\mathbb{C}}$, U_a is a union of $\mathbb{C}^d_J$ for a certain collection of J.

Toric $\mathbb{T}^n$-action on M In Sect. 2.2, we obtained the symplectic toric manifold $M = \mu_N^{-1}(a)/N = U_a/N_{\mathbb{C}}$. The holomorphic $\mathbb{T}^d_{\mathbb{C}}$-action is effective on U_a, so there is an open embedding of $\mathbb{T}^d_{\mathbb{C}} \cong (\mathbb{C}^*)^d$ in U_a, and it descends to an open embedding of $\mathbb{T}^n_{\mathbb{C}} = \mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}} \cong (\mathbb{C}^*)^n$ in $M = U_a/N_{\mathbb{C}}$. As shown in the diagram

$$\begin{array}{ccccc} N_{\mathbb{C}} & \xhookrightarrow{\rho_{\mathbb{C}}} & \mathbb{T}^d_{\mathbb{C}} \cong (\mathbb{C}^*)^d & \hookrightarrow & U_a \\ & & \Big\uparrow \beta_{\mathbb{C}} \quad \Big\downarrow \alpha_{\mathbb{C}} & & \Big\downarrow \bar{\pi}_a \\ & & \mathbb{T}^n_{\mathbb{C}} = \mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}} \cong (\mathbb{C}^*)^n & \hookrightarrow & M = U_a/N_{\mathbb{C}} \end{array}, \tag{37}$$

this embedding of $\mathbb{T}^n_{\mathbb{C}} \hookrightarrow M = U_a/N_{\mathbb{C}}$ can be factored through $\mathbb{T}^d_{\mathbb{C}}$ if we choose a group homomorphism

$$\alpha_{\mathbb{C}} : \mathbb{T}^n_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}} \tag{38}$$

that is the right inverse of the surjective group homomorphism $\beta_{\mathbb{C}} : \mathbb{T}^d_{\mathbb{C}} \to \mathbb{T}^n_{\mathbb{C}}$, characterized by $\beta_{\mathbb{C}} \circ \alpha_{\mathbb{C}}(t) = t$ for all $t \in \mathbb{T}^n_{\mathbb{C}}$. In particular, $\alpha_{\mathbb{C}}$ is injective, as the right split of the exact sequence $0 \to N_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}} \to \mathbb{T}^n_{\mathbb{C}} \to 0$, and $\mathbb{T}^d_{\mathbb{C}} = \rho_{\mathbb{C}}(N_{\mathbb{C}}) \oplus \alpha_{\mathbb{C}}(\mathbb{T}^n_{\mathbb{C}})$.

More explicitly, $\alpha_{\mathbb{C}}$ is of the form

$$\alpha_{\mathbb{C}}(t_1, \ldots, t_n) = \left(\prod_{m=1}^{n} t_m^{s_1^m}, \ldots, \prod_{m=1}^{n} t_m^{s_d^m} \right) \in \mathbb{T}^d_{\mathbb{C}} \tag{39}$$

where $s_k^m \in \mathbb{Z}$, $1 \le k \le d$, $1 \le m \le n$, and

$$\sum_{k=1}^{d} v_m^k s_k^{m'} = \delta_m^{m'}, \quad 1 \le m, m' \le n. \tag{40}$$

Taking the differential of $\alpha_{\mathbb{C}}$ at identity, we obtain an injective linear map

$$\mathfrak{t}^n_{\mathbb{C}} = \mathbb{C}^n \xrightarrow{(d\alpha_{\mathbb{C}})_1} \mathfrak{t}^d_{\mathbb{C}} = \mathbb{C}^d$$

given by the matrix

$$A = \begin{bmatrix} s^1_1 & \cdots & s^n_1 \\ \vdots & & \vdots \\ s^1_d & \cdots & s^n_d \end{bmatrix}. \tag{41}$$

Equation (40) is equivalent to $BA = \mathbb{I}_n$, where $\mathbb{I}_n$ is the $n \times n$ identity matrix. Applying $\mathrm{Hom}(-, \mathbb{C}^*)$ to (38), we obtain a surjective map

$$\alpha^*_{\mathbb{Z}} : \mathrm{Hom}(\mathbb{T}^d_{\mathbb{C}}, \mathbb{C}^*) \cong \mathbb{Z}^d \to \mathrm{Hom}(\mathbb{T}^n_{\mathbb{C}}, \mathbb{C}^*) \cong \mathbb{Z}^n. \tag{42}$$

Tensoring (42) by $\mathbb{R}$, we obtain a surjective linear map

$$\alpha^* : \mathbb{R}^d \to \mathbb{R}^n.$$

Note that α^* thus defined is the same as the map $(d\alpha_{\mathbb{C}})^*_1$.

Given $\alpha_{\mathbb{C}}$, there is a $\mathbb{T}^n_{\mathbb{C}} \cong \alpha_{\mathbb{C}}(\mathbb{T}^n_{\mathbb{C}})$-action on U_a given by

$$\begin{aligned} \mathbb{T}^n_{\mathbb{C}} \times U_a &\to U_a \\ (t, z) &\mapsto \alpha_{\mathbb{C}}(t) \cdot z, \end{aligned}$$

where the action of $\alpha_{\mathbb{C}}(t) \in \mathbb{T}^d_{\mathbb{C}}$ on $U_a \subset \mathbb{C}^d$ is the usual component-wise multiplication. Because $\alpha_{\mathbb{C}}(\mathbb{T}^n_{\mathbb{C}}) = \mathbb{T}^d_{\mathbb{C}}/\rho_{\mathbb{C}}(N_{\mathbb{C}})$, this $\mathbb{T}^n_{\mathbb{C}}$-action on U_a descends to a $\mathbb{T}^n_{\mathbb{C}} = \mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}}$-action on $M = U_a/N_{\mathbb{C}}$. That is, for each $z \in U_a$, $t \in \mathbb{T}^n_{\mathbb{C}}$ sends $\widetilde{\pi}_a(z) \in M$ to $\widetilde{\pi}_a(\alpha_{\mathbb{C}}(t) \cdot z)$. This $\mathbb{T}^n_{\mathbb{C}}$-action on M restricts to an effective Hamiltonian $\mathbb{T}^n \cong U(1)^n$-action on the symplectic manifold (M, ω_a), and we discuss its moment map below.

Moment Map of the $\mathbb{T}^n$-action, Moment Polytope, and the Action-Angle Coordinates By the above construction of the $\mathbb{T}^n_{\mathbb{C}}$-action on M (as descending from a composition of $\alpha_{\mathbb{C}}$ and the usual $\mathbb{T}^d_{\mathbb{C}}$-action on $\mathbb{C}^d$), the moment map $\mu_a : M \to \mathbb{R}^n$ for the Hamiltonian $\mathbb{T}^n$-action on M, which is unique to the addition of a constant vector in $\mathbb{R}^n$, is given by

$$\mu_a \circ \pi_a(z) = \alpha^* \circ \mu_{\mathbb{T}^d}(z), \quad z \in \mu_N^{-1}(a), \tag{43}$$

hence

$$\mu_a \circ \widetilde{\pi}_a(z) = \alpha^* \circ \mu_{\mathbb{T}^d} \circ R_a(z), \quad z \in U_a. \tag{44}$$

where R_a was the deformation retract defined in Definition 1. The theorem below provides some important properties about the moment map μ_a.

Theorem 3 (Atiyah [4] and Guillemin-Sternberg [18] Convexity Theorem for Compact Symplectic Toric Manifolds) *If (M^{2n}, ω_a) is a compact symplectic toric manifold carrying an effective Hamiltonian $\mathbb{T}^n$-action as above, with the moment map $\mu_a : M \to \text{Lie}(\mathbb{T}^n)^* = \mathbb{R}^n$, then the level sets of μ_a are connected, and the image $\Delta := \mu_a(M)$ is a convex polytope (known as the moment polytope of M) that is equal to the convex hull of the image under μ_a of the fixed points of the $\mathbb{T}^n$-action.*

Delzant further completes the description of the moment polytope Δ in the following theorem.

Theorem 4 (Delzant [11] Classification Theorem for Compact Symplectic Toric Manifolds) *Compact toric symplectic manifolds (M^{2n}, ω_a) up to $\mathbb{T}^n$-equivariant symplectomorphism are in one-to-one correspondence with Delzant polytopes $\Delta = \mu_a(M^{2n}) \subset \mathbb{R}^n$ up to translation. Delzant polytopes are convex polytopes satisfying*

- *simple, i.e. each vertex has n edges,*
- *rational, i.e. the vectors normal to each facet are generated by a vector in $\mathbb{Z}^n$, and*
- *smooth, i.e. the n integral normal vectors, one for each of the n facets adjacent to a vertex, form a basis of $\mathbb{Z}^n$.*

In other words, the one-to-one correspondence in the above theorem says that if $(M_1, \omega_1, \mathbb{T}^n, \mu_1)$ and $(M_2, \omega_2, \mathbb{T}^n, \mu_2)$ are equivariantly symplectomorphic toric manifolds, then $\mu_1(M_1)$ and $\mu_2(M_2)$ differ by a translation. However, for the same manifold, if we change the basis of $\mathbb{T}^n$, then that is equivalent to a $\text{GL}(n, \mathbb{Z})$-action on $\text{Lie}(\mathbb{T}^n)^* \cong \mathbb{R}^n$, which changes the shape of the polytope $\Delta \subset \mathbb{R}^n$; this is discussed more in Section "Change of Basis of $\mathbb{T}^n_{\mathbb{C}}$ and the Embedding of $\mathbb{T}^n_{\mathbb{C}}$ in Each $\mathbb{C}^n$-chart" and Example 14.

We now explain Delzant's observation that the primitive column vectors of B (i.e. the matrix form of $(d\beta_{\mathbb{C}})_1$ given in Eq. (9)) are inward pointing normal vectors to the facets of Δ. Let $\beta : \mathbb{R}^d \to \mathbb{R}^n$ be the restriction to $\mathbb{R}$ of $(d\beta_{\mathbb{C}})_1 : \mathbb{C}^d \to \mathbb{C}^n$. Applying $\text{Hom}(-, \mathbb{R})$ to β gives us the dual linear map $\beta^* : \mathbb{R}^n \to \mathbb{R}^d$. From Eq. (43) and the fact that $\alpha_{\mathbb{C}}$ is a right inverse of $\beta_{\mathbb{C}}$, we can obtain defining equations of the moment polytope, as follows. As in Eq. (17), each $(\mu_{\mathbb{T}^d})_j = \frac{1}{2}|z_j|^2 + \kappa_j$ for some constant $\kappa_j \in \mathbb{R}$. The relation between κ and a is $(d\rho_{\mathbb{C}})_1^*(\kappa) = -a$, we take the 0 level set of $(d\rho_{\mathbb{C}})_1^* \mu_{\mathbb{T}^d} = (d\rho_{\mathbb{C}})_1^* \sum_{j=1}^d (\frac{1}{2}|z_j|^2 e^j + \kappa_j e^j) = (d\rho_{\mathbb{C}})_1^* (\sum_{j=1}^d \frac{1}{2}|z_j|^2 e^j) - a = 0 = \mu_N - a$ where e^j form the standard basis on $\mathbb{R}^d$. Let $L : \mathbb{R}^n \to \mathbb{R}^d$ be

$$L(\xi)_j := \beta^*(\xi)_j - \kappa_j = \langle v^j, \xi \rangle - \kappa_j = \sum_{k=1}^{n} v_k^j \xi_k - \kappa_j, \tag{45}$$

where v^j, $j = 1, \ldots d$, are the column vectors of B and $(\xi_1, \ldots, \xi_n) \in \Delta = \mu_a(M) \subset \mathbb{R}^n$ are the moment map coordinates $(\mu_{a,1}, \ldots \mu_{a,n})$.

Recall we have the short exact sequence

$$0 \to \mathbb{R}^n \xrightarrow{\beta^*} \mathbb{R}^d \xrightarrow{(d\rho_{\mathbb{C}})_1^*} \mathfrak{n}^* \to 0. \tag{46}$$

So the kernel of $(d\rho_{\mathbb{C}})_1^*$ is precisely the image of β^*.

By the definition of symplectic reduction, the kernel can be identified isomorphically with the values $\Re(z) = \Re(z_1, \ldots, z_d) \in \mathbb{R}^d_{\geq 0}$ such that $\frac{1}{2}\Re(z_j)^2$ satisfy the affine linear equation

$$0 = (d\rho_{\mathbb{C}})_1^* \left(\frac{1}{2}\Re(z_1)^2 + \kappa_1, \ldots, \frac{1}{2}\Re(z_n)^2 + \kappa_n \right) = (d\rho_{\mathbb{C}})_1^* \circ \mu_{\mathbb{T}^d}(\Re(z_1), \ldots, \Re(z_n)),$$

hence $\Re(z) \in \mu_N^{-1}(a)$. In other words, there is a one-to-one correspondence with points $\xi \in \mathbb{R}^n$:

$$\beta^*(\xi) = \frac{1}{2}\Re(z)^2 + \kappa. \tag{47}$$

Thus, for $\xi \in \Delta$ we have that $L(\xi)_j = \frac{1}{2}|\Re(z_j)|^2 \geq 0$, which leads to the following corollary. (A similar way to think about this is if we restrict the domain of $\beta_{\mathbb{C}}$ to the $(\mathbb{C}^*)^n$ part of an affine coordinate chart such as those given in Eq. (67), then $\alpha_{\mathbb{C}}$ is the inverse of $\beta_{\mathbb{C}}$ on that chart. Then $\beta^*(\xi) = \beta^*\alpha^*\mu_{\mathbb{T}^d}(z) = \mu_{\mathbb{T}^d}(z)$.)

Corollary 1 *The moment polytope Δ is given by*

$$\Delta = \left\{ \xi \in \mathbb{R}^n \mid L(\xi)_j = \langle v^j, \xi \rangle - \kappa_j \geq 0 \right\}, \tag{48}$$

where v^j, $j = 1, \ldots d$, are the column vectors of B. So each of the d facets of Δ is given by

$$\mathcal{F}_j = \{L(\xi)_j = \langle v^j, \xi \rangle - \kappa_j = 0\} \cap \Delta, \tag{49}$$

with v^j being a primitive inward pointing normal vector.

Remark 2 This is a result proven in [8, Equation (6)]. The proof here is based off of [16, Theorem 3.3] where $\beta = (d\beta_{\mathbb{C}})_1|_{\mathbb{R}^d}$, $Z_r^\epsilon = \mathbb{R}^d_+ \ni (e^{r_1}, \ldots, e^{r_d})$ where $e^{r_j} = \Re(z_j)$, $\iota = (d\rho_{\mathbb{C}})_1$ and their $\alpha_k = \iota^* e_k$, $s_j = (e^{r_j})^2/2$, $\lambda_j = \kappa_j$ though $\lambda = a$ and one considers the 0 level set, $x \in \mathbb{R}^n$ there is ξ here, and $h = \mu_{\mathbb{T}^d}$.

As a consequence of Theorem 3, we have a fibration $\mu_a : M^{2n} \to \Delta$ whose fibers are connected. For ξ in the boundary $\partial\Delta$, if ξ belongs to a face of codimension r, then $\mu_a^{-1}(\xi)$ consists of fixed points of a r-dimensional subtorus of $\mathbb{T}^n$. So the fibers of μ_a above $\partial\Delta$ are degenerate smaller tori. Any point ξ in the interior of the

polytope $\mathring{\Delta}$ is a regular value of μ_a, meaning that $d\mu_a$ is surjective and so the n Hamiltonian vector fields generated by the components of μ_a are linearly independent, so the fiber above ξ is $\mathbb{T}^n$. Because $\mathbb{T}^n$ is abelian, these vector fields commute, hence the components of μ_a Poisson commute, and so the $\mathbb{T}^n$-orbit is isotropic of dimension n. That means each fiber $\mu_a^{-1}(\xi) \cong \mathbb{T}^n$ is a Lagrangian submanifold of (M, ω). Hence we have the following symplectomorphism identifying $\mu_a^{-1}(\mathring{\Delta})$ with the total space of a Lagrangian torus fibration

$$\left(\mu_a^{-1}(\mathring{\Delta}), \omega_a\right) \cong \left(\mathring{\Delta} \times \mathbb{T}^n, \sum_{j=1}^{n} d\xi_j \wedge d\theta_j\right) \tag{50}$$

where (ξ, θ) is known as the angle-action coordinates, with ξ a coordinate on $\mathring{\Delta}$ and θ a coordinate on the Lie algebra of $\mathbb{T}^n$, so $e^{i\theta} \in \mathbb{T}^n$. This will be discussed in more detail in Sect. 4 on the Kähler potential.

Description of U_a and Complex Toric Coordinates In Sect. 2.1, below Theorem 2, we mentioned Audin [5] and Guillemin [17] gave a description of U_a, which we summarize here.

If f is a face of Δ of codimension r, then f is at the intersection of r facets $\mathcal{F}_{j_1} \cap \cdots \cap \mathcal{F}_{j_r}$. So f corresponds to an ordered multi-index $J_f = (j_1, \ldots, j_r)$ of the form given in Eq. (34). Then U_a that satisfies the properties for Theorem 2 is described by

$$U_a = \bigcup_f \mathbb{C}^d_{J_f}, \tag{51}$$

for f a face of our moment polytope $\Delta = \mu_a(M)$ and $\mathbb{C}^d_{J_f}$ as defined in Eq. (35). For each face f and $z \in \mathbb{C}^d_{J_f}$, the stabilizer of z is $(\mathbb{T}^d_{\mathbb{C}})_{J_f}$ given in Eq. (36).

The open face that is the interior $\mathring{\Delta}$ of the polytope is of codimension 0, and corresponding to that we have the $\mathbb{T}^d_{\mathbb{C}}$-orbit $\mathbb{C}^d_{\emptyset} = (\mathbb{C}^*)^d$ on which the $\mathbb{T}^d_{\mathbb{C}}$-action is free, and thus descends to a single $\mathbb{T}^n_{\mathbb{C}}$-orbit on which $\mathbb{T}^n_{\mathbb{C}}$-action is free. Hence we have

$$\mu_a^{-1}(\mathring{\Delta}) \cong \mathbb{T}^n_{\mathbb{C}}. \tag{52}$$

Definition 2 (Notation for Complex Toric Coordinates) Let $(u_1, \ldots, u_n)$ be the complex coordinates on the Lie algebra $\mathfrak{t}^n_{\mathbb{C}} \cong \mathbb{C}^n$ of $\mathbb{T}^n_{\mathbb{C}} \cong (\mathbb{C}^*)^n$. Also, denote $u_j := x_j + i\theta_j$, so θ_j is the coordinate on the Lie algebra of $\mathbb{T}^n$. The exponential map $\exp : \mathfrak{t}^n_{\mathbb{C}} \to \mathbb{T}^n_{\mathbb{C}}$

$$u = (u_1, \ldots, u_n) \mapsto t = (t_1, \ldots, t_n) = (e^{u_1}, \ldots, e^{u_n}).$$

gives the complex toric coordinates t. □

With this notation, the identification in Eq. (52) is given by

$$\mu_a(t_1, \ldots, t_n) = \mu_a(e^{x_1+i\theta_1}, \ldots, e^{x_n+i\theta_n}) = (\xi_1, \ldots, \xi_n) \in \mathring{\Delta}. \tag{53}$$

This gives the exchange between the complex toric coordinates and the moment map coordinates established by the combination of Eqs. (52) and (50),

$$\mathbb{T}^n_{\mathbb{C}} \cong \mu_a^{-1}(\mathring{\Delta}) \cong \mathring{\Delta} \times \mathbb{T}^n. \tag{54}$$

Example 9 ($\mathbb{P}^n$) We continue Examples 2, 3, and 5 on $\mathbb{P}^n$. For the choice of B in Eq. (14), which corresponds to the $\beta_{\mathbb{C}}$ in Eq. (15), one possible right inverse A and its corresponding map $\alpha_{\mathbb{C}} : \mathbb{T}^n_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}}$ are

$$A = \left[\begin{array}{c} \mathbb{I}_n \\ \hline 0 \cdots 0 \end{array}\right], \quad \alpha_{\mathbb{C}}(t_1, \ldots, t_n) = (t_1, \ldots t_n, 1). \tag{55}$$

(We'll discuss more about the different choices of A in Example 14.) Then $t \in \mathbb{T}^n_{\mathbb{C}}$ acts on $U_a = \mathbb{C}^{n+1} - \{0\}$ via $\alpha_{\mathbb{C}}(t) \in \mathbb{T}^d_{\mathbb{C}}$, which acts on $z \in U_a \subset \mathbb{C}^{n+1}$ by component-wise multiplication

$$(t_1, \ldots, t_n)\cdot(z_1, \ldots, z_{n+1}) = (t_1, \ldots t_n, 1)\cdot(z_1, \ldots, z_{n+1}) = (t_1 z_1, \ldots, t_n z_n, z_{n+1}).$$

This restricts to a Hamiltonian $\mathbb{T}^n = U(1)^n$-action on $\mathbb{C}^{n+1}$ with the standard symplectic form, and the moment map is

$$\alpha^* \circ \mu_{\mathbb{T}^d} : \mathbb{C}^{n+1} \to \mathbb{R}^n, \quad (\alpha^* \circ \mu_{\mathbb{T}^d})(z_1, \ldots, z_{n+1}) = \frac{1}{2}(|z_1|^2, \ldots, |z_n|^2),$$

up to addition of a constant vector in $\mathbb{R}^n$. The above $\mathbb{T}^n_{\mathbb{C}}$-action on $U_a = \mathbb{C}^{n+1} - \{0\}$ descends to a $\mathbb{T}^n_{\mathbb{C}}$-action on $\mathbb{P}^n = U_a/N_{\mathbb{C}} = \mu_N^{-1}(a)/N$, where $N_{\mathbb{C}} = \mathbb{C}^*$ and $N = U(1)$, as

$$(t_1, \ldots, t_n) \cdot [z_1 : \cdots : z_{n+1}] = [t_1 z_1 : \cdots : t_n z_n : z_{n+1}].$$

This $\mathbb{T}^n_{\mathbb{C}}$-action restricts to a Hamiltonian $\mathbb{T}^n$-action on $(\mathbb{P}^n, \omega_a)$ for any $a > 0$ and with the moment map $\mu_a : \mathbb{P}^n \to \mathbb{R}^n$, which (up to a constant in $\mathbb{R}^n$) written in terms of the homogeneous coordinates is

$$\begin{aligned} \mu_a([z_1 : \cdots : z_{n+1}]) &= (\alpha^* \circ \mu_{\mathbb{T}^d} \circ R_a)(z_1, \ldots, z_{n+1}) \\ &= \tfrac{a(|z_1|^2, \ldots, |z_n|^2)}{|z_1|^2 + \cdots + |z_{n+1}|^2} = (\xi_1, \ldots, \xi_n), \end{aligned} \tag{56}$$

where recall that the formula for R_a is given by Eq. (21). The image of this moment map is then the moment polytope

$$\mathring{\Delta} = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n : \xi_1 > 0, \ldots, \xi_n > 0, a - \xi_1 - \cdots - \xi_n > 0\},$$

and one can see that the inward pointing primitive normal vectors to the facets are indeed the column vectors of B. For $n = 2$, this moment polytope is shown in Fig. 1a. □

Example 10 (The Blowup $\mathbb{P}^2(3)$ of $\mathbb{P}^2$ at the Three Torus Fixed Points) The projective plane blown up at one point is isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, see [15, p 479–480] or note that blow-up of a complex surface at a vertex of the polytope replaces the point with $\mathbb{P}^1$ hence the triangle for $\mathbb{P}^2$ becomes a quadrilateral. Then we can perform two toric blow-ups at the fixed points corresponding to two opposite vertices of the rectangle which is the moment map of $\mathbb{P}^1 \times \mathbb{P}^1$, to obtain a hexagon. (See [12, p 40–41] for the toric blow up and [31, Section 7] for the symplectic blow up.) So in this case we know $n = 2$ and $d = 6$. Here is one choice of B, giving the moment polytope in Fig. 1c.

$$B = \begin{bmatrix} 0 & 0 & -1 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\beta_{\mathbb{C}}(\widetilde{t}_1, \ldots, \widetilde{t}_6) = (\widetilde{t}_3^{-1}\widetilde{t}_4\widetilde{t}_5\widetilde{t}_6^{-1}, \widetilde{t}_1\widetilde{t}_2^{-1}\widetilde{t}_5^{-1}\widetilde{t}_6),$$

$$\rho_{\mathbb{C}}(\lambda_1, \ldots, \lambda_4) = (\lambda_1\lambda_3\lambda_4^{-1},\ \lambda_1,\ \lambda_2,\ \lambda_2\lambda_3^{-1}\lambda_4,\ \lambda_3,\ \lambda_4).$$

We can make the following choice for $\alpha_{\mathbb{C}}$:

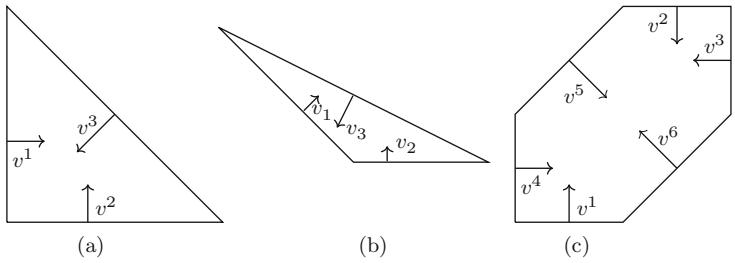


Fig. 1 In the three figures above, the labeling, v^j, of the inward pointing normal vectors to the facets matches with the ordering of the columns of B in Examples 9 ($n = 2$) and 14, B' in Example 14, and B in Example 10, respectively. (**a**) Moment polytope for $\mathbb{P}^2$. (**b**) $\mathbb{P}^2$, different $\mathbb{T}^2$ basis. (**c**) Moment polytope for $\mathbb{P}^2(3)$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \alpha_{\mathbb{C}}(t_1, t_2) = (t_2, 1, 1, t_1, 1, 1).$$

Again, the choice of $\alpha_{\mathbb{C}}$ is not unique. Here $\mu_N : \mathbb{C}^6 \to \mathbb{R}^4$ is given by

$$\mu_N(z_1, \ldots, z_6) = \tfrac{1}{2}\Big(|z_1|^2 + |z_2|^2, |z_3|^2 + |z_4|^2, |z_1|^2 - |z_4|^2 + |z_5|^2, -|z_1|^2 + |z_4|^2 + |z_6|^2\Big),$$

and we will not try finding R_a for this example here. It involves solving a third degree polynomial. We will get a feel for the complexity of solving for R_a in Sect. 3.5 below for $K_{\mathbb{P}^1}$; the formula gets very complicated. □

3.2 Canonical Bundle K_M Continued

In Sect. 2.3, from $M = U_a/N_{\mathbb{C}} = \mu_N^{-1}(a)/N$, we constructed K_M as quotients $K_M = U_a^+/N_{\mathbb{C}} = (\mu_N^+)^{-1}(a)/N$, where $U_a^+ = U_a \times \mathbb{C}$. In this section, we written down the general formulation for the moment map of the $\mathbb{T}^{n+1}$-action on K_M. Extending the theory for the construction of the moment map from M to K_M is straightforward, the main difficulty is keeping track of the extra p coordinate corresponding to the fiber direction. However, writing down the moment map explicitly is a hard problem in general, as we'll illustrate in Sect. 3.5.

Definition 3 We define $\alpha_{\mathbb{C}}^+ : \mathbb{T}_{\mathbb{C}}^{n+1} \to \mathbb{T}_{\mathbb{C}}^{d+1}$ by

$$\alpha_{\mathbb{C}}^+(t_1, \ldots, t_n, t_{n+1}) = \left(\prod_{m=1}^n t_m^{s_1^m}, \ldots, \prod_{m=1}^n t_m^{s_d^m}, \prod_{m=1}^n t_m^{-\sum_{k=1}^d s_k^m} \cdot t_{n+1}\right). \tag{57}$$

Then $\beta_{\mathbb{C}}^+ \circ \alpha_{\mathbb{C}}^+(t) = t$ for all $t \in \mathbb{T}_{\mathbb{C}}^{n+1}$, i.e. $\alpha_{\mathbb{C}}^+$ is a right inverse of the surjective group homomorphism $\beta_{\mathbb{C}}^+ : \mathbb{T}_{\mathbb{C}}^{d+1} \to \mathbb{T}_{\mathbb{C}}^{n+1}$. Taking the differential of $\alpha_{\mathbb{C}}^+$ at the identity, we obtain an injective linear map

$$\mathfrak{t}_{\mathbb{C}}^{n+1} = \mathbb{C}^{n+1} \xrightarrow{(d\alpha_{\mathbb{C}}^+)_1} \mathfrak{t}_{\mathbb{C}}^{d+1} = \mathbb{C}^{d+1} \tag{58}$$

given by the matrix

$$A^+ = \begin{bmatrix} s_1^1 & \cdots & s_1^n & 0 \\ \vdots & & & \vdots \\ s_d^1 & \cdots & s_d^n & 0 \\ -\sum_{k=1}^{d} s_k^1 & \cdots & -\sum_{k=1}^{d} s_k^n & 1 \end{bmatrix}.$$

Then $B^+A^+ = \mathbb{I}_{n+1}$, where $\mathbb{I}_{n+1}$ is the $(n+1) \times (n+1)$ identity matrix. Below we write the matrix for a given choice of $\alpha_{\mathbb{C}}^+$ for a couple of examples:

Example 11 ($K_{\mathbb{P}^n}$)

$$B^+ = \left[\begin{array}{c|cc} & -1 & 0 \\ \mathbb{I}_n & \vdots & \vdots \\ & -1 & 0 \\ \hline 1 \ldots 1 & 1 & 1 \end{array}\right], \quad A^+ = \left[\begin{array}{c|c} & 0 \\ \mathbb{I}_n & \vdots \\ & 0 \\ \hline 0 \ldots 0 & 0 \\ \hline -1 \ldots -1 & 1 \end{array}\right].$$

This choice, with all 1's in the last row of B^+, is often used in the polytopes defining SYZ mirrors in mirror symmetry (discussed in Sect. 5). We can make other choices for B^+, and we discuss one such choice in Example 15. □

Example 12 ($K_{\mathbb{P}^2(3)}$) Extending Example 10 on $\mathbb{P}^2(3)$, for $K_{\mathbb{P}^2(3)}$ we can take

$$B^+ = \begin{bmatrix} 0 & 0 & -1 & 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, Q^+ = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -2 & -1 & -1 \end{bmatrix}, A^+ = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1 & 1 \end{bmatrix},$$

and the maps are expressed as follows

$$\beta_{\mathbb{C}}^+(\tilde{t_1}, \tilde{t_2}, \tilde{t_3}, \tilde{t_4}, \tilde{t_5}, \tilde{t_6}, \tilde{t_7}) = (\tilde{t_3}^{-1}\tilde{t_4}\tilde{t_5}\tilde{t_6}^{-1}, \tilde{t_1}\tilde{t_2}^{-1}\tilde{t_5}^{-1}\tilde{t_6}, \tilde{t_1}\tilde{t_2}\tilde{t_3}\tilde{t_4}\tilde{t_5}\tilde{t_6}\tilde{t_7}),$$

$$\rho_{\mathbb{C}}^+(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (\lambda_1\lambda_3\lambda_4^{-1}, \lambda_1, \lambda_2, \lambda_2\lambda_3^{-1}\lambda_4, \lambda_3, \lambda_4, \lambda_1^{-2}\lambda_2^{-2}\lambda_3^{-1}\lambda_4^{-1}), \text{ and}$$

$$\alpha_{\mathbb{C}}^+ = (t_2, 1, 1, t_1, 1, 1, t_1^{-1}t_2^{-1}t_3).$$

□

Furthermore, we have an analogue in the K_M case of Eqs. (43) and (44) by the same arguments. They are

$$\mu_a^+ \circ \pi_a^+(z) = (\alpha^+)^* \circ \mu_{\mathbb{T}^{d+1}}(z), \quad z \in (\mu_N^+)^{-1}(a), \tag{59}$$

and

$$\mu_a^+ \circ \widetilde{\pi}_a^+(z) = (\alpha^+)^* \circ \mu_{\mathbb{T}^{d+1}} \circ R_a^+(z), \quad z \in U_a \times \mathbb{C}. \tag{60}$$

The image of μ_a^+ is again a convex polyhedron Δ^+.

3.3 Holomorphic Coordinate Charts for M

In this section, we describe holomorphic coordinate charts for the complex manifold $M = U_a/N_{\mathbb{C}}$. This will help us better understand both M and the $\mathbb{T}^n_{\mathbb{C}}$-action. Examples 13 and 14 are provided at the end for illustrating the content of this section.

Open Covering of M Any vertex v of the moment polytope Δ is at the intersection of n facets $\mathcal{F}_{j_1} \cap \cdots \cap \mathcal{F}_{j_n}$, where $J_v = (j_1, \ldots, j_n)$ is an ordered multi-index of the form in Eq. (34) and each facet is given by Eq. (49). For any $z \in \mathbb{C}^d_{J_v}$, there is an open neighborhood of z

$$\widetilde{V}_{J_v} := \{(z_1, \ldots, z_d) \in \mathbb{C}^d \mid z_j \neq 0 \text{ if } j \notin J_v\} \subset U_a, \tag{61}$$

and this descends to an open set

$$V_{J_v} := \{[z_1 : \ldots : z_d] \in M = U_a/N_{\mathbb{C}} \mid z_j \neq 0 \text{ if } j \notin J_v\} \subset M = U_a/N_{\mathbb{C}}. \tag{62}$$

Open sets of this kind, one for each vertex v of Δ, give an open covering of

$$M = \bigcup_v V_{J_v}. \tag{63}$$

Holomorphic Charts and Inhomogeneous Coordinates For each V_{J_v}, we have a biholomorphic map $\varphi^{J_v} : V_{J_v} \to \mathbb{C}^n$. To describe φ^{J_v}, let us assume, without loss of generality by renumbering the facets, that $J_v = (1, \ldots, n)$.

Then

$$V_{J_v} = \left\{[z_1 : \ldots : z_d] \in M = U_a/N_{\mathbb{C}} \mid z_j \neq 0 \text{ if } j \in \{n+1, \ldots, d\}\right\}.$$

Now, we would like to scale $z_{n+1}, \ldots, z_d$ to 1 using the $N_{\mathbb{C}}$-action and the notations for it in Eq. (5). Pick $\lambda_1, \ldots \lambda_{d-n} \in N_{\mathbb{C}}$ such that

$$\prod_{\ell=1}^{d-n} \lambda_\ell^{Q_{n+1}^\ell} = z_{n+1}^{-1}, \quad \ldots, \quad \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_d^\ell} = z_d^{-1}. \tag{64}$$

This is doable for the following reason. We know that the point $(0, \ldots, 0, 1, \ldots, 1)$, with the first n entries being 0 and the last $d-n$ entries being 1, is a point in U_a. The $N_{\mathbb{C}}$-action on this point is given by

$$(\lambda_1, \ldots, \lambda_{d-n}) \cdot (0, \ldots, 0, 1, \ldots, 1) = \left(0, \ldots, 0, \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_{n+1}^\ell}, \ldots, \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_d^\ell}\right).$$

Because the $N_{\mathbb{C}}$-action is free, from the above equation we see that the following map is an isomorphism

$$\begin{aligned} N_{\mathbb{C}} = (\mathbb{C}^*)^{d-n} &\to (\mathbb{C}^*)^{d-n} \\ (\lambda_1, \ldots, \lambda_{d-n}) &\mapsto \left(\prod_{\ell=1}^{d-n} \lambda_\ell^{Q_{n+1}^\ell}, \ldots, \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_d^\ell}\right). \end{aligned}$$

This shows that there is a choice of $\lambda_1, \ldots, \lambda_{d-n}$ such that Eq. (64) holds. Then we have

$$V_{J_v} = \left\{ \left[y_1^{J_v} = \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_1^\ell} z_1 : \ldots : y_n^{J_v} = \prod_{\ell=1}^{d-n} \lambda_\ell^{Q_n^\ell} z_n : 1 : \cdots : 1 \right] \right\} \subset M = U_a / N_{\mathbb{C}}. \tag{65}$$

So we get the map

$$\varphi^{J_v} : V_{J_v} \to \mathbb{C}^n, \quad [z_1 : \cdots : z_d] \mapsto y^{J_v} = (y_1^{J_v}, \ldots, y_n^{J_v}). \tag{66}$$

The origin $y^{J_v} = 0$ of this $\mathbb{C}^n$-chart corresponds to the preimage of the vertex v,

$$p^{J_v} = \mu_a^{-1}(v) = [0 : \cdots : 0 : 1 : \cdots : 1] \in V_{J_v},$$

which is a fixed point of the $\mathbb{T}_{\mathbb{C}}^n$-action.

Equations (65) and (66) identify $\varphi^{J_v}(V_{J_v}) = \mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}]$ with the affine subspace

$$U_{J_v} := \{(z_1, \ldots, z_d) \mid z_j = 1 \text{ if } j \notin J_v\} \subset \widetilde{V}_{J_v} \subset \mathbb{C}^d, \tag{67}$$

with $y_k^{J_v} = z_{j_k}$ for $k = 1, \ldots, n$. Again, when using $J_v = (1, \ldots, n)$, we have $U_{J_v} = \{(z_1, \ldots, z_n, 1, \ldots, 1)\} \subset \mathbb{C}^d$, and its identification with $\mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}]$ is given by $z_k = y_k^{J_v}$, $k = 1, \ldots, n$. Because of this identification, below in section we will often conveniently write $U_{J_v} = \mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}]$.

To note the terminology, the coordinates $z = (z_1, \ldots, z_d) \in U_a \subset \mathbb{C}^d$ is called the homogeneous coordinates on M. For each vertex v, the coordinates $y^{J_v} = (y_1^{J_v}, \ldots, y_n^{J_v})$ on the $\mathbb{C}^n$-chart associated to v, more precisely on the image of the chart map $\varphi^{J_v}(V_{J_v})$, is called the inhomogeneous coordinates. Because of the identification of $(y_k^{J_v})$ with the coordinates $(z_{j_k}) \in U_{J_v}$ on the affine subspace, y^{J_v} are also called the affine coordinates in the literature.

Change of Basis of $\mathbb{T}^n_{\mathbb{C}}$ and the Embedding of $\mathbb{T}^n_{\mathbb{C}}$ in Each $\mathbb{C}^n$-chart Our matrix B is a linear transformation $\mathfrak{t}^d_{\mathbb{C}} \to \mathfrak{t}^n_{\mathbb{C}}$, where $\mathfrak{t}^n_{\mathbb{C}} = \mathrm{Lie}(\mathbb{T}^n_{\mathbb{C}})$, so changing the basis of $\mathfrak{t}^n_{\mathbb{C}}$ is equivalent to performing a row operation on B to get a new matrix $\widehat{B}$.

Let us continue to assume that v is the vertex such that $J_v = (1, \ldots, n)$, so the first n column vectors $\{v^1, \ldots, v^n\}$ of B are linearly independent. The remaining column vectors $v^{n+1}, \ldots, v^d$ can then be expressed as linear combinations of $\{v^1, \ldots, v^n\}$ as

$$v^{n+k} = \sum_{j=1}^{n} c_{jk} v^j, \tag{68}$$

for $k = 1, \ldots, d - n$. Now let us take

$$\{v^1, \ldots, v^n\}$$

to be the new basis for $\mathfrak{t}^n_{\mathbb{C}}$. We see that B in the old basis is now

$$\widehat{B} = PB = \left[\, \mathbb{I}_n \,\middle|\, C \,\right] \tag{69}$$

in the new basis, where P is the $n \times n$ change of basis matrix with P^{-1} being the matrix whose columns are $v^1, \ldots, v^n$, in that order, and $C = (c_{jk})$ is a $n \times (d - n)$ matrix consisting of the constants (c_{jk}) from Eq. (68) above. Note for the moment polytope, since the first n columns of B are the vectors normal to the facets meeting at the vertex v, this change of basis of $\mathfrak{t}^n_{\mathbb{C}}$ transforms those normal vectors from $\{v^1, \ldots, v^n\}$ to the column vectors of $\mathbb{I}_n$. In other words, it straightens out that corner.

One choice of right inverse to $\widehat{B}$ is

$$\widehat{A}^{J_v} = \left[\begin{array}{c} \mathbb{I}_n \\ \hline 0 \end{array}\right].$$

This is not the only choice of right inverse to $\widehat{B}$, but this is the choice that equips the $\mathbb{C}^n$-chart at this vertex v with the standard $\mathbb{T}^n_{\mathbb{C}}$-action. Indeed, following the setup in Eqs. (39) and (41), this $\widehat{A}^{J_v}$ corresponds to $\widehat{\alpha}^{J_v}_{\mathbb{C}}$ where

$$\widehat{\alpha}^{J_v}_{\mathbb{C}}(t_1, \ldots, t_n) = (t_1, \ldots, t_n, 1, \ldots, 1) \in U_{J_v} \subset U_a.$$

The $\widehat{\alpha}^{J_v}_{\mathbb{C}}(\mathbb{T}^n_{\mathbb{C}})$-action on U_a descends to the following standard $\mathbb{T}^n_{\mathbb{C}}$-action on $\varphi^{J_v}(V_{J_v}) = \mathbb{C}^n$

$$(t_1, \ldots, t_n) \cdot \left(y_1^{J_v}, \ldots, y_n^{J_v}\right) = \left(t_1 y_1^{J_v}, \ldots, t_n y_n^{J_v}\right).$$

The point $y^{J_v} = 0$ is the toric fixed point, which corresponds to the moment map preimage $p^{J_v} = \mu_a^{-1}(v)$ of the vertex $v \in \Delta$. This action is free on $(\mathbb{C}^*)^n$, i.e. where $y_j^{J_v} \neq 0$ for all $j = 1, \ldots, n$, so this action embeds $\mathbb{T}^n_{\mathbb{C}} \cong (\mathbb{C}^*)^n$ via $\widehat{\alpha}^{J_v}_{\mathbb{C}}$ in $\varphi^{J_v}(V_{J_v}) = \mathbb{C}^n$. Combining this embedding with the identification of U_{J_v} with $\mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}]$ we discussed around Eq. (67), we have that $\widehat{\alpha}^{J_v}_{\mathbb{C}}$ gives the following isomorphism

$$\begin{array}{rccl} \widehat{\alpha}^{J_v}_{\mathbb{C}} : \mathbb{T}^n_{\mathbb{C}} & \xrightarrow{\cong} & U_{J_v} \backslash \{z_k = 0\}_{k=1}^n = \mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}] \backslash \{y_k^{J_v} = 0\}_{k=1}^n \\ t_k & \mapsto & t_k = y_k^{J_v}, \quad \text{for } k = 1, \ldots n. \end{array}$$

The diagram below summarizes the change of basis we made on $\mathfrak{t}^n_{\mathbb{C}} = \mathrm{Lie}(\mathbb{T}^n_{\mathbb{C}})$,

$$\begin{array}{ccccc} \mathfrak{t}^d_{\mathbb{C}} & \xrightarrow{B} & \mathfrak{t}^n_{\mathbb{C}} \cong \bigoplus_{j=1}^n \mathbb{C} e^j & \xrightarrow{A^{J_v} = \widehat{A}^{J_v} P} & \mathfrak{t}^d_{\mathbb{C}} \\ \| & & \downarrow P & & \| \\ \mathfrak{t}^d_{\mathbb{C}} & \xrightarrow{\widehat{B} = PB} & \mathfrak{t}^n_{\mathbb{C}} \cong \bigoplus_{j=1}^n \mathbb{C} v^j & \xrightarrow{\widehat{A}^{J_v}} & \mathfrak{t}^d_{\mathbb{C}} \end{array}, \tag{70}$$

where e^j is the jth standard basis vector. For the chosen $\widehat{A}^{J_v}$ above, the corresponding A^{J_v} in our original basis of $\mathfrak{t}^n_{\mathbb{C}}$ satisfies $BA^{J_v} = \widehat{B}\widehat{A}^{J_v} = \mathbb{I}_n$, so

$$A^{J_v} = \widehat{A}^{J_v} P = \left[\begin{array}{c} P \\ \hline 0 \end{array}\right]. \tag{71}$$

Corresponding to this A^{J_v}, we get an $\alpha^{J_v}_{\mathbb{C}}$. This is the choice of $\alpha^{J_v}_{\mathbb{C}}$ such that the $\alpha^{J_v}_{\mathbb{C}}(\mathbb{T}^n_{\mathbb{C}})$-action on U_a descends to the standard $\mathbb{T}^n_{\mathbb{C}}$-action on the $\varphi^{J_v}(V_{J_v}) = \mathbb{C}^n$-chart. The point $0 \in \mathbb{C}^n$ in this chart is the fixed point of this action and it corresponds to the moment map preimage $p^{J_v} = \mu_a^{-1}(v)$ of the vertex. This $\mathbb{T}^n_{\mathbb{C}}$-action is free on the $(\mathbb{C}^*)^n$ part of the chart, which gives the isomorphism

$$\begin{array}{rccl} \alpha^{J_v}_{\mathbb{C}} : \mathbb{T}^n_{\mathbb{C}} & \xrightarrow{\cong} & U_{J_v} \backslash \{z_k = 0\}_{k=1}^n = \mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}] \backslash \{y_k^{J_v} = 0\}_{k=1}^n \\ t_k & \mapsto & \alpha^{J_v}_{\mathbb{C}}(t)_k = y_k^{J_v}, \quad \text{for } k = 1, \ldots n. \end{array} \tag{72}$$

This relationship is very useful because it gives an identification of the inhomogeneous coordinates y^{J_v} on each of the $\mathbb{C}^n$-chart with the complex toric coordinates $t \in \mathbb{T}^n_{\mathbb{C}}$.

Transition Functions Suppose v is a vertex at the intersection of $\bigcap_{\ell=1}^n \mathcal{F}_{j_\ell}$ and $\tilde{v}$ is another vertex at the intersection of $\bigcap_{k=1}^n \mathcal{F}_{j_k}$, so $J_v = (j_\ell)_{\ell=1}^n$ and $J_{\tilde{v}} = (j_k)_{k=1}^n$. Then the overlap $\varphi^{J_{\tilde{v}}}(V_{J_v} \cap V_{J_{\tilde{v}}})$ can be identified with

$$\varphi^{J_{\tilde{v}}}(V_{J_v} \cap V_{J_{\tilde{v}}}) = U_{J_{\tilde{v}}} \backslash \{z_j = 0\}_{j \notin J_v} = \mathbb{C}[y_1^{J_{\tilde{v}}}, \ldots, y_n^{J_{\tilde{v}}}] \backslash \{y_\ell^{J_{\tilde{v}}} = 0\}_{j_\ell \notin J_v}.$$

Similarly, $\varphi^{J_v}(V_{J_v} \cap V_{J_{\tilde{v}}})$ can be identified with

$$\varphi^{J_v}(V_{J_v} \cap V_{J_{\tilde{v}}}) = U_{J_v} \backslash \{z_j = 0\}_{j \notin J_{\tilde{v}}} = \mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}] \backslash \{y_k^{J_v} = 0\}_{j_k \notin J_{\tilde{v}}}.$$

Note that $\varphi^{J_{\tilde{v}}}(V_{J_v} \cap V_{J_{\tilde{v}}}) = (\mathbb{C}^*)^n$ if $J_v \cap J_{\tilde{v}} = \emptyset$; otherwise, it contains $(\mathbb{C}^*)^n$ as a proper subset.

By Eq. (72), we see that $\alpha_{\mathbb{C}}^{J_v}$ and $\alpha_{\mathbb{C}}^{J_{\tilde{v}}}$ give the following isomorphisms,

$$\mathbb{T}^n_{\mathbb{C}} \overset{\alpha_{\mathbb{C}}^{J_{\tilde{v}}}}{\cong} U_{J_{\tilde{v}}} \backslash \{z_j = 0\}_{j \in J_{\tilde{v}}} = \mathbb{C}[y_1^{J_{\tilde{v}}}, \ldots, y_n^{J_{\tilde{v}}}] \backslash \{y_j^{J_{\tilde{v}}} = 0\}_{j=1}^n \subset \varphi^{J_{\tilde{v}}}(V_{J_v} \cap V_{J_{\tilde{v}}}),$$

$$\mathbb{T}^n_{\mathbb{C}} \overset{\alpha_{\mathbb{C}}^{J_v}}{\cong} U_{J_{\tilde{v}}} \backslash \{z_j = 0\}_{j \in J_v} = \mathbb{C}[y_1^{J_v}, \ldots, y_n^{J_v}] \backslash \{y_j^{J_v} = 0\}_{j=1}^n \subset \varphi^{J_v}(V_{J_v} \cap V_{J_{\tilde{v}}}).$$

Note again that the "$\subset$" above is "$=$" if $J_v \cap J_{\tilde{v}} = \emptyset$; otherwise, it is "$\subsetneq$".

So the transition function

$$\varphi^{J_v} \circ (\varphi^{J_{\tilde{v}}})^{-1} : \varphi^{J_{\tilde{v}}}(V_{J_v} \cap V_{J_{\tilde{v}}}) \longrightarrow \varphi^{J_v}(V_{J_v} \cap V_{J_{\tilde{v}}})$$

restricted to the $(\mathbb{C}^*)^n \cong \mathbb{T}^n_{\mathbb{C}}$ part is exactly $\alpha_{\mathbb{C}}^{J_v} \circ (\alpha_{\mathbb{C}}^{J_{\tilde{v}}})^{-1}$.

Now let $J_v = (1, \ldots, n)$ and below we discuss more explicitly the relationship between the inhomogeneous coordinates y^{J_v} and $y^{J_{\tilde{v}}}$ in the overlap of the two charts. Denote by $\tilde{v}^k := v^{j_k}$, $k = 1, \ldots, n$, the column vector of B that is the normal vector to the facet $\mathcal{F}_{j_k}$. Then

$$\tilde{v}^j = \sum_{k=1}^n d_{jk} v^k, \tag{73}$$

where $D = (d_{jk}) \in \mathrm{GL}(n, \mathbb{Z})$. In fact, because $\det(DD^{-1}) = \det(D)\det(D^{-1}) = 1$, and that $\det(D)$ and $\det(D^{-1})$ are both integers, we must have $\det(D) = \pm 1$. We can further assume that $\det(D) = 1$, i.e. $D \in \mathrm{SL}(n, \mathbb{Z})$, since otherwise if $\det(D) = -1$, we can interchange $\tilde{v}^1$ and $\tilde{v}^2$ to make $\det(D) = 1$.

Then the transition map for the inhomogeneous coordinates is given by

$$y_k^{J_v} = \prod_{j=1}^{n} (y_j^{J_{\tilde{v}}})^{d_{jk}} \quad \text{for } k = 1, \ldots, n. \tag{74}$$

Summary To summarize: in [16] and [5], the starting point is a Delzant moment polytope, from which the symplectic and complex geometry of M can be read off. Let the inward normals to its facets $\mathcal{F}_1, \ldots, \mathcal{F}_d$ be $v^1, \ldots, v^d$. Because the space of normal vectors to each facet of Δ is generated by a vector in $\mathbb{Z}^n$, we can always choose each v^j to be a primitive integral vector that is the inward pointing normal vector to the facet. There is a standard open covering of M, corresponding to vertices of Δ, and transition functions, which are obtained by the following pieces of information:

- Any subset $J \subseteq \{1, \ldots, d\}$ indexes orbits $\mathbb{T}^d_{\mathbb{C}} \cdot (z_1, \ldots, z_d) \subset \mathbb{C}^d$ such that $z_j = 0$ when $j \in J$.
- Each vertex v of the moment polytope is the intersection $\mathcal{F}_{j_1} \cap \cdots \cap \mathcal{F}_{j_n}$ of n facets of the polytope Δ so corresponds to an index set $J_v := (j_1, \ldots, j_n) \subset \{1, \ldots, d\}$.
- Let $p^{J_v} := \mu_a^{-1}(\mathcal{F}_{j_1} \cap \ldots \cap \mathcal{F}_{j_n})$ be a toric fixed point of the residual $\mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}} \cong \mathbb{T}^n_{\mathbb{C}}$-action on M, obtained by quotienting $U_a \subset \mathbb{T}^d_{\mathbb{C}}$ by the action of $N_{\mathbb{C}}$.
- $(\mathbb{C}^*)^n$ is the dense open $\mathbb{T}^n_{\mathbb{C}}$-orbit corresponding to the $J = \emptyset$ orbit, modulo $N_{\mathbb{C}}$, in the quotient $U_a/N_{\mathbb{C}} = M$.
- For the index set J_v, the primitive inward pointing normal vectors v^{j_k}, $k = 1, \ldots, n$, give the j_kth column of B for each k. Different choices of B correspond to $\mathrm{GL}(n, \mathbb{Z})$ transformations of Δ.
- $\alpha_{\mathbb{C}}^{J_v} : \mathbb{T}^n_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}}$ passes to an embedding $\widetilde{\pi}_a \circ \alpha_{\mathbb{C}}^{J_v} : \mathbb{T}^n_{\mathbb{C}} \xrightarrow{\alpha_{\mathbb{C}}^{J_v}} U_a \xrightarrow{\widetilde{\pi}_a} M$ into an open $\mathbb{C}^n$ set centered at p^{J_v} in M, which we may denote $U_{J_v} \subset M$:
  - V_{J_v} consists of points $[z_1 : \ldots : z_d] \in U_a/N_{\mathbb{C}} = M$ such that the components of $(z_{j_1}, \ldots, z_{j_n}) \in \mathbb{C}^n$ may go to 0 and the other homogeneous coordinates are always nonzero,

$$V_{J_v} := \{[z_1 : \ldots : z_d] \in U_a/N_{\mathbb{C}} = M \mid z_j \neq 0 \text{ if } j \notin J_v\},$$

  - we can express transition functions in terms of the $y_k^{J_v}$ on the overlap of two $\mathbb{C}^n$ charts; for example in $\mathbb{P}^2$ in Fig. 1a corresponding to the vertex at the right angle we have $\alpha_{\mathbb{C}}(t_1, t_2) = (t_1, t_2, 1)$ extends to $\mathbb{C}^2$ and passing to the quotient we obtain $\mathbb{C}^2 \hookrightarrow \{[t_1 : t_2 : 1] \in \mathbb{P}^2 \mid (t_1, t_2) \in \mathbb{C}^2\} \subset \mathbb{P}^2$, where the transition function is defined on $\mathbb{C} \times \mathbb{C}^*$ (for more details see Example 14),
  - we choose the unique representative for each $[z_1 : \ldots : z_d] \in V_{J_v}$ which has been scaled to have 1's in all the $d - n$ strictly nonzero coordinates,
  - the remaining scaled coordinates are the affine (inhomogeneous) coordinates $y_1^{J_v}, \ldots, y_n^{J_v}$ on U_{J_v},

- the V_{J_v} with charts U_{J_v} give a standard open cover of M indexed by the toric fixed points v, in bijection with vertices of the moment polytope at the intersection of the J_vth facets, and
- transition functions on the dense open orbit of M are of the form $\alpha_{\mathbb{C}}^{J_v} \circ (\alpha_{\mathbb{C}}^{J_{\tilde{v}}})^{-1}$, where $\mathbb{T}_{\mathbb{C}}^n$ embeds into each U_{J_v} under the corresponding $\alpha_{\mathbb{C}}$ and we may use coordinates $(y_1^{J_v}, \ldots, y_n^{J_v})$ and $(y_1^{J_{\tilde{v}}}, \ldots, y_n^{J_{\tilde{v}}})$ to express the transition maps, again see Example 14 for the case of $\mathbb{P}^2$.

Examples

Example 13 ($\mathbb{P}^n$) We continue Example 9. Consider the vertex v of Δ given by $(\xi_1, \ldots, \xi_n) = (0, \ldots, 0)$, i.e. $v = \mathcal{F}_1 \cap \cdots \cap \mathcal{F}_n$ is the intersection of the first n facets whose normal vectors are the first n column vectors of B in Eq. (14). The $\mathbb{C}^n$-chart corresponding to this vertex is then

$$\begin{aligned} V_{J_v} &= \{[z_1 : \cdots : z_{n+1}] \in \mathbb{P}^n \mid z_{n+1} \neq 0\} \\ &= \left\{ \left[y_1 = \tfrac{z_1}{z_{n+1}} : \cdots : y_n = \tfrac{z_n}{z_{n+1}} : 1 \right] \in \mathbb{P}^n \right\}, \end{aligned}$$

and $\varphi^{J_v} : V_{J_v} \to \mathbb{C}^n$ given by $[z_1 : \cdots : z_{n+1}] \mapsto (y_1, \ldots y_n)$ is a biholomorphic map, with $y = 0$ being a fixed point of the $\mathbb{T}_{\mathbb{C}}^n$-action. The image $\varphi^{J_v}(V_{J_v}) = \mathbb{C}[y_1, \ldots, y_n]$ can be identified with the affine subspace

$$U_{J_v} = \{(z_1, \ldots z_n, 1)\} \subset \mathbb{C}^{n+1}$$

with $y_k = z_k$ for $k = 1, \ldots, n$.

Following the diagram in Eq. (37), we see that the map $\alpha_{\mathbb{C}}$ chosen in Eq. (55) gives an embedding of $\mathbb{T}_{\mathbb{C}}^n$ in $\mathbb{P}^n$ by

$$\mathbb{T}_{\mathbb{C}}^n \overset{\alpha_{\mathbb{C}}}{\hookrightarrow} U_{J_v} \to \mathbb{P}^n, \quad (t_1, \ldots, t_n) \mapsto (t_1, \ldots, t_n, 1) \mapsto [t_1 : \cdots : t_n : 1].$$

We see that $\alpha_{\mathbb{C}}$ is an isomorphism onto $U_{J_v} \backslash \{z_k = 0\}_{k=1}^n$. This choice of $\alpha_{\mathbb{C}}$, along with the identification of U_{J_v} with $\mathbb{C}[y_1, \ldots, y_n]$, gives the relation $t_k = y_k$ between the complex toric coordinates t and the inhomogeneous coordinates y, which illustrates the identification stated in Eq. (72). This identification, along with the notations given in Definition 2, $t_j = e^{u_j} = e^{x_j + i\theta_j}$, we can write Eq. (56) as

$$\mu_a(y_1, \ldots, y_n) = (\xi_1, \ldots, \xi_n) = \frac{a(|y_1|^2, \ldots, |y_n|^2)}{|y_1|^2 + \cdots + |y_n|^2 + 1} = \frac{a(e^{2x_1}, \ldots, e^{2x_n})}{e^{2x_1} + \cdots + e^{2x_n} + 1},$$

so

$$(x_1, \ldots, x_n) = \frac{1}{2}\Big(\log \xi_1 - \log(a - \xi_1 - \cdots - \xi_n), \ldots, \log \xi_n - \log(a - \xi_1 - \cdots - \xi_n) \Big).$$

□

Example 14 ($\mathbb{P}^2$) We specialize the above example to $\mathbb{P}^2$ to further demonstrate the change of coordinates between charts. The last paragraph of this example illustrates the change of basis of $\mathbb{T}^2_{\mathbb{C}}$. Using the matrix B we chose in Eq. (14), below we list all possible matrices A such that $BA = \mathbb{I}_n$ and its corresponding map $\alpha_{\mathbb{C}}$

$$B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}, \quad A = \begin{bmatrix} b & c \\ b & c \\ b & c \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \alpha_{\mathbb{C}}(t_1, t_2) = (t_1^{b+1} t_2^c, t_1^b t_2^{c+1}, t_1^b t_2^c),$$

where $b, c \in \mathbb{Z}$.

Like we discussed in Example 9, the moment polytope is the same as pictured in Fig. 1a. At each vertex v, there is a choice of $b, c \in \mathbb{Z}$ such that the above formula gives $\alpha_{\mathbb{C}}^{J_v}$, which gives rise to the standard $\mathbb{T}^2_{\mathbb{C}}$-action on the $\mathbb{C}^2$-chart at that vertex. Below we compare the $\mathbb{C}^2$-charts at the vertices indexed by $\{1, 2\}$ and $\{1, 3\}$, respectively.

The vertex $\{1, 2\}$ is at the lower left corner of the of the triangle, corresponding to the intersection of the two facets normal to columns 1 and 2 of B. The chart at this vertex is the same as the one presented in Example 13. We have

$$\begin{aligned}
&\varphi^{\{1,2\}} : V_{\{1,2\}} = \{[z_1 : z_2 : z_3] \in \mathbb{P}^2 \mid z_3 \neq 0\} \longrightarrow U_{\{1,2\}} = \{(z_1, z_2, 1) \in \mathbb{C}^3\}, \\
&\varphi^{\{1,2\}}([z_1 : z_2 : z_3]) = \left(y_1^{\{1,2\}}, y_2^{\{1,2\}}\right) = \left(\tfrac{z_1}{z_3}, \tfrac{z_2}{z_3}\right), \\
&\alpha_{\mathbb{C}}^{\{1,2\}} : \mathbb{T}^2_{\mathbb{C}} \xrightarrow{\cong} U_{\{1,2\}} \backslash \{z_j = 0\}_{j=1,2} = \mathbb{C}\left[y_1^{\{1,2\}}, y_2^{\{1,2\}}\right] \setminus \left\{y_k^{\{1,2\}} = 0\right\}_{k=1,2} \\
&\alpha_{\mathbb{C}}^{\{1,2\}}(t_1, t_2) = (t_1, t_2, 1) \Longrightarrow \left(t_1 = y_1^{\{1,2\}}, t_2 = y_2^{\{1,2\}}\right).
\end{aligned}$$

For vertex $\{1, 3\}$, which is at the top of the triangle, we have

$$\begin{aligned}
&\varphi^{\{1,3\}} : V_{\{1,3\}} = \{[z_1 : z_2 : z_3] \in \mathbb{P}^2 \mid z_2 \neq 0\} \longrightarrow U_{\{1,3\}} = \{(z_1, 1, z_3) \in \mathbb{C}^3\}, \\
&\varphi^{\{1,3\}}([z_1 : z_2 : z_3]) = \left(y_1^{\{1,3\}}, y_2^{\{1,3\}}\right) = \left(\tfrac{z_1}{z_2}, \tfrac{z_3}{z_2}\right), \\
&\alpha_{\mathbb{C}}^{\{1,3\}} : \mathbb{T}^2_{\mathbb{C}} \xrightarrow{\cong} U_{\{1,3\}} \backslash \{z_j = 0\}_{j=1,3} = \mathbb{C}\left[y_1^{\{1,3\}}, y_2^{\{1,3\}}\right] \setminus \left\{y_k^{\{1,3\}} = 0\right\}_{k=1,2}, \\
&\alpha_{\mathbb{C}}^{\{1,3\}}(t_1, t_2) = (t_1 t_2^{-1}, 1, t_2^{-1}) \Longrightarrow \left(t_1 t_2^{-1} = y_1^{\{1,3\}}, t_2^{-1} = y_2^{\{1,3\}}\right).
\end{aligned}$$

Using the identification between the complex toric coordinates t with the inhomogeneous coordinates y on the two charts above, we can see that the coordinate change between these two charts is given by

$$y_1^{\{1,2\}} = y_1^{\{1,3\}} \left(y_2^{\{1,3\}}\right)^{-1}, \quad y_2^{\{1,2\}} = \left(y_2^{\{1,3\}}\right)^{-1}$$

on the overlap

$$\varphi^{\{1,3\}}\left(V_{\{1,2\}} \cap V_{\{1,3\}}\right) = \{(z_1, 1, z_3) \mid z_3 \neq 0\} = \mathbb{C}\left[y_1^{\{1,3\}}, y_2^{\{1,3\}}\right] \backslash \{y_2^{\{1,3\}} = 0\}.$$

The above transition map is precisely $\alpha_{\mathbb{C}}^{\{1,2\}} \circ \left(\alpha_{\mathbb{C}}^{\{1,3\}}\right)^{-1}$, except that $\left(\alpha_{\mathbb{C}}^{\{1,3\}}\right)^{-1}$ is only defined on the $(\mathbb{C}^*)^2$, and the above overlap set is $\mathbb{C} \times \mathbb{C}^*$, which is slightly bigger. So the transition map is a slight extension of the map defined using the $\alpha_{\mathbb{C}}$'s.

In this paragraph, we illustrate the change of $\mathbb{T}^2_{\mathbb{C}}$-basis. Changing the basis of the range of B, which is $\mathfrak{t}^2 = \mathrm{Lie}(\mathbb{T}^2_{\mathbb{C}})$, is equivalent to left multiplying B by an element in $\mathrm{GL}(2,\mathbb{Z})$ with determinant ± 1. For example, taking $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathrm{GL}(2,\mathbb{Z})$, we get a new $B' = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & -2 \end{bmatrix}$. Then the corresponding moment polytope is a triangle that is slanted, with inward normals given by columns of B', as in Fig. 1b. □

3.4 Justification of Choices for K_M

To determine B^+ and Q^+, we start by considering a vertex v of the moment polytope Δ for M. As in Sect. 3.3, this corresponds to an index J_v of size n. After possible renumbering, we may assume $J_v = (1, \ldots, n)$. In particular, we have a $\mathbb{Z}$-basis

$$\mathcal{B} := \{v^1, \ldots, v^n\} \subset \mathbb{Z}^n$$

for $\mathbb{Z}^n$ consisting of inward normals to the first n facets. Therefore each of the remaining inward normals $v^{n+1}, \ldots, v^d$ can be expressed as a linear combination of vectors in the basis $\mathcal{B}$ via an $n \times (d-n)$ matrix $C = (c_{jk})_{1 \le j \le n, 1 \le k \le d-n}$, where the kth column expresses v^{n+k} as a vector with respect to $\mathcal{B}$, namely

$$v^{n+k} = \sum_{j=1}^{n} c_{jk} v^j. \tag{75}$$

Now we would like to answer the question, *what are the inward normals to the polytope Δ^+ for K_M?* These make up the columns of B^+. Recall the correspondence on M between vertices of Δ, multi-indices J, and coordinate charts U_J, from Sect. 3.3. Let $y_1, \ldots, y_n$ denote coordinates on one coordinate chart corresponding to $J = J_v$ centered around toric fixed point $q \in M \hookrightarrow K_M$ and $\tilde{y}_1, \ldots, \tilde{y}_n$ coordinates on another coordinate chart corresponding to choices with a tilde. In K_M, each chart has an additional coordinate y_q and $y_{\tilde{q}}$. First let D denote the $n \times n$ change of basis matrix between inward normals corresponding to each chart of M, namely between $\mathcal{B}$ and $\{\tilde{v}^1, \ldots, \tilde{v}^n\}$ (the inward normals to the facets at the intersection of the second vertex):

$$\tilde{v}^j = \sum_{k=1}^{n} d_{jk} v^k. \tag{76}$$

where now the jth row expresses $\tilde{v}^j$ as a linear combination of the vectors v^k. Note that $D = (d_{jk}) \in \mathrm{GL}(n, \mathbb{Z})$ has determinant 1 or -1 for the same reason as in Eq. (73). If it has determinant -1, interchange $\tilde{v}^1$ and $\tilde{v}^2$; this way we may assume $\det D = +1$. Under the identifications of Eq. (53), the transition functions turn the coefficients into exponents so that in coordinates

$$y_k = \prod_{j=1}^{n} \tilde{y}_j^{d_{jk}}, \qquad k = 1, \ldots, n. \tag{77}$$

Note that if D is the change of basis on vectors, then $(D^{-1})^T$ is the change of coordinates.

An element locally trivializing K_M may be written $y_q dy_1 \wedge \cdots \wedge dy_n$. On the overlap of two charts, we have two different sets of coordinates with which to express a point:

$$y_q dy_1 \wedge \cdots \wedge dy_n = y_{\tilde{q}} d\tilde{y}_1 \wedge \cdots \wedge d\tilde{y}_n. \tag{78}$$

Taking the log of Eq. (77) and differentiating, we obtain

$$\begin{aligned} \log y_k = \sum_{j=1}^{n} d_{jk} \log \tilde{y}_j \implies \frac{dy_k}{y_k} &= \sum_{j=1}^{n} d_{jk} \frac{d\tilde{y}_j}{\tilde{y}_j} \\ \implies \frac{dy_1}{y_1} \wedge \cdots \wedge \frac{dy_n}{y_n} &= \frac{d\tilde{y}_1}{\tilde{y}_1} \wedge \cdots \wedge \frac{d\tilde{y}_n}{\tilde{y}_n}, \end{aligned} \tag{79}$$

where the last equality holds because $\det D = +1$. We can now apply this to Eq. (78) to determine $y_{\tilde{q}}$.

$$\begin{aligned} y_q dy_1 \wedge \cdots \wedge dy_n = y_{\tilde{q}} d\tilde{y}_1 \wedge \cdots \wedge d\tilde{y}_n = y_{\tilde{q}} dy_1 \wedge \cdots \wedge dy_n \frac{\tilde{y}_1 \cdots \tilde{y}_n}{y_1 \cdots y_n} \\ \implies y_q = y_{\tilde{q}} \frac{\tilde{y}_1 \cdots \tilde{y}_n}{y_1 \cdots y_n} = y_{\tilde{q}} \prod_{j=1}^{n} \tilde{y}_j^{1 - \sum_{k=1}^{n} d_{jk}} \end{aligned} \tag{80}$$

by Eq. (77). We now have the information needed to write down B^+, which we prove below.

Lemma 1 *Given a matrix B for M of the following form:*

$$B = \left[\mathbb{I}_n \middle| C \right], \tag{81}$$

the following gives a choice of B^+ for K_M:

$$B^+_{(n+1)\times(d+1)} = \left[\begin{array}{c|c|c} \mathbb{I}_n & C & \mathbf{0} \\ \hline 0 & 1-\sum_{j=1}^n c_{j,1} \ \ldots \ 1-\sum_{j=1}^n c_{j,d-n} & 1 \end{array}\right]. \tag{82}$$

Corollary 2 *One choice of Q^+ satisfying $B^+Q^+ = 0$, for the B^+ in Lemma 1 for K_M, is*

$$Q^+_{(d+1)\times(d-n)} := \left[\begin{array}{c} -C \\ \hline \mathbb{I}_{d-n} \\ \hline \left(\sum_{j=1}^n c_{j,1}\right)-1 \ \ \ldots \ \ \left(\sum_{j=1}^n c_{j,d-n}\right)-1 \end{array}\right] \tag{83}$$

and one choice of A^+ satisfying $B^+A^+ = \mathbb{I}_{n+1}$ is

$$A^+_{(d+1)\times(n+1)} := \left[\begin{array}{c|c} \mathbb{I}_n & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array}\right]. \tag{84}$$

Remark 3 Other choices of Q^+ differ from this one via right multiplication by an element in $\mathrm{GL}(d-n, \mathbb{Z})$, and they all satisfy the property that the sum of the entries in each column must be 0, justifying the choice of Q^+ we presented in Eq. (24).

Remark 4 Note that the sum of the entries in a column of B^+ equals 1. Namely, the total weight of this $N_{\mathbb{C}}$ action is 1. (We will see this play a role later on in the mirror symmetry section in that the $N_{\mathbb{C}}$ action preserves $z_1 \ldots z_d p$, which is the *superpotential*. It also preserves the Calabi-Yau condition of K_M, in which case the above-mentioned n-form is global.)

Proof of Lemma 1 In coordinates $(y_1, \ldots, y_n, y_q)$ on K_M, the Hamiltonian $\mathbb{T}^{n+1}$-action restricts to the $\mathbb{T}^n$-action on M when $y_q = 0$. In particular, this tells us that for $i_0 : M \to K_M$, we have $i_0^* \mu_a^+ = \mu_a$ and $\Delta^+ \cap \{\xi_{n+1} = \kappa_{d+1}\} = \Delta$ (for κ as defined in Eq. (49)). Thus to answer our question regarding the inward normals of Δ^+, we have that Δ^+ has $d+1$ facets whose normals are denoted by $v^{j,+}$, for $j = 1, \ldots, d+1$. For $j = 1, \ldots, d$, the first n coordinates of $v^{j,+}$ are the same as that of v^j, i.e.

$$v^{j,+} = (v^j, v^{j,+}_{n+1}), \qquad 1 \le j \le d, \tag{85}$$

and below we will discuss how to determine the last coordinate, $v^{j,+}_{n+1}$ based on our changed of coordinates calculation above. The last facet, which corresponds to $j = d+1$, is in the $\xi_{n+1} = \kappa_{d+1}$ plane, so its primitive inward normal vector is $v^{d+1,+} := [0 \ldots 0 \ 1]^T$. This last facet intersects every vertex of Δ^+. (See also Remark 1 and Eq. (60) for discussions on relating M and K_M.) This can be visualized in the case of $K_{\mathbb{P}^1}$ in Fig. 2.

At a vertex v^+ of Δ^+ corresponding to v for Δ, we hence have a basis of inward normal vectors

$$\mathcal{B}^+ := \{v^{1,+}, \ldots, v^{n,+}, v^{d+1,+}\}. \tag{86}$$

With respect to this basis $\mathcal{B}^+$, by Eq. (75) and Eq. (85) we can write B^+ as

$$B^+ = \left[\begin{array}{c|ccc|c} \mathbb{I}_n & & C & & \mathbf{0} \\ \hline \mathbf{0} & v_{n+1}^{n+1,+} & \ldots & v_{n+1}^{d,+} & 1 \end{array}\right],$$

since there is no change of basis in the last vector $v^{d+1,+}$ according to Eq. (80). To fill in the last row, it remains to write each $v^{n+k,+}$, $k = 1, \ldots, d-n$, in terms of the basis vectors in $\mathcal{B}^+$. So for each k, we take $v^{n+k,+}$, let $\tilde{v}^{1,+} = v^{n+k,+}$, and extend it to a $\mathbb{Z}$-basis $\{\tilde{v}^{1,+}, \ldots, \tilde{v}^{n,+}, v^{d+1,+}\}$ of $\mathbb{Z}^{n+1}$ corresponding to a vertex $\tilde{v}^+$. Let $\tilde{v} = \{\tilde{v}^1, \ldots, \tilde{v}^n\}$ be the first n coordinates of $\{\tilde{v}^{1,+}, \ldots, \tilde{v}^{n,+}\}$, and we look at what Eq. (80) say for this choice of $\tilde{v}$. So, $\tilde{v}$ is related to v by Eq. (76). Furthermore, because of the choice that $\tilde{v}^{1,+} = v^{n+k,+}$, by comparing Eq. (76) and Eq. (75), we see that the first row $[d_{11}, \ldots, d_{1n}]$ of D is the same as the transpose of the k-th column $[c_{1k}, \ldots, c_{nk}]$ of C, i.e. $d_{1j} = c_{jk}$ for $j = 1, \ldots, n$. Hence we obtain

$$\begin{aligned} v^{n+k,+} = \tilde{v}^{1,+} &= \sum_{j=1}^{n} d_{1j} v^{j,+} + \left(1 - \sum_{j=1}^{n} d_{1j}\right) v^{d+1,+} \\ &= \sum_{j=1}^{n} c_{jk} v^{j,+} + \left(1 - \sum_{j=1}^{n} c_{jk}\right) v^{d+1,+}, \qquad k = 1, \ldots, d-n, \end{aligned} \tag{87}$$

where in the first line of the above calculation, the first term follows from Eq. (76) and the second term follows from the coordinate change of y_q in Eq. (80). This proves the lemma. □

Example 15 ($K_{\mathbb{P}^n}$) In Example 11, we presented one choice of B^+. Following the prescription given above in this section, we obtain another choice for B^+ below, which differs from the previous one by an invertible linear transformation over $\mathbb{Z}$. (These two different choices of B^+ correspond to two different sets of $\mathbb{T}^{n+1}$-basis, so their moment polytopes look different; see Fig. 2b for the case of $n = 1$.) One choice of A^+ such that $B^+A^+ = \mathbb{I}_{n+1}$ is given below

$$B^+ = \left[\begin{array}{ccc|c} & & & -1 \\ & \mathbb{I}_n & & \vdots \\ & & & -1 \\ \hline 0 & \ldots & 0 & n+1 \end{array}\right], \qquad A^+ := \left[\begin{array}{ccc|c} & & & 0 \\ & \mathbb{I}_n & & \vdots \\ & & & 0 \\ \hline 0 & \ldots & 0 & 0 \\ \hline 0 & \ldots & 0 & 1 \end{array}\right]. \tag{88}$$

□

3.5 Moment Map for $K_{\mathbb{P}^n}$ in Homogeneous Coordinates

Recall in Example 7, we obtained $K_{\mathbb{P}^n}$ as the quotient

$$K_{\mathbb{P}^n} = U_a^+/\mathbb{C}^* = (\mu_N^+)^{-1}(a)/U(1),$$

where $U_a^+ = (\mathbb{C}^{n+1} - \{0\}) \times \mathbb{C}$ and

$$(\mu_N^+)^{-1}(a) = \{(z_1, \ldots, z_{n+1}, p) \in \mathbb{C}^{n+2} \mid |z_1|^2 + \cdots + |z_{n+1}|^2 - (n+1)|p|^2 = 2a\},$$

and note that we use $a > 0$ as this is the chamber that gives $K_{\mathbb{P}^n}$. For any $(z, p) \in U_a^+$, there is a unique $\lambda_a(z, p) \in \mathbb{R}_{>0}$ such that $\lambda_a(z, p) \cdot (z, p) \in (\mu_N^+)^{-1}(a)$ under the action $\lambda_a \cdot (z, p) = (\lambda_a z, \lambda_a^{-n-1} p)$ defined by Eq. (28). This defines the following deformation retraction (same as Eq. (33))

$$R_a^+ : (\mathbb{C}^{n+1} - \{0\}) \times \mathbb{C} \to (\mu_N^+)^{-1}(a), \quad R_a^+(z, p) = (\lambda_a(z, p)z, \lambda_a(z, p)^{-n-1} p).$$

The $\mathbb{T}^{n+1}$-action on $K_{\mathbb{P}^n}$ and Its Moment Map Corresponding to the A^+ in Example 15, we have the map $\alpha_{\mathbb{C}}^+ : \mathbb{T}_{\mathbb{C}}^{n+1} \to \mathbb{T}_{\mathbb{C}}^{n+2}$ given by

$$\alpha_{\mathbb{C}}^+(t_1, \ldots, t_{n+1}) = (t_1, \ldots t_n, 1, t_{n+1}),$$

and we get a $\mathbb{T}_{\mathbb{C}}^{n+1}$-action on $\mathbb{C}^{n+2}$ via $\alpha_{\mathbb{C}}^+(\mathbb{T}_{\mathbb{C}}^{n+1})$ by

$$(t_1, \ldots, t_{n+1}) \cdot (z_1, \ldots, z_{n+1}, p) = (t_1 z_1, \ldots, t_n z_n, z_{n+1}, t_{n+1} p).$$

As usual, this restricts to a Hamiltonian $\mathbb{T}^{n+1}$-action on $\mathbb{C}^{n+2}$ in coordinates z, p with the standard symplectic form $\frac{i}{2}\left(\sum_{k=1}^{n+1} dz_k \wedge d\bar{z}_k + dp \wedge d\bar{p}\right)$ and moment map (up to a constant)

$$\mu = (\alpha^+)^* \circ \mu_{\mathbb{T}^{n+1}} : \mathbb{C}^{n+2} \to \mathbb{R}^{n+1}, \quad \mu(z_1, \ldots, z_{n+1}, p) = \frac{1}{2}(|z_1|^2, \ldots, |z_n|^2, |p|^2).$$

Because this $\mathbb{T}_{\mathbb{C}}^{n+1}$-action on U_a^+ commutes with the $\mathbb{C}^*$-action on U_a^+, we get that $\mathbb{T}_{\mathbb{C}}^{n+1}$ acts on the quotient $K_{\mathbb{P}^n} = U_a^+/\mathbb{C}^*$ by

$$(t_1, \ldots, t_n, t_{n+1}) \cdot [z_1 : \ldots : z_{n+1} : p] = [t_1 z_1 : \ldots : t_n z_n : z_{n+1} : t_{n+1} p]. \tag{89}$$

This restricts to a Hamiltonian $\mathbb{T}^{n+1}$-action on $(K_{\mathbb{P}^n}, \omega_a)$. Using Eq. (44), we find the moment map $\mu_a : K_{\mathbb{P}^n} \to \mathbb{R}^{n+1}$ is given by

$$\begin{aligned} \mu_a([z_1, \ldots, z_{n+1}, p]) &= \mu \circ R_a^+(z_1, \ldots, z_{n+1}, p) \\ &= \tfrac{1}{2}(\lambda_a(z, p)^2|z_1|^2, \ldots, \lambda_a(z, p)^2|z_n|^2, \lambda_a(z, p)^{-2(n+1)}|p|^2). \end{aligned} \tag{90}$$

Computing $\lambda_a(z,p)^2$ Now we try to compute $\lambda_a^2(z,p)$ in order to obtain a more explicit formula for μ_a in terms of the homogeneous coordinates (z,p). Let $|z|^2 = \sum_{k=1}^{n+1} |z_k|^2$. By setting $\mu_N^+(z_1,\ldots,z_{n+1},p)$ equal to a, we get that $x := \lambda_a(z,p)^2|z|^2$ satisfies

$$x^{n+2} - 2ax^{n+1} - (n+1)|z|^{2(n+1)}|p|^2 = 0.$$

So

$$|z|^{2(n+1)}|p|^2 = \frac{1}{n+1}(x^{n+2} - 2ax^{n+1}) =: f(x).$$

$$f'(x) = \frac{n+2}{n+1}x^{n+1} - 2ax^n$$

$f'(x) > 0$ for $x > \frac{2(n+1)a}{n+2}$. So there is a smooth function

$$g : \left(-\frac{2a}{(n+1)(n+2)}\left(\frac{2a(n+1)}{n+2}\right)^{n+1}, \infty\right) \longrightarrow \left(\frac{2(n+1)a}{n+2}, \infty\right)$$

which is the inverse of

$$f : \left(\frac{2(n+1)a}{n+2}, \infty\right) \longrightarrow \left(-\frac{2a}{(n+1)(n+2)}\left(\frac{2a(n+1)}{n+2}\right)^{n+1}, \infty\right).$$

Furthermore, $g'(t) > 0$, and

$$g(0) = 2a, \quad \lim_{t\to+\infty} g(t) = +\infty.$$

Because g is the inverse of f, we get that $x = g(|z|^{2(n+1)}|p|^2)$, and by definition

$$\lambda_a(z,p)^2 = \frac{x}{|z|^2} = \frac{g(|z|^{2(n+1)}|p|^2)}{|z|^2}.$$

Using the above formula for $\lambda_a(z,p)^2$ and continuing Eq. (90) for the moment map μ_a, we get

$$\begin{aligned}\mu_a([z_1,\ldots,z_{n+1},p]) &= \mu \circ R_a^+(z_1,\ldots,z_{n+1},p)\\ &= \left(\frac{g(|z|^{2(n+1)}|p|^2)}{2}\frac{|z_1|^2}{|z|^2},\ldots,\frac{g(|z|^{2(n+1)}|p|^2)}{2}\frac{|z_n|^2}{|z|^2},\frac{1}{2(n+1)}(g(|z|^{2(n+1)}|p|^2) - 2a)\right).\end{aligned} \tag{91}$$

Recall that $g(0) = 2a$, so

$$\mu_a([z_1, \ldots, z_{n+1}, 0]) = \frac{a}{|z_1|^2 + \cdots + |z_{n+1}|^2}(|z_1|^2, \ldots, |z_n|^2, 0).$$

So when $p = 0$ we obtain the same moment map as above in Example 9 for $\mathbb{P}^n$.

Explicit Formula for When $n = 1$ We next derive an explicit expression of the inverse function g of f when $n = 1$. In this case, $x = \lambda(z, p)^2|z|^2 > 0$ satisfies

$$x^3 - 2ax^2 - 2|z|^4|p|^2 = 0. \tag{92}$$

If $p \neq 0$, let $y = x^{-1}$. Then y is the unique real root of the following depressed cubic equation.

$$y^3 + \frac{a}{|z|^4|p|^2}y - \frac{1}{2|z|^4|p|^2} = 0.$$

By Cardano's formula,

$$\begin{aligned} y &= \sqrt[3]{\frac{1}{4|z|^4|p|^2} + \sqrt{\frac{1}{16|z|^8|p|^4} + \frac{a^3}{27|z|^{12}|p|^6}}} + \sqrt[3]{\frac{1}{4|z|^4|p|^2} - \sqrt{\frac{1}{16|z|^8|p|^4} + \frac{a^3}{27|z|^{12}|p|^6}}} \\ &= \frac{\sqrt{a}}{\sqrt{3}|z|^2|p|}\left(\sqrt[3]{\sqrt{1 + \frac{27|z|^4|p|^2}{16a^3}} + \frac{3\sqrt{3}|z|^2|p|}{4a\sqrt{a}}} - \sqrt[3]{\sqrt{1 + \frac{27|z|^4|p|^2}{16a^3}} - \frac{3\sqrt{3}|z|^2|p|}{4a\sqrt{a}}}\right) \\ &= \frac{1}{2a} \times \frac{3}{\left(\sqrt{1+\rho_a(z,p)^2}+\rho_a(z,p)\right)^{\frac{2}{3}} + \left(\sqrt{1+\rho_a(z,p)^2}-\rho_a(z,p)\right)^{\frac{2}{3}} + 1} \end{aligned}$$

where

$$\rho_a(z, p) = \frac{3\sqrt{3}|z|^2|p|}{4a\sqrt{a}}$$

is invariant under the $\mathbb{C}^*$-action on $(\mathbb{C}^2 - \{0\}) \times \mathbb{C}$, and $\rho_a(z, 0) = 0$.

Using the relations that $x = g(|z|^{2(n+1)}|p|^2)$ (here $n = 1$) and $x = y^{-1}$, we get that

$$\begin{aligned} &g(|z|^4|p|^2) \\ &= \frac{2a}{3}\left(\left(\sqrt{1 + \rho_a(z, p)^2} + \rho_a(z, p)\right)^{\frac{2}{3}} + \left(\sqrt{1 + \rho_a(z, p)^2} - \rho_a(z, p)\right)^{\frac{2}{3}} + 1\right) \\ &= 2a + \frac{2a}{3}\left(\sqrt[3]{\sqrt{1 + \frac{27|z|^4|p|^2}{16a^3}} + \frac{3\sqrt{3}|z|^2|p|}{4a\sqrt{a}}} - \sqrt[3]{\sqrt{1 + \frac{27|z|^4|p|^2}{16a^3}} - \frac{3\sqrt{3}|z|^2|p|}{4a\sqrt{a}}}\right)^2. \end{aligned}$$

We can then substitute this into Eq. (91) to get

$$\mu_a([z_1, z_2, p]) = \left(\frac{g(|z|^4|p|^2)}{2}\frac{|z_1|^2}{|z_1|^2+|z_2|^2}, \frac{g(|z|^4|p|^2)-2a}{4}\right) =$$

$$= \left(\frac{a|z_1|^2}{|z|^2} \times \frac{\left(\sqrt{1+\rho_a(z,p)^2}+\rho_a(z,p)\right)^{\frac{2}{3}} + \left(\sqrt{1+\rho_a(z,p)^2}-\rho_a(z,p)\right)^{\frac{2}{3}} + 1}{3},\right.$$

$$\left.\frac{a}{2} \times \frac{\left(\sqrt{1+\rho_a(z,p)^2}+\rho_a(z,p)\right)^{\frac{2}{3}} + \left(\sqrt{1+\rho_a(z,p)^2}-\rho_a(z,p)\right)^{\frac{2}{3}} - 2}{3}\right).$$

The moment polytope, which is the image of this moment map μ_a is shown in Fig. 2b.

So far in this section, the $\mathbb{T}^2$-action on $K_{\mathbb{P}^1}$ is given by Eq. (89), which arose from the choice of B^+ and A^+ given in Example 15. If we change the basis of $\mathbb{T}^2$ and use the B^+ and A^+ given in Example 11, we get the $\mathbb{T}^2$-action on $K_{\mathbb{P}^1}$ given by

$$(t_1, t_2) \cdot [z_1 : z_2 : p] = [t_1 z_1 : z_2 : t_1^{-1} t_2 p],$$

then the moment map becomes

$$\mu'_a([z_1, z_2, p]) = \left(\frac{g(|z|^4|p|^2)}{2}\frac{|z_1|^2}{|z_1|^2+|z_2|^2} - \frac{g(|z|^4|p|^2)-2a}{4}, \frac{g(|z|^4|p|^2)-2a}{4}\right) \tag{93}$$

and its image is shown in Fig. 2a. In particular, when $p = 0$,

$$\mu'_a([z_1, z_2, 0]) = \left(\frac{a|z_1|^2}{|z_1|^2+|z_2|^2}, 0\right).$$

Note that the first component is the moment map for $\mathbb{P}^1$, and indeed its image is the line segment connecting $(0, 0)$ and $(a, 0)$ in Fig. 2a, b.

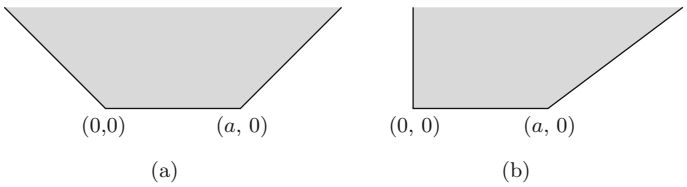


Fig. 2 Two moment polytopes for $K_{\mathbb{P}^1}$. (**a**) Image of the moment map $\mu'_a : K_{\mathbb{P}^1} \to \mathbb{R}^2$. (**b**) Image of the moment map $\mu_a : K_{\mathbb{P}^1} \to \mathbb{R}^2$

4 Kähler Potential

There is a natural Kähler potential that toric varieties admit from toric geometry, as follows. With the data of a moment polytope, they come equipped with an ample line bundle defined by a sum of the toric divisors weighted according to the moment polytope; hence they admit a Kähler potential induced from sections of that line bundle, see [10, Proposition 4.3.3] and [21, Example 4.1.2(i)] for details. The Kähler potential in this case is a function of the complex affine coordinates. However, many calculations become simpler in the moment map coordinates. We saw above the existence of a symplectic form ω_a coming from symplectic reduction. We will see in this section that the Legendre transform of the Kähler potential one obtains from symplectic reduction is computable. Although the two symplectic forms are different in general, they can have similar properties such as being in the same Kähler class.

More specifically, the focus of this chapter is to find a Kähler potential F over $\mathring{\Delta}$ for ω_a in terms of the affine toric coordinates, as well as its Legendre transform G in terms of the action-angle coordinates, specifically the moment map coordinates. First we recall some notation. The diagram below is a summary of the identifications given by Eqs. (50), (52), and (53), and the notations introduced in Definition 2,

$$\begin{array}{ccccc} \mathfrak{t}^n_{\mathbb{C}} & \xrightarrow{\exp} & \mathbb{T}^n_{\mathbb{C}} & \cong \mu_a^{-1}(\mathring{\Delta}) \cong & \mathring{\Delta} \times \mathbb{T}^n \\ \cup & & \cup & & \cup \\ u = (u_j = x_j + i\theta_j)_{j=1}^n & \mapsto & (e^{u_1}, \ldots, e^{u_n}) & & (\xi, e^{i\theta}) = (\xi_j, e^{i\theta_j})_{j=1}^n \\ & & \| & & \\ & & (t_1, \ldots t_n) & \longmapsto & \mu_a(t) = \xi \in \mathring{\Delta} \end{array}. \tag{94}$$

Note that since $\mathbb{T}^n$ embeds in $\mathbb{T}^n_{\mathbb{C}}$ as $U(1)^n$, $\theta = (\theta_1, \ldots, \theta_n)$ is a coordinate on the Lie algebra of $\mathbb{T}^n_{\mathbb{C}}$. The above diagram identifies $\mu_a^{-1}(\mathring{\Delta})$ with $\mathbb{T}^n_{\mathbb{C}}$ and $\mathring{\Delta} \times \mathbb{T}^n$, and hence endows $\mu_a^{-1}(\mathring{\Delta})$ with the complex toric coordinates t and the action-angle coordinates (ξ, θ), respectively.

Kähler Potential F By Guillemin [16, Theorem 4.3], there exists a $\mathbb{T}^n$-invariant Kähler potential F on $(\mathbb{C}^*)^n \subset M$ such that $\omega_a|_{(\mathbb{C}^*)^n} = 2i\partial\overline{\partial}F$. Contracting the symplectic form with $\partial/\partial\theta$, using that F is not a function of θ since it is $\mathbb{T}^n$-invariant (that is, a function $\mathbb{R}^n \to \mathbb{R}$ on the norms of the affine toric coordinates of $(\mathbb{C}^*)^n$), plugging into the left hand side of $\iota_{\partial/\partial\theta_j}\omega_a = -d\mu_{a,j}$, and integrating we see that $F : \mathbb{R}^n \to \mathbb{R}$ satisfies

$$\mu_a(t_1, \ldots, t_n) = \mu_a(e^{x_1+i\theta_1}, \ldots, e^{x_n+i\theta_n}) = \left(\frac{\partial F}{\partial x_1}(x), \ldots, \frac{\partial F}{\partial x_n}(x)\right). \tag{95}$$

Therefore by Guillemin [16, Theorem 3.3] and the sentence following it, the map $x \mapsto \left(\frac{\partial F}{\partial x_1}(x), \ldots, \frac{\partial F}{\partial x_n}(x)\right)$ is a diffeomorphism from $\mathbb{R}^n$ to the interior $\mathring{\Delta}$ of the

moment polytope Δ. In particular, let $\xi_j = \frac{\partial F}{\partial x_j}$. Then

$$\frac{\partial \xi_j}{\partial x_k} = \frac{\partial^2 F}{\partial x_j \partial x_k}. \tag{96}$$

Legendre Transform G In the other direction, there is a function $G : \mathring{\Delta} \to \mathbb{R}$ such that the inverse map of the diffeomorphism

$$\mathbb{R}^n \longrightarrow \mathring{\Delta}, \quad (x_1, \ldots, x_n) \mapsto \left(\frac{\partial F}{\partial x_1}(x), \ldots, \frac{\partial F}{\partial x_n}(x) \right) \tag{97}$$

is

$$\mathring{\Delta} \longrightarrow \mathbb{R}^n, \quad (\xi_1, \ldots, \xi_n) \mapsto \left(\frac{\partial G}{\partial \xi_1}(\xi), \ldots, \frac{\partial G}{\partial \xi_n}(\xi) \right) = (x_1, \ldots, x_n). \tag{98}$$

Specifically, G is the *Legendre transform* of F, namely (up to a constant)

$$F(x) + G(\xi) = \langle x, \xi \rangle \tag{99}$$

so that differentiating Eq. (99) with respect to ξ we see that $\frac{\partial G}{\partial \xi_i} = x_i$. Furthermore

$$\frac{\partial x_j}{\partial \xi_k} = \frac{\partial^2 G}{\partial \xi_j \partial \xi_k}. \tag{100}$$

From the relations in Eqs. (96) and (100), we see that the Hessian of G is the inverse of the Hessian of F. Also from Eqs. (96) and (100), we see that the complex structure in the (x_j, θ_j) coordinates is the standard $\left(\begin{array}{c|c} 0 & -\mathbb{I}_n \\ \hline \mathbb{I}_n & 0 \end{array}\right)$, hence the complex structure in the (ξ_j, θ_j) coordinates is $\left(\begin{array}{c|c} 0 & -[\partial_j \partial_k F] \\ \hline [\partial_j \partial_k G] & 0 \end{array}\right)$. We now discuss the Kähler form and metric.

Recall in Eq. (50), $\mathring{\Delta} \times \mathbb{T}^n$ with the symplectic form $\sum_{j=1}^n d\xi_j \wedge d\theta_j$ is symplectomorphic to $\mu_a^{-1}(\mathring{\Delta})$ with the symplectic form ω_a. We explain why the symplectic form is standard in the action-angle coordinates.

Lemma 2 *The symplectic form for M can be written (over the interior of the moment polytope) in the action-angle coordinates as*

$$\omega_a = \sum_{j=1}^{n} d\mu_{a,j} \wedge d\theta_j.$$

Proof We impose conditions on the coefficients of terms in ω_a. There are three conditions: $\iota_{\partial/\partial\theta_j}\omega_a = -d\mu_{a,j}$, fibers of the moment map are Lagrangian with tangent space spanned by the $\partial/\partial\theta_j$, and ω_a is Kähler so compatible with the almost complex structure J induced from multiplication by i on the affine coordinates. The first condition is the definition of the moment map, where $\partial/\partial\theta_j$ is the infinitesimal action of t_j on M defined by taking the derivative at the identity as in Eqs. (41) and (42), see also [39, Definition 22.1]. That is, $\theta_{j=1,\dots n}$ are the imaginary parts of the complex coordinates on the Lie algebra of $\mathbb{T}^n_{\mathbb{C}}$ and μ, θ are the action-angle coordinates. In particular, $\partial/\partial\theta_j$ corresponds to rotating the jth affine toric coordinate y_j by t_j, as described in the paragraph above Eq. (72).

As discussed in the paragraph following Corollary 1, the second condition is true by the Arnold-Liouville theorem which is stated, for example, in [39, Theorem 18.12]. In more detail, $\mathbb{T}^n$ is abelian so has trivial Lie bracket and

$$\omega_a\left(\frac{\partial}{\partial\theta_i}, \frac{\partial}{\partial\theta_j}\right) = \left[\frac{\partial}{\partial\theta_i}, \frac{\partial}{\partial\theta_j}\right] = 0 = \{\mu_{a,i}, \mu_{a,j}\}$$

where $\{,\}$ denotes the Poisson bracket. But the tangent space of a μ_a-fiber is precisely the kernel of $d\mu_a$, so by the first condition:

$$T\mu_a^{-1}(c) = \ker(d\mu_a) = \bigcap_{j=1}^{n} \ker(\iota_{\partial/\partial\theta_j}\omega_a) \supseteq span\{\partial/\partial\theta_j\}_{j=1}^n.$$

Since a fiber has dimension n, the last containment is equality. The fibers of the moment map are then $\mathbb{T}^n$-orbits of the $\mathbb{T}^n$-action $(e^{i\theta_1}, \dots, e^{i\theta_n}) \cdot (y_1, \dots, y_n) = (e^{i\theta_1} y_1, \dots, e^{i\theta_n} y_n)$, where $(t_1, \dots, t_n) = (e^{i\theta_1}, \dots, e^{i\theta_n})$, therefore as ω_a is toric invariant, ω_a restricted to a fiber is 0. That is, the fibers are Lagrangian.

For the third condition, we saw that multiplication by i on $x \in \mathfrak{t}^n_{\mathbb{C}}$ induces, under the Jacobian transformation between ξ and x, $J = \left(\begin{array}{c|c} 0 & -[\partial_j\partial_k F] \\ \hline [\partial_j\partial_k G] & 0 \end{array}\right)$ in the (ξ_j, θ_j) coordinates, where $[\partial_j\partial_k F] = [\partial_j\partial_k G]^{-1}$ hence $J^2 = -\mathbb{I}_{2n}$.

Now we prove the statement of the lemma. Let $\omega_a = \sum_{j,k=1}^{n} a_{jk} d\mu_{a,j} \wedge d\theta_k + \sum_{j\le k,j,k=1}^{n} b_{jk} d\theta_j \wedge d\theta_k + c_{jk} d\mu_{a,j} \wedge d\mu_{a,k}$. Then $\omega_a|_{span\{\partial/\partial\theta_j\}_{j=1}^n} \equiv 0 \implies b_{jk} = 0$ and by the definition of the moment map, $\iota_{\partial/\partial\theta_k}\omega_a = \sum_{j=1}^{n} -a_{jk} d\mu_{a,j} = -d\mu_{a,k} \implies a_{kk} = 1, a_{jk} = 0 \quad \forall j \neq k$. Wrapping up, compatibility with J means that $J^*\omega_a = \omega_a$ thus

$$\sum_{j=1}^{n} d\mu_{a,j} \wedge d\theta_j + \sum_{\substack{j,k=1 \\ j \le k}}^{n} c'_{jk} d\theta_j \wedge d\theta_k = \sum_{j=1}^{n} d\mu_{a,j} \wedge d\theta_j + \sum_{\substack{j,k=1 \\ j \le k}}^{n} c_{jk} d\mu_{a,j} \wedge d\mu_{a,k}$$

$$\therefore c_{jk} = 0 \implies \omega_a = \sum_{k=1}^{n} d\mu_{a,k} \wedge d\theta_k.$$

where c'_{jk} is a function of $[\partial_j \partial_k G]$. □

Below we will use ω_a to interchangeably denote the symplectic forms on $\mathbb{T}^n_{\mathbb{C}}$ and $\mathring{\Delta} \times \mathbb{T}^n$, which we view as ω_a written in two different set of coordinates. For K_M the same holds true replacing n with $n+1$.

Corollary 3

$$\omega_a = \sum_{j=1}^{n} d\xi_j \wedge d\theta_j = \sum_{j,k=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_k} dx_j \wedge d\theta_k = \frac{i}{2} \sum_{j,k=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_k} du_j \wedge d\bar{u}_k. \tag{101}$$

In the above, we have written ω_a both in terms of the complex coordinates $u_j = x_j + i\theta_j$ via the identification of $\mu^{-1}(\mathring{\Delta}) \cong \mathbb{T}^n_{\mathbb{C}}$, and in terms of the action-angle coordinates (ξ_j, θ_j) via the identification of $\mu^{-1}(\mathring{\Delta}) \cong \mathring{\Delta} \times \mathbb{T}^n$.

Corollary 4 *The Riemannian metric compatible with ω_a and the complex structure J is*

$$g_a = \sum_{j,k=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_k} (dx_j dx_k + d\theta_j d\theta_k) = \sum_{j,k=1}^{n} \frac{\partial^2 G}{\partial \xi_j \partial \xi_k} d\xi_j d\xi_k + \sum_{j,k=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_k} d\theta_j d\theta_k, \tag{102}$$

which we've also written both in terms of the complex coordinates and the action-angle coordinates.

To recap, let ω be a $\mathbb{T}^n$-invariant Kähler form on M. Then the action of $\mathbb{T}^n$ on (M, ω) is Hamiltonian with a moment map $\mu : M \to \mathbb{R}^n$ which is unique up to addition of a constant vector in the target $\mathbb{R}^n$. The image Δ of μ is a convex polytope known as the moment polytope, and the moment map determines smooth functions $F : \mathbb{R}^n \to \mathbb{R}$ and $G : \mathring{\Delta} \to \mathbb{R}$ up to adding constants, as described above. Let g be the Riemannian metric on M determined by the symplectic structure ω and the complex structure on M. Then g is given by the right hand side of (102). When ω is the symplectic structure coming from symplectic reduction, $G(\xi)$ is given explicitly in [16]. Below we summarize a way of deriving Guillemin's formula for $G(\xi)$ due to Calderbank-David-Gauduchon [8].

Lemma 3 *The Legendre transform $G(\xi)$ of the potential function $F(x)$ is*

$$G(\xi) = \frac{1}{2} \sum_{j=1}^{d} L_j(\xi) \log L_j(\xi).$$

Proof The standard flat metric on $\mathbb{C}^d$ written in polar coordinates, $z_j = r_j e^{i\varphi_j}$, is

$$g = \sum_{j=1}^{d} dz_j d\bar{z}_j = \sum_{j=1}^{d} dr_j^2 + r_j^2 d\varphi_j^2 = \sum_{j=1}^{d} \frac{d\mu_j^2}{2\mu_j} + 2\mu_j d\varphi_j^2 = \tilde{g} + \sum_{j=1}^{d} 2\mu_j d\varphi_j^2, \tag{103}$$

where $\mu_j = \frac{1}{2} r_j^2 = \frac{1}{2}|z_j|^2$ is the jth component of the moment map $\mu_{\mathbb{T}^d}$ given by Eq. (17). This metric is compatible with the standard symplectic form ω of Eq. (16) and the standard complex structure on $\mathbb{C}^d$. The first term in the sum above in Eq. (103) in polar coordinates, denoted $\tilde{g}$, is the metric on the image of the moment map, and the 2nd term is the metric on the torus fiber. Note that $\tilde{g}$ can be written in terms of the Hessian of a potential function $\widetilde{G}$ as follows

$$\tilde{g} = \sum_{j=1}^{d} \frac{d\mu_j^2}{2\mu_j} = \sum_{j,k=1}^{d} \frac{\partial^2 \widetilde{G}}{\partial \mu_j \partial \mu_k} d\mu_j d\mu_k, \text{ where } \widetilde{G} = \sum_{j=1}^{d} \mu_j \log \mu_j.$$

Recall that the metric g_a in Eq. (102) is the canonical metric defined by symplectic reduction from $\mathbb{C}^d$ with the flat metric g. The first term of g_a on the right-most expression of Eq. (102), which is

$$\tilde{g}_a = \sum_{j,k=1}^{n} \frac{\partial^2 G}{\partial \xi_j \partial \xi_k} d\xi_j d\xi_k,$$

is a Riemannian metric on $\mathring{\Delta}$ (the unique Riemannian metric such that $\mu_a : (\mathbb{T}^n_{\mathbb{C}}, g_a) \to (\mathring{\Delta}, \tilde{g}_a)$ is a Riemannian submersion). By construction and because $L(\xi)$ defined in Eq. (45) satisfies $L(\xi)_j = \frac{1}{2}|z_j|^2 = \mu_j$, we have

$$\tilde{g}_a = L^* \tilde{g} = \sum_{j=1}^{d} \frac{dL_j^2}{2L_j}, \quad G = L^* \widetilde{G} = \frac{1}{2} \sum_{j=1}^{d} L_j \log L_j. \tag{104}$$

□

Example 16 ($\mathbb{P}^n$) Recall $\alpha_{\mathbb{C}}$ from Example 9 and μ_a from Eq. (56), leading to the conclusion at the end of Example 13 that

$$(\xi_1, \ldots, \xi_n) = \frac{a(e^{2x_1}, \ldots, e^{2x_n})}{e^{2x_1} + \cdots + e^{2x_n} + 1}$$

and

$$(x_1, \ldots, x_n) = \frac{1}{2}\Big(\log \xi_1 - \log(a - \xi_1 - \cdots - \xi_n), \ldots, \log \xi_n - \log(a - \xi_1 - \cdots - \xi_n) \Big).$$

Thus, integrating, we find by Eqs. (95) and (98):

$$F(x_1, \ldots, x_n) = \frac{a}{2} \log(e^{2x_1} + \cdots + e^{2x_n} + 1) = \frac{a}{2} \log(|t_1|^2 + \cdots + |t_n|^2 + 1),$$

$$G(\xi_1, \ldots, \xi_n) = \frac{1}{2}\Big(\xi_1 \log \xi_1 + \cdots + \xi_n \log \xi_n + (a - \xi_1 - \cdots - \xi_n) \log(a - \xi_1 - \cdots - \xi_n)\Big).$$

□

These calculations apply for K_M as well, replacing n with $n + 1$ everywhere.

Note that if ω is the symplectic structure coming from symplectic reduction, and if we know the moment polytope Δ to begin with, then according to Equations (104) and (45), we can readily write down the explicit formulas $G(\xi)$ from the combinatorics of the moment polytope Δ. Below we describe two perspectives for writing down $F(x)$. Both are less straightforward for the exact same reason, if not impossible, compared to finding $G(\xi)$.

1. If we could compute the moment map, μ_a, then we could find $F(x)$ via Eq. (95). However, as we saw in Sect. 3.5 (also mentioned in Example 10 for $\mathbb{P}^2(3)$), the moment map can be extremely complicated. In Example 15, our ability to explicitly compute the moment map in the $n = 1$ case boils down to being able to solve the cubic polynomial (Eq. (92)). For $n > 2$, it's not possible to compute explicitly, at least not using the method of Sect. 3.5.
2. Once we know $G(\xi)$, we can find $F(x)$ via Legendre transform, i.e. Eqs. (98) and (99). However, in order to write down $F(x)$ using Eq. (99), we need to write ξ in terms of x, and $\xi(x)$ is exactly the moment map, which is complicated as we just discussed. One might also think to use Eq. (98), which is a system of n nonlinear equations in $(x_1, \ldots, x_n)$ and $(\xi_1, \ldots, \xi_n)$. Similarly, it's complicated, if not impossible, to use that to find $\xi(x)$.

Remark 5 (Comparing with the Kähler Structure Defined by an Ample Line Bundle) The toric manifolds from Delzant's construction are projective, so there is another natural Kähler form obtained from using sections of an ample line bundle. This remark addresses the following question: *is the Guillemin construction of a symplectic form from symplectic reduction the same as the Kähler structure from toric geometry given an ample line bundle?* They are different in general as noted in Remark 5, however they are in the same Kähler class by Guillemin [16, Equation (1.7)]. We first consider the 1-dimensional case. The ample line bundles on $\mathbb{P}^1$ are $\mathcal{O}_{\mathbb{P}^1}(k)$, where k is a positive integer, and the space of sections of $\mathcal{O}_{\mathbb{P}^1}(k)$ can be identified with the space of homogeneous polynomials in two variables z_1, z_2 of degree k:

$$H^0(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(k)) = \bigoplus_{m=0}^{k} \mathbb{C} z_1^m z_2^{k-m}.$$

Let $\mathbb{C}^*$ act on z_1 and z_2 with weights 1 and 0 respectively. The moment map in Section 4.2 of Fulton [12] is given by

$$\hat{\mu}_k([z_1 : z_2]) = \frac{\sum_{m=0}^{k} m|z_1^m z_2^{k-m}|^2}{\sum_{m=0}^{k} |z_1^m z_2^{k-m}|^2}.$$

The image of $\hat{\mu}_k$ is the closed interval $[0, k] \subset \mathbb{R}$.

When $[z_1 : z_2]$ are both non-zero, define $x = \log|z_1/z_2|$. Then letting $a = k$ in μ_a from Example 9 we get,

$$\mu_k([z_1 : z_2]) = \frac{dF_k}{dx}(x), \quad \hat{\mu}_k([z_1 : z_2]) = \frac{d\hat{F}_k}{dx}(x),$$

where the hat denotes the potential obtained from sections of the ample line bundle $\mathcal{O}(k)$ and without the hat means we are using symplectic reduction on $S^3(\sqrt{2k})/S^1$:

$$F_k(x) = \frac{k}{2}\log(1 + e^{2x}), \quad \hat{F}_k(x) = \frac{1}{2}\log\left(\sum_{m=0}^{k} e^{2mx}\right).$$

In particular, $F_1(x) = \hat{F}_1(x)$ however they are different for larger k. Here the Kähler potential F_k for the symplectic form ω_k defining the moment map $\mu_k : \mathbb{P}^1 \to \mathbb{R}$ is such that

$$\omega_k = d\mu_k \wedge d\theta = \frac{d^2F_k}{dx^2} dx \wedge d\theta$$

and similarly for the case without the hat. Therefore we see that $\hat{\omega}_1 = \omega_1$ since the Kähler potentials are equal. When $k > 1$ is an integer, $\hat{\omega}_k \neq \omega_k$ but they represent the same class in $H^2(\mathbb{P}^1, \mathbb{R}) \cong \mathbb{R}$. Also note that here k is an integer, so the images of $\hat{\mu}_k$, which is $[0, k]$, does not gives all possible $[0, a]$, where $a \in \mathbb{R}$, that μ_a does.

More generally for higher dimensional projective spaces, let $\hat{\omega}_k$ be the symplectic form on $\mathbb{P}^n$ determined by the ample line bundle $\mathcal{O}_{\mathbb{P}^n}(k)$, where $k > 0$. Then $\hat{\omega}_k$ is the pullback of ω_1 on $\mathbb{P}^{\binom{k+n}{k}}$ under the degree k embedding $\mathbb{P}^n \hookrightarrow \mathbb{P}^{\binom{k+n}{k}}$. In particular, $\hat{\omega}_1 = \omega_1$. When $k > 1$, $\hat{\omega}_k \neq \omega_k$ but they represent the same class in $H^{1,1}(\mathbb{P}^n; \mathbb{R}) \cong \mathbb{R}$.

Remark 6 (Properties of $\mathbb{P}^n$) Of interest to symplectic geometers is the choice of symplectic cohomology class, for example when considering global homological mirror symmetry one considers the Kähler cone $\mathcal{K}_M$ of all possible Kähler classes on the symplectic manifold M. Information about $\mathcal{K}_M$ can be obtained using the Calabi-Yau theorem (stated for example in [21, Theorem 4.B.19]), that for compact Kähler manifolds there is a bijection between $\mathcal{K}_M$ and the set of Kähler forms ω with $\omega^n = \lambda \times vol$ for some $\lambda \in \mathbb{R}_{>0}$. As an example, consider the Fano manifold $\mathbb{P}^n$. The symplectic area of a projective line $\mathbb{P}^1 \subset \mathbb{P}^n$ is

$$\int_{\mathbb{P}^1} \omega_a = 2\pi a$$

and the cone of Kähler classes is one-dimensional corresponding to this parameter a. When $a = 1$, we obtain an integral cohomology class.

Furthermore, if $c_1(M) = [\alpha]$ is represented by a closed real $(1, 1)$-form α and we've fixed a choice of Kähler class β on M, there is a unique Kähler structure g on M so that $\alpha = \mathrm{Ric}(g)$ and $[\alpha] = \omega_g$ the Kähler form determined by the metric g. This is [21, Proposition 4.B.21]. Note that given any two of a metric, complex structure, and symplectic form on M which are compatible, the third is uniquely determined, see [21, p 29]. For example, the symplectic form ω_a and the standard complex structure on $\mathbb{P}^n$ determine a Kähler metric g_a on $\mathbb{P}^n$ which satisfies the Kähler-Einstein equation:

$$\mathrm{Ric}(g_a) = \frac{n+1}{a}\omega_a.$$

5 Connection to Mirror Symmetry

We give some context for mirror symmetry in which the above calculations play a role. We start with background.

5.1 Mirror Symmetry for Calabi-Yau Manifolds

Mirror symmetry relates the symplectic (resp. complex) geometry of a Calabi-Yau manifold X to the complex (resp. symplectic) geometry of a mirror Calabi-Yau manifold $\check{X}$ of the same dimension. Let (X, ω, J) be a Calabi-Yau manifold, where ω is the symplectic structure and J is the complex structure, and let $(\check{X}, \check{\omega}, \check{J})$ be the mirror Calabi-Yau manifold. Kontsevich's [27] Homological Mirror Symmetry (HMS) conjecture predicts the following equivalences of triangulated categories:

$$D^{\pi} Fuk(X, \omega) \cong D^b Coh(\check{X}, \check{J}), \tag{105}$$

$$D^b Coh(X, J) \cong D^{\pi} Fuk(\check{X}, \check{\omega}), \tag{106}$$

where $D^{\pi} Fuk$ is the split-closed derived Fukaya category (derived by taking the homotopy category so one obtains a triangulated category) and $D^b Coh$ is the bounded derived category of coherent sheaves.

In [38], N. Sheridan proved the equivalence (105) when X is a smooth Calabi-Yau hypersurface in the projective space $\mathbb{P}^n$. In this case, the mirror Calabi-Yau manifold $\check{X}$ is (a crepant resolution of) a Calabi-Yau hypersurface in the orbifold

$\mathbb{P}^n/G$, where $G = (\mathbb{Z}_{n+1})^n$. (The $n = 2$ case was first proved by Polishchuk-Zaslow [34], and the $n = 3$ case was first proved by P. Seidel [37].)

More generally, let X be a smooth Calabi-Yau hypersurface in a toric Fano manifold M. There is a one-to-one correspondence between projective Gorenstein Fano toric varieties and isomorphism classes of reflexive lattice polytopes Δ (see [10, Theorem 8.3.4], and in 2-dimensions, the 16 reflexive lattice polygons are listed in [10, p. 382]). The Batyrev mirror $\check{X}$ is (a crepant resolution of) a Calabi-Yau hypersurface in the Gorenstein toric Fano variety $\check{M}$ defined by the dual reflexive polytope $\check{\Delta}$ [7]. When $n > 3$, $H^{1,1}(X) = H^{1,1}(M) = H^2(M)$, and the Kähler moduli of the compact Calabi-Yau $(n-1)$-fold can be identified with the Kähler moduli of the ambient compact toric Fano manifold M.

5.2 Mirror Symmetry for Landau-Ginzburg Models

A smooth Calabi-Yau hypersurface X in a toric Fano manifold M can be identified with the critical locus of a holomorphic function W on the total space K_M of the canonical line bundle of M as follows. When M is Fano, K_M^* is ample and for a generic section s of K_M^*, the zero locus $X := s^{-1}(0) \subset M$ is a smooth anti-canonical divisor hence a manifold. In particular X is a compact Calabi-Yau manifold of complex dimension $n-1$. By the adjunction formula, $K_X \cong (K_M \otimes \mathcal{O}_M(X))|_X \cong \mathcal{O}_X$, which implies $K_M^*|_X \cong \mathcal{O}_M(X)|_X$. For example, when $M = U_a/N_{\mathbb{C}}$ where $U_a \subset \mathbb{C}^d$, let $s(z_1, \ldots, z_d) \in \mathbb{C}[z_1, \ldots, z_d]$ be a polynomial in $z_1, \ldots, z_d$ such that the rational function $\frac{s(z_1,\ldots,z_d)}{z_1 \cdots z_d}$ is invariant under the $N_{\mathbb{C}}$-action on $\mathbb{C}^d$. Then $s(z_1, \ldots, z_d)$ defines a section of the anti-canonical line bundle K_M^*.

Define the holomorphic function $W : K_M \to \mathbb{C}$ by $W(z, p) = \langle p, s(z)\rangle$, where $z \in M$, $p \in (K_M)_z$ (the fiber of K_M over z), and $\langle -, - \rangle$ is the pairing between dual vector spaces. (For example, for the polynomial s in the example of the previous paragraph, we see that $ps(z_1, \cdots, z_d) \in \mathbb{C}[z_1, \ldots, z_d, p]$ is invariant under the $N_{\mathbb{C}}$-action on $\mathbb{C}^{d+1}$ so descends to a well-defined holomorphic function $W : K_M \to \mathbb{C}$.) The critical locus of W is hence given by

$$\begin{aligned} \mathrm{Crit}(W) &= \{[z_1, \ldots, z_d, p] \in K_M : dW(z_1, \ldots, z_d, p) = 0\} \\ &= \{[z_1, \ldots, z_d, p] \in K_M : p = s(z_1, \ldots, z_d) = 0\}. \end{aligned}$$

Namely, the critical locus of W is exactly $X \subset M \subset K_M$ where the second inclusion is by the zero section: $\mathrm{Crit}(W) = s^{-1}(0) = X$. The pair (K_M, W) is an example of a *Landau-Ginzburg (LG) model* and W is known as the *superpotential*.

Example 17 (Fermat Surface a the Critical Locus of a Superpotential on $K_{\mathbb{P}^n}$) Let $s(z_1, \ldots, z_{n+1}) \in \mathbb{C}[z_1, \ldots, z_{n+1}]$ be a homogeneous polynomial of degree $n+1$, for example the Fermat polynomial in $(n+1)$-variables $\sum_{j=1}^{n+1} z_j^{n+1}$. Then ps is invariant under the $\mathbb{C}^*$-action on $(\mathbb{C}^{n+1} - \{0\}) \times \mathbb{C}$ and descends to a holomorphic

function

$$W : K_{\mathbb{P}^n} \to \mathbb{C}, \quad [z_1, \ldots, z_n, p] \mapsto ps(z_1, \ldots, z_{n+1})$$

which, in the case of the Fermat polynomial, has the following critical locus

$$\begin{aligned} \text{Crit}(W) &= \{[z_1, \ldots, z_{n+1}, p] \in K_{\mathbb{P}^n} : dW(z_1, \ldots, z_{n+1}, p) = 0\} \\ &= \{[z_1, \ldots, z_{n+1}, p] \in K_{\mathbb{P}^n} : \sum_{j=1}^{n+1} z_j^{n+1} = p = 0\} = s^{-1}(0) \cong X_{n+1}, \end{aligned}$$

where $X_{n+1} = \{[z_1, \ldots, z_{n+1}] \in \mathbb{P}^n : \sum_{j=1}^{n+1} z_j^{n+1} = 0\}$ is the Fermat Calabi-Yau hypersurface in $\mathbb{P}^n$. □

Similarly, the mirror $\check{X}$ can be identified with the critical locus of a holomorphic function $\check{W}$ on the total space $K_{\check{M}}$ of the canonical line bundle of $\check{M}$. The LG model $(K_{\check{M}}, \check{W})$ is mirror to the LG model (K_M, W). A natural formulation of the homological mirror symmetry conjecture in this setting is the following equivalences of triangulated categories:

$$\mathcal{W}(K_M, W) \cong MF(K_{\check{M}}, \check{W}), \tag{107}$$

$$MF(K_M, W) \cong \mathcal{W}(K_{\check{M}}, \check{W}), \tag{108}$$

where wrapping in non-compact fibers leads one to take $\mathcal{W}(-,-)$ the fiber-wise wrapped Fukaya category (being defined in [1], see also [22]) of the LG model, and MF is the category of matrix factorizations.

The B-model $MF(K_M, W = \langle p, s(z)\rangle)$ on the LG model is equivalent to the B-model $D^bCoh(X)$ on the Calabi-Yau hypersurface $X = s^{-1}(0) = \text{Crit}(W) \subset M$, as a consequence of Orlov's generalized Knörrer periodicity theorem [33]. An A-model version of this Knörrer periodicity theorem would be an equivalence between $\mathcal{W}(K_M, W)$ and $D^\pi Fuk(X)$. Recently [22] has proven a version of Knörrer periodicity for the A-model that uses the notion of a partially wrapped Fukaya category. Also see [35, 41]. With these equivalences, Eqs. (107) and (108) is equivalent to Eqs. (105) and (106) when $X = s^{-1}(0)$ is a smooth CY hypersurface in M that is the critical locus of $W = \langle p, s\rangle$ on K_M.

Now let us consider a different LG model of $(K_M, W = z_1 \cdots z_d\, p)$, again $M^n = \mathbb{C}^d /\!\!/ N$ is a toric Fano manifold of dimension n obtained via symplectic reduction from $\mathbb{C}^d$ and $z \in \mathbb{C}^d$ are the homogeneous coordinates. The critical locus of W in this case is a singular CY hypersurface in M defined by $z_1 \cdots z_d = 0$, and it is the preimage $\mu^{-1}(\partial\Delta)$ of the boundary of the moment polytope for M. The LG model $(K_M, W = z_1 \cdots z_d\, p)$ captures the geometry of this singular $\text{Crit}(W)$ via Knörrer periodicity [22, 33], and (K_M, W) turns out to be the generalized SYZ [40] mirror (in the sense of [3]) of a smooth hypersurface Σ in $(\mathbb{C}^*)^n$. For example, if M is a

Fano surface defined by a reflexive polygon with n vertices then (K_M, W) is the generalized SYZ mirror of an n-punctured torus Σ (such as the thrice punctured torus $\Sigma = \{1 + x + y + t/xy = 0\} \subset (\mathbb{C}^*)^2$, whose mirror is the LG model $(K_{\mathbb{P}^2}, W = z_1 z_2 z_3 p)$).

More generally, the canonical bundle K_M is a special case of a toric Calabi-Yau manifold $Y = \mathbb{C}^{d+1} /\!\!/ \mathbb{T}^{d-n}$ of dimension $(n + 1)$. The function $z_1 \cdots z_d z_{d+1}$ on $\mathbb{C}^{d+1}$ descends to a well-defined holomorphic function $W : Y \to \mathbb{C}$. It was first proposed by Hori-Vafa [20] and then proven in the SYZ framework by Abouzaid-Auroux-Katzarkov [3] that LG-models given by (Y, W) for Y a toric CY manifold are generalized SYZ mirrors to hypersurfaces Σ in toric varieties. The HMS prediction would be the equivalence of the following triangulated categories:

$$\mathcal{W}(Y, W) \cong D^b Coh(\Sigma) \tag{109}$$

$$MF(Y, W) \cong \mathcal{W}(\Sigma), \tag{110}$$

where $\mathcal{W}(\Sigma)$ is the wrapped Fukaya category of Σ, wrapped due to Σ being a noncompact Liouville manifold. To define the fiber-wise wrapped Fukaya category $\mathcal{W}(Y, W)$ one would need to view $W : Y \to \mathbb{C}$ as a symplectic fibration, which requires a good understanding of the symplectic structure on Y. Equivalence (109) is the subject of the work in preparation by Abouzaid-Auroux [1] when Σ is an algebraic hypersurface in $(\mathbb{C}^*)^n$ and Y is a toric Calabi-Yau $(n+1)$-fold. In the other direction, the third author [28] proved the equivalence (110) when $n = 2$ where $\Sigma \subset (\mathbb{C}^*)^2$ is a punctured Riemann surface via decomposition into pair-of-pants (thrice punctured spheres) and applying the result of [2], which establishes (110) when Σ is a punctured sphere. Lekili-Polishchuk [29] proved the equivalence (110) when $\Sigma \subset (\mathbb{C}^*)^n$ is a generalized higher dimensional pair-of-pants. A version of the equivalence in Eq. (110) is proven from the microlocal perspective [13] using localization results from [14].

Further generalizations are given in [3, Section 10], such as to the SYZ mirrors for hypersurfaces Σ of abelian varieties per the speculation of Seidel [36]. In [9], the second author proved a HMS result for genus-2 compact Riemann surfaces Σ_2 that are hypersurfaces in an abelian variety $V = (\mathbb{C}^*)^2 / \Gamma_B$ ($\Gamma_B \cong \mathbb{Z}^2$) and its generalized SYZ mirror (Y, v_0). Here, she considers a LG model (Y, v_0) where $Y = \widetilde{Y} / \Gamma_B$ is the quotient of a toric Calabi-Yau threefold $\widetilde{Y}$ of infinite type by the free action of Γ_B. The main result in [9] is a fully-faithful embedding

$$D^b Coh(\Sigma_2) \hookrightarrow H^0 \mathcal{FS}(Y, v_0). \tag{111}$$

where $\mathcal{FS}$ denotes the Fukaya-Seidel category with compact fibers that are abelian varieties degenerating to the following critical locus; the non-compact Calabi-Yau threefold Y contains a "banana" configuration of three 2-spheres $C_1 \cup C_2 \cup C_3$ that intersect at two triple intersection points. The Kähler moduli of Y is three-dimensional and the real Kähler parameters are given by the symplectic areas A_i of C_i. The complex moduli $\mathcal{M}_2$ of Σ_2 is a complex orbifold of dimension 3. In [9], the

second author considers a one-parameter family of symplectic structures on Y with $A_1 = A_2 = A_3$, which corresponds to a one-parameter family of complex structures on Σ_2.

In [6], we extend the work [9] to any genus 2 curve in $\mathcal{M}_2$. To do that, we need to construct more general symplectic structures on Y where the areas of the three 2-spheres may vary independently, which inspired us to do the exercises that are in this paper. This is naturally done in the action-angle coordinates using Guillemin's Kähler potential. The canonical bundle K_M is a local model for $\widetilde{Y}$, and in the example below, we express the superpotential also in the angle-action coordinates.

Example 18 (The Superpotential $W : K_M \to \mathbb{C}$ in Action-Angle Coordinates) Let $s(z_1, \ldots, z_d) = z_1 \cdots z_d$ and $W : K_M \to \mathbb{C}[z_1, \ldots, z_d, p] \mapsto pz_1 \cdots z_d$ a morphism between smooth toric varieties. We are interested in the A-model on the Landau-Ginzburg model (K_M, W), so we would like to view W as a symplectic fibration. We write it in terms of the action-angle coordinates (ξ, θ) instead of the complex homogeneous coordinates $(z_1, \ldots, z_d, p)$. In particular, since the superpotential is defined on the symplectic quotient K_M, it is independent of the choice of representative in the homogeneous coordinates. We could write it in terms of the affine coordinates corresponding to the Jth chart where the complement of the Jth z_i's are scaled to 1, so it would still be the product of all the inhomogeneous coordinates, or we could write it in terms of the coordinates on the $\mathbb{T}^n_{\mathbb{C}}$ which parametrizes all those charts via choices of $\alpha_{\mathbb{C}}$. In that case, it would just be projection onto the last coordinate, t_{n+1}.

More specifically, let $v : \Delta^+ \times \mathbb{T}^{n+1} \to \mathbb{C}$ be the composition

$$v : \Delta^+ \times \mathbb{T}^{n+1} \xrightarrow{\cong} \mathbb{T}^{n+1}_{\mathbb{C}} \xrightarrow{\alpha_{\mathbb{C}}} \mathbb{T}^{d+1}_{\mathbb{C}} \hookrightarrow U_a \times \mathbb{C} \xrightarrow{\widetilde{\pi}_a^+} K_M \xrightarrow{W} \mathbb{C}. \tag{112}$$

where $\widetilde{\pi}_a^+ : U_a \times \mathbb{C} \to K_M$ is the projection map defined in Eq. (27). Namely we compose the exponential map

$$(\xi_1, \ldots, \xi_{n+1}, e^{i\theta_1}, \cdots, e^{i\theta_{n+1}}) \mapsto (e^{\xi_1 + i\theta_1}, \ldots, e^{\xi_n + i\theta_{n+1}})$$

with $\alpha_{\mathbb{C}}(t_1, \ldots, t_{n+1})$ which, up to permutation is $(y_1^J, \ldots, y_{n+1}^J, 1, \ldots, 1)$, and then multiply everything together to obtain $\prod_{k=1}^{n+1} y_k^J$. Now recall Definition 3, which defines the $\alpha_{\mathbb{C}}$ injection for K_M, namely what the y_k^J are as functions of t_k. Plugging in for t_k, we see that

$$\begin{aligned} v(\xi_1, \ldots, \xi_{n+1}, e^{i\theta_1}, \cdots, e^{i\theta_{n+1}}) &= z_1 \ldots z_d p = y_1^J \ldots y_{n+1}^J \\ &= t_{n+1} = \exp\left(\frac{\partial G^+}{\partial \xi_{n+1}}(\xi) + i\theta_{n+1}\right). \end{aligned} \tag{113}$$

Namely the superpotential is projection onto the last coordinate in the parameters $(t_1, \ldots, t_{n+1})$ because that equals the product of the coordinates in $\alpha_{\mathbb{C}}(t_1, \ldots, t_{n+1})$. □

5.3 Monodromy in Mirror Symmetry

Landau-Ginzburg models, of which (K_M, W) is a prototype example, are important in physics and math, in particular in mirror symmetry. In particular, moment maps are an example of the more general notion of a *Lagrangian torus fibration* which is the input needed to obtain a mirror when using the prescription given by SYZ mirror symmetry [40]. Thus in the case of K_M there are two interesting fibrations, one is the Lagrangian torus fibration coming from the moment map we've described, and the other is (K_M, W) which is a singular symplectic fibration with e.g. the symplectic form coming from symplectic reduction. For manifolds which are Fano or of general type ($c_1 < 0$), their mirrors are Landau-Ginzburg models [3, 19]. So it is crucial to be able to understand the Fukaya category of a Landau-Ginzburg model, which is the algebraic input for the symplectic side of homological mirror symmetry. To do so, one may start by looking at the Floer theory of (K_M, W) for noncompact Lagrangians which interact well with $W : K_M \to \mathbb{C}$. Two types of non-compact Lagrangians are often used.

One type of non-compact Lagrangian considered consists of those in [35] which are *thimbles* obtained by parallel transporting a Lagrangian in a fiber of W to the singular locus over 0; this works when W is a Lefschetz fibration. However, for more complicated singular fibers, this process could produce a singular Lagrangian if the fiber Lagrangian degenerates to a submanifold which is of codimension more than 1 from what it was in the generic fiber. In this case, Lagrangians which are *U-shaped* may be used; they go around the critical value(s) of W, where often that value is 0. Therefore, it is of interest to know the monodromy around the origin of $W : K_M \to \mathbb{C}$. This is what brought us to do the work in this paper, as we are interested in the Fukaya category of a certain LG-model in forthcoming work [6].

Lastly, the theory of the symplectic quotient extends not only to noncompact manifolds, but to toric varieties of *infinite type*. Specifically, one example is the following: let Y be the example mentioned in the paragraph surrounding Eq. (111). This is denoted $X_{(1,1)}$ in the SYZ mirror symmetry paper by Kanazawa and Lau [23]. It can be equipped with the following symplectic form, which is a scaled version of that from symplectic reduction, to account for the infinitely many facets.

Definition 4 (Definition of ω from [23]) There exists an open covering $\{U_v\}$ of the moment polyhedron $\Delta_{\tilde{Y}}$, and non-negative bump functions ρ_v on $\mathbb{R}^3$ which are supported on U_v and identically 1 in a smaller neighborhood of the boundary stratum, such that the following

$$\tilde{G}(y) := \frac{1}{2} \sum_{v \in \Sigma(1)} \rho_v(y) L_v(y) \log L_v(y) \tag{114}$$

is a finite sum at each point $y \in \mathring{\Delta}_{\tilde{Y}}$, and whose Legendre transform is Γ_B equivariant and defines a Kähler potential for Y, (namely its second-order derivatives are

positive). Here, as earlier, $L_v \geq 0$ denotes the half spaces whose intersection defines $\Delta_{\tilde{Y}}$.

Computing monodromy in this case requires a bit more finesse and the computation is done by splitting into three regions: one is around a vertex of the polytope which is modeled on an open neighborhood around $0 \in \mathbb{C}^3$, one is in a neighborhood of the center of a hexagon modeled on $\mathrm{tot}\left(K_{\mathbb{CP}^2(3)}\right)$, and the third is the remaining region between these two which is harder to compute directly but can be estimated and computed up to Hamiltonian isotopy.

In conclusion, toric varieties provide a rich source of symplectic manifolds, including compact, noncompact, and of infinite type, and this article describes how to compute information about them of interest to symplectic geometers.

6 Notation

This is an index for notation appearing throughout the text, with some comparisons to notation used in [16].

Section 2: Toric Manifolds as Symplectic Quotients

- X in Guillemin [16] is called M or K_M here
- $\cdot$ denotes a group action as specified in the context
- $\mathbb{I}_k$ denotes the $k \times k$ identity matrix
- $\mathbb{T}^k_{\mathbb{C}}$ is the complex algebraic torus $(\mathbb{C}^*)^k$ of dimension k
- $\mathbb{T}^k$ is the real torus of dimension k obtained as the maximal compact subgroup $U(1)^k$ of $\mathbb{T}^k_{\mathbb{C}}$ by restricting to norm 1 coordinates
- $n = \dim_{\mathbb{C}} M$, $n + 1 = \dim_{\mathbb{C}} K_M$
- Maps involved in symplectic reduction:
  - $\rho_{\mathbb{C}} : N_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}}$ is an embedding and $Q = (Q^l_k)_{k,l}$ is the matrix representing its linearization
  - $\beta_{\mathbb{C}} : \mathbb{T}^d_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}}/N_{\mathbb{C}} \cong \mathbb{T}^n_{\mathbb{C}}$ is the quotient map and $B = (v^k_m)_{m,k}$ is the matrix representing its linearization
  - superscripts denote the column, subscripts denote the row
- $\mu_{\mathbb{T}^d}$ is the moment map for the standard Hamiltonian $\mathbb{T}^d$-action on $\mathbb{C}^d$
- μ_N is the moment map for the Hamiltonian N-action on $\mathbb{C}^d$
- $U_a \subset \mathbb{C}^d$ is the open set where $N_{\mathbb{C}}$ acts freely and which contains $\mu_N^{-1}(a)$
- $U_a^+ = U_a \times \mathbb{C}$
- $R_a : U_a \to \mu_N^{-1}(a)$ retracts U_a onto $\mu_N^{-1}(a)$ via a choice of $\lambda_a(z) \in N_{\mathbb{C}}$ for each $z \in U_a$
- symbols with a $+$ refer to the K_M case
- Z in [16] is a level set, here called $\mu_N^{-1}(a)$ for M or $\mu_N^{+-1}(a)$ for K_M
- $z_1, \ldots, z_d$ are the homogeneous coordinates on $\mathbb{C}^d$ (with additional coordinate p in the case of K_M)

- $r_1, \ldots, r_d, \varphi_1, \ldots, \varphi_d$ are the polar coordinates of the homogeneous coordinates
- Δ denotes the moment polytope (when bounded, for compact M) and Δ^+ denotes the moment polyhedron (when unbounded, for noncompact K_M)
- Projection maps:

$$\begin{cases} \pi_a : (\mu_N)^{-1}(a) \to M = (\mu_N)^{-1}(a)/N, \\ \widetilde{\pi}_a : U_a \to M = U_a/N_{\mathbb{C}}, \end{cases}$$

defined in Theorems 1 and 2, and

$$\begin{cases} \pi_a^+ : (\mu_N^+)^{-1}(a) \to K_M = (\mu_N^+)^{-1}(a)/N, \\ \widetilde{\pi}_a^+ : U_a \times \mathbb{C} \to K_M = (U_a \times \mathbb{C})/N_{\mathbb{C}}, \end{cases}$$

defined in Eq. (27)

Section 3: Toric Actions and Moment Maps

- $J = (j_1, \ldots, j_r)$ denotes a multi-index where $j_l \in \{1, \ldots, d\}$ are strictly increasing (and will index a choice of r facets)
- $\mathbb{C}^d_J = \mathrm{orb}_{\mathbb{T}^d_{\mathbb{C}}}(z) \cong \mathbb{C}^{d-r}$ denotes the orbit under the standard $\mathbb{T}^d_{\mathbb{C}}$-action of any point $(z_1, \ldots, z_d)$ with $z_j = 0 \iff j \in J$
- $(\mathbb{T}^d_{\mathbb{C}})_J = \mathrm{stab}_{\mathbb{T}^d_{\mathbb{C}}}(z) \cong (\mathbb{C}^*)^r$, for any fixed choice $z \in \mathbb{C}^d_J$
- $\alpha_{\mathbb{C}} : \mathbb{T}^n_{\mathbb{C}} \to \mathbb{T}^d_{\mathbb{C}}$ is a right inverse to $\beta_{\mathbb{C}}$ and $A = (s^m_k)_{k,m}$ is the matrix representing its linearization
- μ_a is the moment map of the Hamiltonian $\mathbb{T}^n$-action on $\mu^{-1}(a)/N$
- $(d\rho_{\mathbb{C}})^*_1(\kappa) = -a$
- l_i in [16] is called L_i here, which denotes the affine linear defining equation of the ith facet of polyhedron Δ
- $\mathcal{F}_j$ denotes the jth facet of the moment polytope
- $J_f = (j_1, \ldots, j_r)$ indexes the facets which intersect with polytope Δ in a face f, e.g. $r = n$ when $f = v$ a vertex
- $t_1, \ldots, t_n$ are the inhomogeneous coordinates on the dense $(\mathbb{C}^*)^n = \mathbb{T}^n_{\mathbb{C}}$
- $\mathring{\Delta}$ denotes the interior of the polytope
- $u_j = \log t_j$ are coordinates on the Lie algebra $\mathrm{Lie}\, \mathbb{T}^n_{\mathbb{C}} = \mathfrak{t}^n_{\mathbb{C}}$
- $x_1, \ldots, x_n$ and $\theta_1, \ldots, \theta_n$ are the polar coordinates of the inhomogeneous coordinates $(t_1, \ldots, t_n) \in (\mathbb{C}^*)^n$
- $\xi_1, \ldots, \xi_n$ are the moment map coordinates
- ξ_i, θ_i are the action-angle coordinates
- Notation for Sect. 3.3: Holomorphic coordinate charts for M
  - v is a vertex of the polytope Δ
  - For vertex v and corresponding indexing set J_v,

$$\widetilde{V}_{J_v} = \{(z_1, \ldots, z_d) \in \mathbb{C}^d \mid z_j \neq 0 \text{ if } j \notin J_v\} \cong (\mathbb{C}^*)^{d-n} \times \mathbb{C}^n \subset U_a$$

- $V_{J_v} = \widetilde{V}_{J_v}/N_{\mathbb{C}} \cong \mathbb{C}^n \subset M$ is the open chart of M corresponding to vertex v, the union of which cover M
- $y_1^{J_v}, \ldots, y_n^{J_v}$ (defined in Eq. (65)) denote affine coordinates on the chart V_{J_v}
- $\varphi^{J_v} : V_{J_v} \to \mathbb{C}^n$ is the chart map defined by $(y_1^{J_v}, \ldots, y_n^{J_v})$
- $p^{J_v} = \mu_a^{-1}(v)$ is a $\mathbb{T}^n_{\mathbb{C}}$ toric fixed point corresponding to vertex v, and it's the center of the chart φ^{J_v}
- $U_{J_v} = \{(z_1, \ldots, z_d) \in \mathbb{C}^d \mid z_j = 1 \text{ if } j \notin J_v\} \subset \widetilde{V}_{J_v} \subset \mathbb{C}^d$ is the slice of the $N_{\mathbb{C}}$ action given by the representatives in Eq. (65), where the letter U is used because it is a subset of U_a
- $\widehat{B}$, $\widehat{A}^{J_v}$ denote choices of bases for matrices B, A so that the corresponding $\alpha_{\mathbb{C}}^{J_v}$ gives the standard $\mathbb{T}^n_{\mathbb{C}}$ action, as described in the paragraph of Section "Change of Basis of $\mathbb{T}^n_{\mathbb{C}}$ and the Embedding of $\mathbb{T}^n_{\mathbb{C}}$ in Each $\mathbb{C}^n$-chart"

Section 4: Kähler Potential

- F is the Kähler potential and is a function of the x_j
- G is the Legendre transform of F and is a function of the ξ_j
- the jth coordinate function of $\mu_a(t_1, \ldots, t_n)$ is ξ_j and equals $\partial F/\partial x_j$

Section 5: Connection to Mirror Symmetry

- $W : K_M \to \mathbb{C}$ denotes the superpotential
- $s \in \Gamma(M, K_M^*)$ is a generic section
- $X = \mathrm{Crit}(W)$ is the critical locus of the superpotential
- $\Sigma \subset (\mathbb{C}^*)^n$ is a hypersurface whose generalized SYZ mirror is a LG model (K_M, W)
- Σ_2 is the genus 2 curve
- (Y, v_0) is the SYZ mirror of $(\mathrm{Bl}_{\Sigma \times \{0\}} V \times \mathbb{C}, y)$, both LG models
- $\widetilde{Y}$ is the universal cover of Y, a toric variety of infinite type
- v is W written as a function of the moment map coordinates ξ_j

Acknowledgments We thank Chiu-Chu Melissa Liu for her significant involvement in an earlier draft of this paper and for her continued support. We would like to thank Ana Rita Pires for her encouragement and comments on an earlier draft. We thank the referees for their helpful and detailed comments. We are also grateful to the 2019 Research Collaboration Conference for Women in Symplectic and Contact Geometry and Topology (WiSCon), during which this project was initiated, and the Institute for Computational and Experimental Research in Mathematics (ICERM) at Brown University for hosting this conference.

References

1. Abouzaid, M., Auroux, D.: Homological mirror symmetry for hypersurfaces in $(\mathbb{C}^*)^n$ (in preparation)
2. Abouzaid, M., Auroux, D., Efimov, A.I., Katzarkov, L., Orlov, D.: Homological mirror symmetry for punctured spheres. J. Amer. Math. Soc. **26**(4), 1051–1083 (2013). https://doi.org/10.1090/S0894-0347-2013-00770-5

3. Abouzaid, M., Auroux, D., Katzarkov, L.: Lagrangian fibrations on blowups of toric varieties and mirror symmetry for hypersurfaces. Publ. Math. Inst. Hautes Études Sci. **123**, 199–282 (2016). https://doi.org/10.1007/s10240-016-0081-9
4. Atiyah, M.F.: Convexity and commuting Hamiltonians. Bull. London Math. Soc. **14**(1), 1–15 (1982). https://doi.org/10.1112/blms/14.1.1
5. Audin, M.: The topology of torus actions on symplectic manifolds, *Progress in Mathematics*, vol. 93. Birkhäuser Verlag, Basel (1991). https://doi.org/10.1007/978-3-0348-7221-8. Translated from the French by the author
6. Azam, H., Cannizzo, C., Lee, H., Liu, C.: Global homological mirror symmetry for genus two curves (in preparation)
7. Batyrev, V.V.: Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties. J. Algebraic Geom. **3**(3), 493–535 (1994)
8. Calderbank, D.M.J., David, L., Gauduchon, P.: The Guillemin formula and Kähler metrics on toric symplectic manifolds. J. Symplectic Geom. **1**(4), 767–784 (2003). http://projecteuclid.org/euclid.jsg/1092749568
9. Cannizzo, C.: Categorical mirror symmetry on cohomology for a complex genus 2 curve. Advances in Mathematics **375**, 107,392 (2020). https://doi.org/10.1016/j.aim.2020.107392. https://www.sciencedirect.com/science/article/pii/S0001870820304205
10. Cox, D.A., Little, J.B., Schenck, H.K.: Toric varieties, *Graduate Studies in Mathematics*, vol. 124. American Mathematical Society, Providence, RI (2011). https://doi.org/10.1090/gsm/124
11. Delzant, T.: Hamiltoniens périodiques et images convexes de l'application moment. Bull. Soc. Math. France **116**(3), 315–339 (1988). http://www.numdam.org/item?id=BSMF_1988__116_3_315_0
12. Fulton, W.: Introduction to toric varieties, *Annals of Mathematics Studies*, vol. 131. Princeton University Press, Princeton, NJ (1993). https://doi.org/10.1515/9781400882526. The William H. Roever Lectures in Geometry
13. Gammage, B., Shende, V.: Mirror symmetry for very affine hypersurfaces (2018). https://arxiv.org/abs/1707.02959
14. Ganatra, S., Pardon, J., Shende, V.: Covariantly functorial wrapped Floer theory on Liouville sectors. Publ. Math. Inst. Hautes Études Sci. **131**, 73–200 (2020). https://doi.org/10.1007/s10240-019-00112-x
15. Griffiths, P., Harris, J.: Principles of algebraic geometry. Wiley Classics Library. John Wiley & Sons, Inc., New York (1994). https://doi.org/10.1002/9781118032527. Reprint of the 1978 original
16. Guillemin, V.: Kaehler structures on toric varieties. J. Differential Geom. **40**(2), 285–309 (1994). http://projecteuclid.org/euclid.jdg/1214455538
17. Guillemin, V.: Moment maps and combinatorial invariants of Hamiltonian T^n-spaces, *Progress in Mathematics*, vol. 122. Birkhäuser Boston, Inc., Boston, MA (1994). https://doi.org/10.1007/978-1-4612-0269-1
18. Guillemin, V., Sternberg, S.: Convexity properties of the moment mapping. Invent. Math. **67**(3), 491–513 (1982). https://doi.org/10.1007/BF01398933
19. Hori, K., Katz, S., Klemm, A., Pandharipande, R., Thomas, R., Vafa, C., Vakil, R., Zaslow, E.: Mirror symmetry, *Clay Mathematics Monographs*, vol. 1. American Mathematical Society, Providence, RI; Clay Mathematics Institute, Cambridge, MA (2003). With a preface by Vafa
20. Hori, K., Vafa, C.: Mirror symmetry (2000). https://arxiv.org/abs/hep-th/0002222
21. Huybrechts, D.: Complex geometry. Universitext. Springer-Verlag, Berlin (2005). An introduction
22. Jeffs, M.: Mirror symmetry and Fukaya categories of singular hypersurfaces (2020). https://arxiv.org/abs/2012.09764
23. Kanazawa, A., Lau, S.C.: Local Calabi-Yau manifolds of type $\tilde{A}$ via SYZ mirror symmetry. J. Geom. Phys. **139**, 103–138 (2019). https://doi.org/10.1016/j.geomphys.2018.12.015
24. Karshon, Y., Lerman, E.: Non-compact symplectic toric manifolds. SIGMA Symmetry Integrability Geom. Methods Appl. **11**, Paper 055, 37 (2015). https://doi.org/10.3842/SIGMA.2015.055

25. Kempf, G., Ness, L.: The length of vectors in representation spaces. In: Algebraic geometry (Proc. Summer Meeting, Univ. Copenhagen, Copenhagen, 1978), *Lecture Notes in Math.*, vol. 732, pp. 233–243. Springer, Berlin (1979)
26. Kirwan, F.C.: Cohomology of Quotients in Symplectic and Algebraic Geometry. (MN-31), Volume 31, vol. 104. Princeton University Press, Princeton, NJ (1984). https://doi.org/10.2307/j.ctv10vm2m8. https://www.jstor.org/stable/j.ctv10vm2m8
27. Kontsevich, M.: Homological algebra of mirror symmetry. In: Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), pp. 120–139. Birkhäuser, Basel (1995)
28. Lee, H.: Homological mirror symmetry for open Riemann surfaces from pair-of-pants decompositions https://arxiv.org/abs/1608.04473
29. Lekili, Y., Polishchuk, A.: Homological mirror symmetry for higher-dimensional pairs of pants. Compos. Math. **156**(7), 1310–1347 (2020). https://doi.org/10.1112/s0010437x20007150
30. Marsden, J., Weinstein, A.: Reduction of symplectic manifolds with symmetry. Rep. Mathematical Phys. **5**(1), 121–130 (1974). https://doi.org/10.1016/0034-4877(74)90021-4
31. McDuff, D., Salamon, D.: Introduction to symplectic topology, second edn. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York (1998)
32. Meyer, K.R.: Symmetries and integrals in mechanics. In: Dynamical systems (Proc. Sympos., Univ. Bahia, Salvador, 1971), pp. 259–272 (1973)
33. Orlov, D.O.: Triangulated categories of singularities and D-branes in Landau-Ginzburg models. Tr. Mat. Inst. Steklova **246**(Algebr. Geom. Metody, Svyazi i Prilozh.), 240–262 (2004)
34. Polishchuk, A., Zaslow, E.: Categorical mirror symmetry: the elliptic curve. Adv. Theor. Math. Phys. **2**(2), 443–470 (1998). https://doi.org/10.4310/ATMP.1998.v2.n2.a9
35. Seidel, P.: Fukaya categories and Picard-Lefschetz theory. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich (2008). https://doi.org/10.4171/063
36. Seidel, P.: Some speculations on pairs-of-pants decompositions and Fukaya categories. In: Surveys in differential geometry. Vol. XVII, *Surv. Differ. Geom.*, vol. 17, pp. 411–425. Int. Press, Boston, MA (2012). https://doi.org/10.4310/SDG.2012.v17.n1.a9
37. Seidel, P.: Homological mirror symmetry for the quartic surface. Mem. Amer. Math. Soc. **236**(1116), vi+129 (2015). https://doi.org/10.1090/memo/1116
38. Sheridan, N.: Homological mirror symmetry for Calabi-Yau hypersurfaces in projective space. Invent. Math. **199**(1), 1–186 (2015). https://doi.org/10.1007/s00222-014-0507-2
39. Cannas da Silva, A.: Lectures on symplectic geometry, *Lecture Notes in Mathematics*, vol. 1764. Springer-Verlag, Berlin (2001). https://doi.org/10.1007/978-3-540-45330-7
40. Strominger, A., Yau, S.T., Zaslow, E.: Mirror symmetry is T-duality. Nuclear Phys. B **479**(1–2), 243–259 (1996). https://doi.org/10.1016/0550-3213(96)00434-8
41. Wehrheim, K., Woodward, C.T.: Exact triangle for fibered Dehn twists. Res. Math. Sci. **3**, Paper No. 17, 75 (2016). https://doi.org/10.1186/s40687-016-0065-x

An Introduction to Weinstein Handlebodies for Complements of Smoothed Toric Divisors

Bahar Acu, Orsola Capovilla-Searle, Agnès Gadbled, Aleksandra Marinković, Emmy Murphy, Laura Starkston, and Angela Wu

2000 Mathematics Subject Classification Primary: 57R17; Secondary: 53D05, 53D10

1 Introduction

A key way to study closed symplectic manifolds is to break them down into two more easily understood parts: a neighborhood of a *divisor* and a complementary *Weinstein domain*. A divisor is a symplectic submanifold of co-dimension 2. One

B. Acu (✉)
Department of Mathematics, ETH Zürich, Zürich, Switzerland
e-mail: bahar.acu@math.ethz.ch

O. Capovilla-Searle
Department of Mathematics, Duke University, Durham, NC, USA
e-mail: orsola.capovilla.searle@duke.edu

A. Gadbled
Université Paris-Saclay, CNRS, Laboratoire de Mathématiques d'Orsay, Orsay, France
e-mail: agnes.gadbled@universite-paris-saclay.fr

A. Marinković
Matematički Fakultet, Belgrade, Serbia
e-mail: aleks@matf.bg.ac.rs

E. Murphy
Department of Mathematics, Northwestern University, Evanston, IL, USA
e-mail: e_murphy@math.northwestern.edu

L. Starkston
Department of Mathematics, UC Davis, Davis, CA, USA
e-mail: lstarkston@math.ucdavis.edu

A. Wu
Department of Mathematics, University College London, London, UK
e-mail: angela.wu.17@ucl.ac.uk

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_4

can allow this submanifold to have certain controlled singularities, such as normal crossing singularities or more general singularities modeled on complex hypersurfaces. Donaldson proved that every symplectic manifold has a divisor [8] and Giroux proved that this divisor can be chosen such that the complement of a regular neighborhood admits the structure of a Weinstein domain [17, 18]. A *Weinstein domain* is a symplectic manifold with convex contact type boundary which can be broken down into symplectic handles modeled and glued as described by Weinstein [24]. The symplectic topology of a Weinstein manifold is encoded in the attaching spheres of the handles. The attaching spheres can be represented by isotropic and Legendrian knots in the front projection. This Weinstein handlebody diagram gives a combinatorial/diagrammatic method to encode a symplectic manifold. There is a calculus of moves which relates different diagrams for equivalent Weinstein manifolds [7, 19].

Recently, there has been increased study of symplectic divisors in symplectic manifolds, particularly in the case when the complement is Weinstein. Some of the motivation comes from homological mirror symmetry, where generalizing the link between coherent sheaves and Fukaya categories to larger classes of manifolds has required one to consider a mirror pair that includes not only a space, but also a divisor [3]. One way to associate a Fukaya category for a divisor pair, is to look at the wrapped Fukaya category of the complement of the divisor. The Weinstein handle decomposition is key to understanding the wrapped Fukaya category, due to recent results that the co-cores of the handles generate the category [5, 20, 21]. The Floer homology of these co-cores is intrinsically tied to the Legendrian DGA of the Weinstein handlebody diagram [4, 9, 10] which is combinatorially calculated by Ekholm-Ng [12].

An important class of symplectic manifolds are *toric manifolds*. These have been studied extensively as they form a large class of examples of integrable systems because of the symmetry provided by the Hamiltonian action of a torus on such manifolds. According to the famous Delzant classification, all compact symplectic toric $2n$-manifolds are uniquely (up to equivariant symplectomorphism) determined by convex n-dimensional polytopes, which correspond to the orbit space of the action. Much of the symplectic information can be encoded in the combinatorics of these polytopes known as Delzant polytopes. Moreover, toric manifolds have their origin in algebraic geometry, and they come by definition with a fibration by tori given by the action, so that they have been among the first cases of interest for homological mirror symmetry, especially in view of the SYZ philosophy (see for instance [1]). Every compact toric symplectic manifold is naturally equipped with a toric divisor. This is precisely the set of all points with non trivial stabilizer and the fixed points of the toric action are normal crossing singularities of the divisor. This can also be understood in terms of the moment map image polytope: the toric divisor is the preimage of the faces of the polytope under the moment map. The complement of a neighborhood of the divisor is symplectomorphic to a Weinstein domain whose completion is T^*T^n. (The complement of the divisor is the preimage of the interior of the polytope under the moment map, which is $T^n \times P$ where P is a convex open subset of $\mathbb{R}^n$.) Hypersurfaces with normal crossing singularities

can naturally be deformed to become less singular at the expense of increasing the topological complexity of the divisor and its complement. A toric manifold, together with its toric divisor or any smoothing of the divisor is a log Calabi-Yau pair which is a convenient setting for studying mirror symmetry of a space with a divisor [16, 22].

A manifold of dimension 4 will have symplectic surface divisors. Normal crossing singularities in this dimension are just positive transverse intersections of two smooth branches, or *nodes*. A deformation of this node smooths out the surface, trading the node for an annular tube which thus joins two different components or increases the genus of the surface. For a toric 4-manifold, the complement of the (fully singular) toric divisor looks like T^*T^2, which has a natural Weinstein structure described by a diagram discovered by Gompf [19].

The toric divisor is the preimage of the facets of the Delzant polytope Δ, and the nodes are the preimages of the vertices of Δ. Since there is a one to one correspondence between the nodes and vertices, we will use the same notation to denote both a node and its moment map image vertex. Each vertex $V \in \Delta$, has a corresponding ray r based at V defined as the sum of the primitive edge vectors of Δ adjacent to V and pointing outward from V.

Definition 1 A toric manifold with a chosen subset $\{V_1, \ldots, V_n\}$ of the nodes is $\{V_1, \ldots, V_n\}$-centered, if the corresponding rays $r_1, \ldots, r_n$ all intersect at a common single point in the interior of its Delzant polytope.

We show in [2, Theorem 4.1] that if a toric 4-manifold is $\{V_1, \ldots, V_n\}$-centered, then the complement of the toric divisor smoothed at the nodes $\{V_1, \ldots, V_n\}$ has a Weinstein structure which we can explicitly describe.

In this article, we explain an algorithm to produce a Weinstein handlebody diagram for the complement of any divisor obtained by smoothing the $\{V_1, \ldots, V_n\}$-nodes of a toric divisor in a $\{V_1, \ldots, V_n\}$-centered toric 4-manifold. We prove the Weinstein handlebody produced by this algorithm is Weinstein homotopic to the complement of the smoothed toric divisor in our later paper [2]. That article also explains how toric moment data determines the input to our algorithm, provides a detailed exploration of the centered hypothesis, and includes many more examples. This article focuses on the most accessible example, as well as showcasing one fun case.

1.1 Main Results

There exists an algorithm to produce a Weinstein handlebody diagram for the complement of a toric divisor smoothed at the subset of nodes $\{V_1, \ldots, V_n\}$ in a $\{V_1, \ldots, V_n\}$-centered toric 4-manifold.

1. Applying this algorithm to $\mathbb{CP}^2$ smoothed in one node yields the self-plumbing of T^*S^2 as illustrated in Fig. 1. Moreover, the same output is obtained for the complement of a toric divisor in *any* toric 4-manifold smoothed in one node.

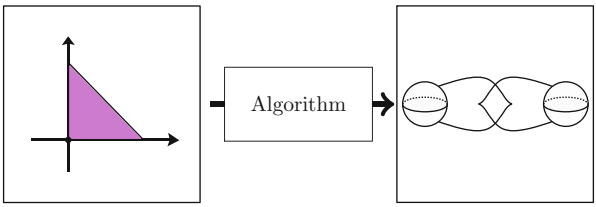


Fig. 1 The input and output diagrams corresponding to the smoothing at one node of the toric divisor in the standard toric $\mathbb{CP}^2$

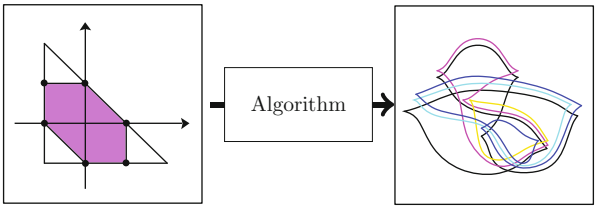


Fig. 2 The input and output diagrams corresponding to the smoothing at six nodes of the toric divisor in a monotone toric $\mathbb{CP}^2\#3\overline{\mathbb{CP}}^2$

2. Applying the algorithm to a monotone $\mathbb{CP}^2\#3\overline{\mathbb{CP}}^2$ smoothed at all six nodes yields a Weinstein manifold constructed by attaching 2-handles to the 4-ball along a 5-component Legendrian link as in Fig. 2.

We would like to note that the handlebody diagram in the first example above has already been observed by Casals-Murphy in [6], viewed as a Weinstein handlebody for the complement of an affine smooth conic in $\mathbb{C}^2$ (which is the same as the complement of a smooth conic together with a generic line in $\mathbb{CP}^2$, which is the once smoothed toric divisor in $\mathbb{CP}^2$). We obtain this diagram from a completely different method and provide a systematic recipe which applies much more generally. This first example provides an accessible way to explain the steps of our more general algorithm.

This paper is organized as follows. In Sect. 2, we give definitions and discuss the relevant preliminary background on Weinstein Kirby calculus. In Sect. 3, we provide a picture into how to see the main structure of the handle attachment needed to obtain the complement of the smoothed divisor by describing the core and co-core of the handle in the smoothing local model. The remainder of the paper demonstrates the algorithm for producing the desired handlebody diagrams. In Sect. 4, we produce a Weinstein handlebody diagram for the complement of a toric divisor smoothed in one node and apply sequences of Kirby calculus moves to simplify the diagrams. In Sect. 5, we present a more complicated example, coming from $\mathbb{CP}^2\#3\overline{\mathbb{CP}}^2$, with the toric divisor smoothed at all six nodes, to showcase the scope of applications and the corresponding Weinstein Kirby diagrams.

2 Weinstein Handlebodies and Kirby Calculus

2.1 Weinstein Handle Structure

A *Liouville vector field* Z for a symplectic manifold (W, ω) is a vector field satisfying $\mathcal{L}_Z\omega = \omega$. By Cartan's formula for the Lie derivative and the fact that the symplectic form is closed, this is equivalent to saying $d(\iota_Z\omega) = \omega$ (here ι denotes the interior product where $\iota_Z\omega(\cdot) = \omega(V, \cdot)$). In particular, when there exists a Liouville vector field, the symplectic structure is exact. The 1-form $\lambda = \iota_Z\omega$ which satisfies $d\lambda = \omega$ is called the *Liouville form*. For an introduction to these ideas, see [23]. The primary use of Liouville vector fields is to glue symplectic manifolds along contact type boundaries. When the Liouville vector field is transverse to the boundary, it defines a contact structure on the boundary and can be used to identify a collared neighborhood of the boundary with a piece of the symplectization of that contact boundary.

When Weinstein defined a model of a handle decomposition for symplectic manifolds [24], he equipped the handle with a Liouville vector field so that the gluing of the handle attachment could be performed using only contact information on the boundary. More specifically, the handle attachment is completely specified by a Legendrian attaching sphere, or an isotropic attaching sphere together with data on its normal bundle. The limitation is that the index of the handle is required to be less than or equal to n in a $2n$ dimensional manifold. In particular, Weinstein 4-manifolds must be built entirely from handles of index 0, 1, and 2.

The model Weinstein handle of index k in dimension $2n$ for $k \leq n$ is a subset of $\mathbb{R}^{2n}$ with coordinates $(x_1, y_1, \cdots, x_n, y_n)$, with the standard symplectic structure $\omega = \sum_j dx_j \wedge dy_j$ and Liouville vector field (Fig. 3)

$$Z_k = \sum_{j=1}^{k} \left(-x_j\partial_{x_j} + 2y_j\partial_{y_j}\right) + \sum_{j=k+1}^{n} \left(\frac{1}{2}x_j\partial_{x_j} + \frac{1}{2}y_j\partial_{y_j}\right).$$

As with smooth handle theory, the handles are in one to one correspondence with the critical points of a Morse function. The Liouville vector field agrees with the gradient of such a Morse function (for some choice of metric), in other words, the Liouville vector field is *gradient-like*. In the model index k handle, the Liouville vector field is the gradient (with the standard Euclidean metric) of the function

$$\phi_k = \sum_{j=1}^{k} \left(-\frac{1}{2}x_j^2 + y_j^2\right) + \sum_{j=k+1}^{n} \left(\frac{1}{4}x_j^2 + \frac{1}{4}y_j^2\right).$$

The handle can be considered to be the subset of $\mathbb{R}^{2n}$ given by $D^k \times D^{2n-k}$ where the first factor corresponds to the coordinates $(x_1, \cdots, x_k)$ and the second

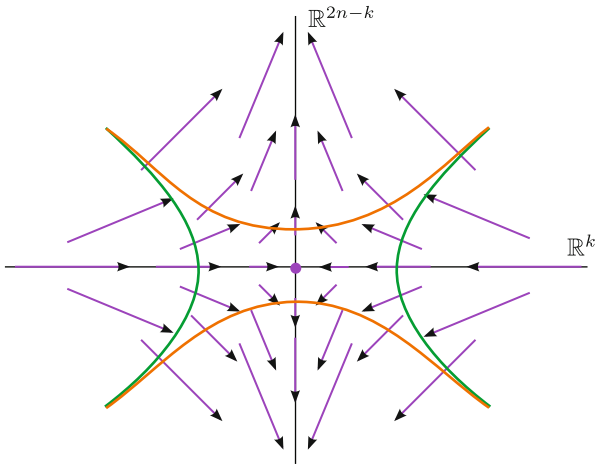


Fig. 3 A sketch of the model for the Weinstein handle of index k, with the associated gradient-like Liouville vector field Z_k

corresponds to the remaining coordinates $(x_{k+1}, \cdots, x_n, y_1, \cdots, y_n)$. The key terminology for important parts of the handle is as follows.

- The *core* of the handle is $D^k \times \{0\}$ where $x_{k+1} = \cdots = x_n = y_1 = \cdots = y_n = 0$. This is the *stable manifold* of flow-lines of Z_k which limit positively towards the zero at the origin.
- The *co-core* of the handle is $\{0\} \times D^{2n-k}$ where $x_1 = \cdots = x_k = 0$. This is the *unstable manifold* of flow-lines of Z_k which limit negatively towards the zero at the origin.
- The *attaching sphere* is the boundary of the core, $S^{k-1} \times \{0\}$. This will be identified with an isotropic sphere in the boundary of the existing manifold to which the handle is attached.
- The *attaching region* is a neighborhood of the attaching sphere $S^{k-1} \times D^{2n-k}$. This is the entire part of the handle which will be glued on to a piece of the boundary of the existing manifold when the handle is attached. Therefore the Liouville vector field Z_k points inward into the handle along this part of the boundary (it is *concave*).
- The *belt sphere* is the boundary of the co-core $\{0\} \times S^{2n-k-1}$. It is a co-isotropic sphere in the boundary of the manifold obtained after attaching the handle.

In general we can piece together the Liouville vector fields on the handles, and put together adjusted versions of the locally defined Morse functions to get a global Morse function on the manifold. A Weinstein structure is often encoded analytically as a quadruple (W, ω, Z, ϕ) where W is a smooth manifold, ω is a symplectic structure on W, Z is a Liouville vector field for ω on W, and ϕ is a Morse function such that Z is the gradient-like for ϕ.

Remark 2 When a manifold with a Weinstein structure has contact type boundary, it is called a *Weinstein domain*. Such a domain can be extended by a cylindrical end to make the Liouville vector field complete to give a non-compact infinite volume *Weinstein manifold*.

2.2 Weinstein Kirby Calculus

The data needed to encode the Weinstein domain are the attaching maps. In dimension 4, the attaching map of a handle is completely determined by the Legendrian or isotropic attaching sphere. The attaching sphere of a 1-handle is a pair of points. Diagramatically we draw a pair of 3-balls implicitly identified by a reflection, representing the attaching region $S^0 \times D^3$. The attaching sphere of a 2-handle is a Legendrian embedded circle (a Legendrian knot). The 4-dimensional 2-handle attachment is determined by the knot together with a framing, but in the Weinstein case, the framing is determined by the contact structure. More specifically, the contact planes along a Legendrian knot determine a framing by taking a vector field transverse to the contact planes. The contactomorphism gluing the attaching region of the 2-handle to the neighborhood of the Legendrian identifies the product framing in the 2-handle with the $tb - 1$ framing. Here tb denotes the contact framing (Thurston-Bennequin number), which is identified with an integer by looking at the difference between the contact framing and the Seifert framing (this must be appropriately interpreted when the diagram contains 1-handles–see [19]).

The diagram we draw should specify the Legendrian attaching knots in S^3 along with the pairs of 3-balls indicating the attachments of the 1-handles. By removing a point away from these attachments, we reduce the picture in S^3 to a picture in $\mathbb{R}^3$. After a contactomorphism, the contact structure on $\mathbb{R}^3$ is $\ker(dz - ydx)$ in coordinates (x, y, z). The *front projection* is the map $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ with $\pi(x, y, z) = (x, z)$. A Legendrian curve in this contact structure is tangent to the contact planes, which happens precisely when the y-coordinate is equal to the slope $\frac{dz}{dx}$ of the front projection. Therefore, Legendrian knots can be recovered from their front projections with the requirement that the diagram has no vertical tangencies (instead it will have cusp singularities where the knot is tangent to the fibers of the projection) and the crossings are always resolved so that the over-strand is the strand with the more negative slope (we orient the y-axis into the page to maintain the standard orientation convention for $\mathbb{R}^3$ so the over-strand is the strand with a more negative y-coordinate). In these front projections, the contact framing tb can be computed combinatorially in terms of the oriented crossings and cusps of the diagram when the diagram is placed in a standard form where the pairs of 3-balls giving the attaching regions of 1-handles are related by a reflection across a vertical axis. Namely, tb of a Legendrian knot is the difference of the writhe of the knot and half the number of cusps in the front projection. For a thorough introduction to Legendrian knots, see [15].

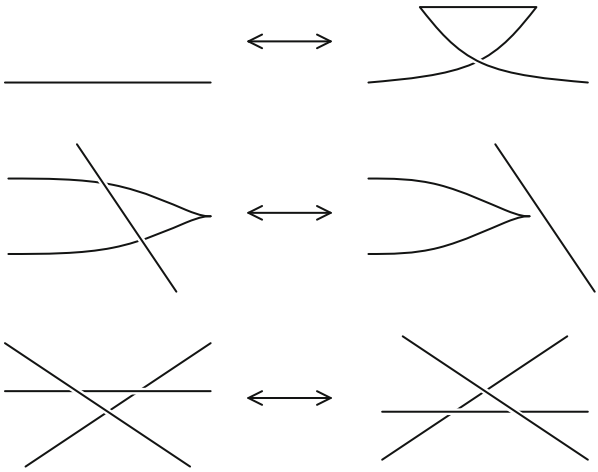


Fig. 4 The Legendrian Reidemeister moves, up to 180° rotation about each axis, where the top, middle, and bottom moves are called Reidemeister I, Reidemeister II, and Reidemeister III, respectively

The set of moves that relate Weinstein handlebody diagrams in Gompf standard form for equivalent Weinstein domains includes Legendrian Reidemeister moves (including how they interact with the 1-handles) listed in [19], see Figs. 4 and 5, as well as *handle slides*, and *handle pair cancellations and additions*, shown in [7]. Given two k-handles, h_1 and h_2, a *handle slide* of h_1 over h_2 is given by isotoping the attaching sphere of h_1, and pushing it through the belt sphere of h_2. We depict a 1-handle slide (along with intermediate Reidemeister and Gompf moves) in Fig. 6 and a 2-handle slide in Fig. 7. A 1-handle h_1 and a 2-handle h_2 can be cancelled, provided that the attaching sphere of h_2 intersects the belt sphere of h_1 transversely in a single point. We call this a *handle cancellation* and the pair of handles a *cancelling pair*. Likewise a cancelling pair can be added to a Weinstein handlebody diagram, as depicted in Fig. 8. When multiple 2-handles intersect a single 1-handle, the simplification in Fig. 9 can be performed to reduce the overall complexity of a Weinstein diagram.

Before approaching our goal of presenting an algorithm to construct Weinstein–Kirby diagrams for complements of smoothed toric divisors, we will start with the unsmoothed case, where the complement has a Liouville completion, T^*T^2. The Legendrian handlebody we present, was originally found by Gompf [19]. It follows from that article that this handlebody gives a Stein/Weinstein structure on the smooth manifold D^*T^2 (which is the trivial bundle $D^2 \times T^2$). More generally, Stein handlebody diagrams are given on the smooth manifolds $D^*\Sigma$ for any surface in [19]. In [2, Theorem 7.1], we show that the Weinstein structures induced on these diagrams are Weinstein homotopic to the canonical co-tangent Weinstein structure on $D^*\Sigma$.

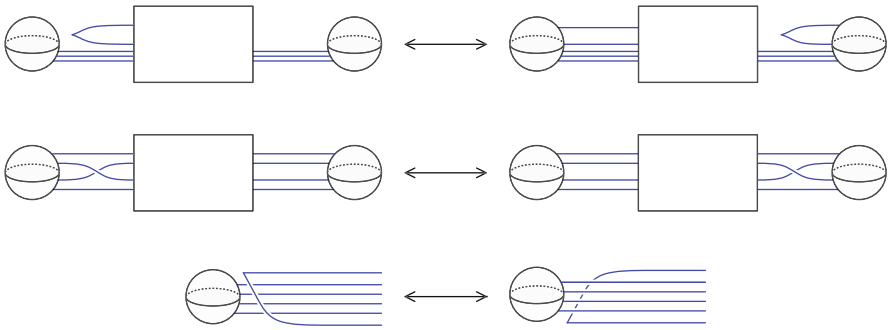


Fig. 5 Gompf's three additional isotopic moves, up to 180° rotation about each axis. The top, middle, and bottom moves are called Gompf move 4, Gompf move 5, and Gompf move 6, respectively

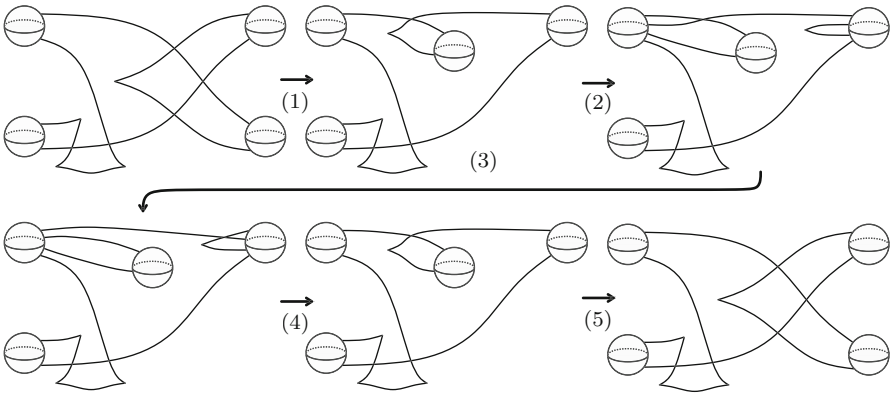


Fig. 6 An example of a 1-handle slide on T^*T^2 consisting of the following sequence of isotopies: (1) Reidemeister 2, (2) Gompf move 4, (3) Gompf move 5, (4) Gompf move 4, (5) Reidemeister 2

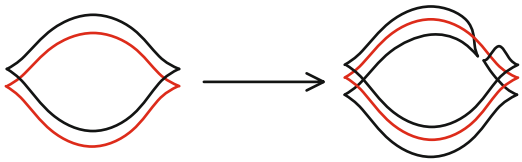


Fig. 7 An example of a 2-handle slide of the black unknot over the red unknot

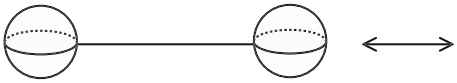


Fig. 8 An example of a 1-handle cancelling with a 2-handle

For T^*T^2 specifically, it is known that there is a unique Weinstein fillable contact structure on the boundary T^3 and one can then deduce that the Gompf handlebody agrees with the canonical symplectic structure by Wendl's result that S^*T^2 has a unique Stein/Weinstein filling up to deformation [25]. To see that the diagram

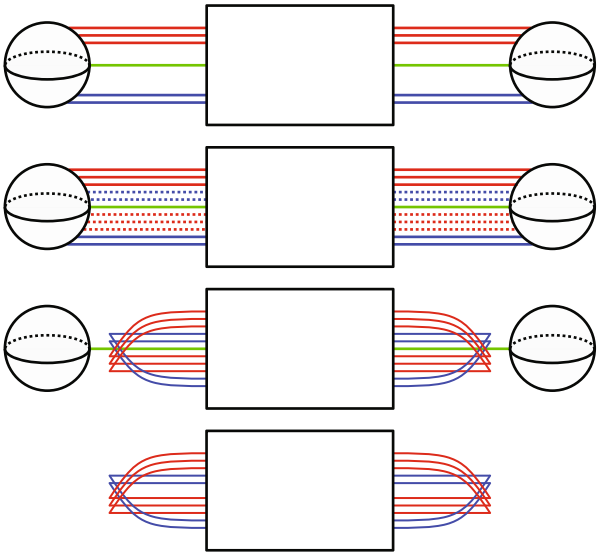


Fig. 9 An example of handle slides and cancellations when multiple 2-handles pass through a 1-handle. Red and blue 2-handles are slid over the central green 2-handle. The green 2-handle is then cancelled with the 1-handle

in Fig. 11 represents $D^*T^2 \cong D^2 \times T^2$ smoothly, we can start with a handle decomposition for T^2 with one 0-handle, two 1-handles and a single 2-handle. Thickening this diagram to a 4-dimensional handlebody yields a disk bundle over T^2 with Euler number e, agreeing with the framing coefficient of the 2-handle attachment. One then needs to put the diagram into Gompf standard form as seen in Fig. 10 by sliding the uppermost attaching ball below the attaching ball on the right so that both 1-handles are related by a reflection across the same vertical axis. Then we must realize the knot as a Legendrian knot by replacing vertical tangencies by cusps and making sure the crossings always have the over-strand corresponding to the more negative slope. The most obvious way to do this yields a Legendrian knot whose Thurston-Bennequin framing is 0 (see the diagram on the right of Fig. 10), so this would correspond to a D^2 bundle over T^2 with Euler number -1. By wrapping one strand around the lower left attaching ball as in Fig. 11, we obtain a smoothly isotopic picture where the new Legendrian has $tb = 1$, so the Euler number is $1 - 1 = 0$ as needed for D^*T^2. Since this is one Weinstein filling of S^*T^2, and we know that such fillings are unique up to deformation, it must agree with the canonical co-tangent symplectic structure on D^*T^2.

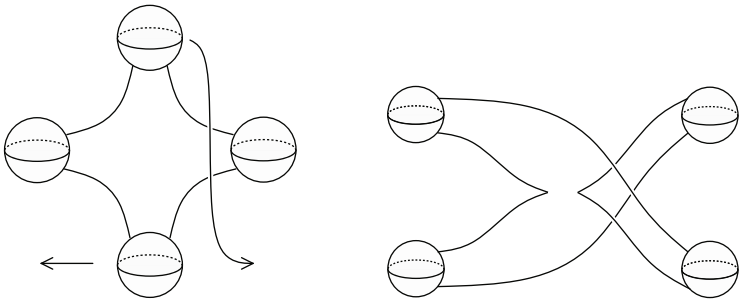


Fig. 10 Diagram depicting how to move the usual picture of T^*T^2 into standard form after inserting the necessary Legendrian data, i.e. replacing vertical tangencies by cusps

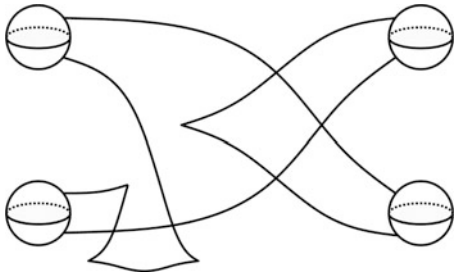


Fig. 11 Stein structure on D^2 bundle over T^2 with framing coefficient $e(T^*T^2) \leq 0$

3 The Local Model for Our Handle Attachment

In this section we study the local model for the smoothing of a normal crossing singularity of a symplectic divisor in dimension 4 and describe the local handle attachment information (core and co-core) for the 2-handle attachment needed to describe the complement of the smoothing of a node. Here we give the intuitive picture behind the following theorem.

Theorem 3 (Theorem 4.1 [2]) *Let (M, ω) be a toric 4-manifold corresponding to Delzant polytope Δ which is $\{V_1, \ldots, V_n\}$-centered. Let D denote the divisor obtained by smoothing the toric divisor at the nodes $V_1, \ldots, V_n$. Then there exist arbitrarily small neighborhoods N of D such that $M \setminus N$ admits the structure of a Weinstein domain.*

Furthermore, $M \setminus N$ is Weinstein homotopic to the Weinstein domain obtained by attaching Weinstein 2-handles to the unit disk cotangent bundle of the torus D^*T^2, along the Legendrian co-normal lifts of co-oriented curves of slope $s(V_1), \ldots, s(V_n)$. Here $s(V_i)$ is equal to the difference of the inward normal vectors of the edges adjacent to V_i in Δ.

A local model in a 4-dimensional manifold M for the normal crossing intersection of two symplectic divisors can be given by a Darboux chart $(\mathbb{C}^2, \omega_{std})$ at the intersection point, where the two divisors are mapped to the two axes

$\Sigma_1 = \{(z_1, z_2) \in \mathbb{C}^2 \mid z_2 = 0\}$ and $\Sigma_2 = \{(z_1, z_2) \in \mathbb{C}^2 \mid z_1 = 0\}$. Smoothing this normal crossing means that locally one substitutes the union of these divisors by the smooth surface

$$\Sigma = \left\{(z_1, z_2) \in \mathbb{C}^2 \mid z_1 \cdot z_2 = \varepsilon^2\right\}$$

for some $\varepsilon > 0$. Topologically, the complement of the smoothed surface differs from the complement of the normal crossing divisors by one 2-handle attachment for each smoothed normal crossing intersection. Under the centered hypothesis (Definition 1), Theorem 3 says that when considering the symplectic and Liouville structures, the difference between the complements also corresponds to a collection of *Weinstein* 2-handle attachments.

In order to encode this handle attachment in the Weinstein Kirby diagram, we need to identify the corresponding Legendrian attaching sphere. We find natural Lagrangian representatives of the co-core and core of the handle in the complement of the smoothed divisor. From there, the attaching sphere is the boundary of the core.

A co-core of a Weinstein 2-handle is characterized (uniquely up to Lagrangian isotopy [14]) as a Lagrangian disk contained in the smooth 2-handle with unknotted Legendrian boundary lying in the boundary of the handle (not the attaching region). In our model, the boundary of the handle is in the boundary of the small neighborhood of the smoothed hypersurface Σ. One can consider, for $0 < \varepsilon' < \varepsilon$, the disk D_1 defined as the image of the map

$$\begin{array}{rcl} \phi : [0, 1] \times [0, 2\pi] \to & \mathbb{C}^2 \\ (r, \theta) \quad \mapsto & \begin{pmatrix} r\varepsilon' e^{i\theta} \\ r\varepsilon' e^{-i\theta} \end{pmatrix} \end{array}$$

The size of ε' is determined by the size of the neighborhood of the smoothed divisor which is deleted. As $\varepsilon' \to \varepsilon$, the open disk of radius ε provides the analog of the co-core in the non-compact complement of the divisor itself.

This disk is Lagrangian as for $r \neq 0$, the derivative of ϕ

$$d_{(r,\theta)}\phi = \begin{pmatrix} \varepsilon' e^{i\theta} & ir\varepsilon' e^{i\theta} \\ \varepsilon' e^{-i\theta} & -ir\varepsilon' e^{-i\theta} \end{pmatrix}$$

is an isomorphism so that the tangent space to D_1 at the point $\phi(r, \theta)$ is spanned by the vectors

$$u_1 = \begin{pmatrix} \varepsilon' e^{i\theta} \\ \varepsilon' e^{-i\theta} \end{pmatrix} \text{ and } u_2 = \begin{pmatrix} ir\varepsilon' e^{i\theta} \\ -ir\varepsilon' e^{-i\theta} \end{pmatrix}$$

and one can check that

$$\omega_{std}(u_1, u_2) = \Im m(\langle u_1, u_2 \rangle) = 0$$

where $\langle \cdot, \cdot \rangle$ is the standard Hermitian product in $\mathbb{C}^2$ (equivalently one can check that $\phi^*(\frac{i}{2}(dz_1 \wedge d\bar{z}_1 + dz_2 \wedge d\bar{z}_2)) = 0$). The point for $r = 0$ corresponds to the origin of $\mathbb{C}^2$. At this point, the two following curves c_1 and c_2 in D_1 parametrized for $t \in (-1, 1)$ by:

$$c_1(t) = \begin{pmatrix} t\varepsilon' \\ t\varepsilon' \end{pmatrix} \text{ and } c_2(t) = \begin{pmatrix} it\varepsilon' \\ -it\varepsilon' \end{pmatrix}$$

give the two independent vectors in the tangent space at the origin of the disk D_1:

$$u_1 = \begin{pmatrix} \varepsilon' \\ \varepsilon' \end{pmatrix} \text{ and } u_2 = \begin{pmatrix} i\varepsilon' \\ -i\varepsilon' \end{pmatrix}$$

One can note that we have again $\omega_{std}(u_1, u_2) = 0$.

The boundary of this disk is the image by ϕ of $\{1\} \times [0, 2\pi]$, that is, the circle

$$B = \left\{ \begin{pmatrix} \varepsilon' e^{i\theta} \\ \varepsilon' e^{-i\theta} \end{pmatrix} \middle| \, \theta \in [0, 2\pi] \right\}.$$

One can choose $\varepsilon' < \varepsilon$ such that the circle B lies on the boundary of the small neighborhood of the smoothed Σ (since $z_1 z_2 = \varepsilon' e^{i\theta} \varepsilon' e^{-i\theta} = \varepsilon'^2$ goes to ε^2 if ε' approaches ε). It limits to the origin when ε' goes to 0, so that B is the belt sphere of the handle and D_1 is indeed its co-core.

The core is characterized (uniquely up to Lagrangian isotopy) as a Lagrangian disk with unknotted boundary in the smooth 2-handle which intersects the co-core transversally at one point and which avoids the boundary of the Weinstein manifold, so in our model, it should avoid the smoothed Σ.

Let D_2 be the disk defined as the image of the map

$$\begin{array}{rcl} \psi : [0, 1] \times [0, 2\pi] \to & & \mathbb{C}^2 \\ (r, \theta) & \mapsto & \begin{pmatrix} r\varepsilon e^{i\theta} \\ r\varepsilon e^{-i(\theta+\pi)} \end{pmatrix} = \begin{pmatrix} r\varepsilon e^{i\theta} \\ -r\varepsilon e^{-i\theta} \end{pmatrix} \end{array}$$

This disk is also Lagrangian as, similarly as before, the tangent space to D_2 at the point $\psi(r, \theta)$ for $r \neq 0$ is spanned by the vectors

$$u_3 = \begin{pmatrix} \varepsilon e^{i\theta} \\ -\varepsilon e^{-i\theta} \end{pmatrix} \text{ and } u_4 = \begin{pmatrix} ir\varepsilon e^{i\theta} \\ ir\varepsilon e^{-i\theta} \end{pmatrix}$$

and one can check that

$$\omega_{std}(u_3, u_4) = \Im m(\langle u_3, u_4 \rangle) = 0.$$

Similarly, at the origin, the two following curves c_3 and c_4 in D_2 parametrized for $t \in (-1, 1)$ by

$$c_3(t) = \begin{pmatrix} t\varepsilon \\ -t\varepsilon \end{pmatrix} \text{ and } c_4(t) = \begin{pmatrix} it\varepsilon \\ it\varepsilon \end{pmatrix}$$

give the two independent vectors in the tangent space at the origin of the disk D_2:

$$u_3 = \begin{pmatrix} \varepsilon \\ -\varepsilon \end{pmatrix} \text{ and } u_4 = \begin{pmatrix} i\varepsilon \\ i\varepsilon \end{pmatrix}$$

Note again that $\omega_{std}(u_3, u_4) = 0$. This disk does not intersect the smoothed Σ as

$$z_1 z_2 = r\varepsilon e^{i\theta}(-r\varepsilon e^{-i\theta}) = -r^2\varepsilon^2 \neq \varepsilon^2.$$

Moreover, D_1 and D_2 intersect at the origin and this intersection is transverse as one can check that the family (u_1, u_2, u_3, u_4) spans $\mathbb{C}^2$ as a real vector space. This shows that D_2 is the core of the handle and the attaching sphere is the image by ψ of $\{1\} \times [0, 2\pi]$, that is,

$$A = \left\{ \begin{pmatrix} \varepsilon e^{i\theta} \\ \varepsilon e^{-i(\theta+\pi)} \end{pmatrix} \middle| \, \theta \in [0, 2\pi] \right\}.$$

Identifying these transverse Lagrangian disks with the core and co-core in Weinstein's model for a 4-dimensional 2-handle, allows us to place a Weinstein structure (a Liouville vector field and corresponding Morse function) on this new piece which lies in the complement of the regular neighborhood of the smoothing. If the polytope is centered with respect to all the nodes we are smoothing, the Weinstein structures on these pieces glue together consistently with a Weinstein structure on the complement of the unsmoothed toric divisor (a Weinstein domain which completes to T^*T^2). It is shown in [2, Section 4.5] how these fit together to give an explicit global Weinstein structure.

In the toric description of the toric manifolds and divisors we consider, the Hamiltonian torus action in the local Darboux model corresponds to the torus action on $\mathbb{C}^2$ given in coordinates: $(e^{i\theta_1}, e^{i\theta_2}) * (z_1, z_2) = (e^{i\theta_1} z_1, e^{i\theta_2} z_2)$. In particular, through the symplectomorphism between the complement of the normal crossing divisor we consider and T^*T^2, the orbit of a point corresponds to the torus T^2 and the quotient space under the Hamiltonian action corresponds to a cotangent fiber. The attaching sphere in the model corresponds in this symplectic identification to a lift of a circle of slope $(1, -1)$ in the base T^2 to the cotangent bundle T^*T^2 (lift corresponding to the point $(\varepsilon^2, \varepsilon^2)$ in the quotient space $(\mathbb{R}_{>0})^2$). The standard model neighborhood corresponds to the standard cone $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, and the corresponding attaching slope is $(1, -1) = (1, 0) - (0, 1)$, the difference of the inward normal vectors. In general, relating any node to the standard model

via an $SL(2,\mathbb{Z})$ transformation, the attaching sphere for the 2-handle corresponding to the smoothing of a normal crossing singularity at a chosen vertex V will be the co-normal lift of a circle of slope $s(V)$, where $s(V)$ is the difference of the inward normal vectors of the edges adjacent to V, (see [2, Lemma 4.2]).

4 The Algorithm Through an Example

Here we will show how to obtain a Weinstein handle diagram for the complement of a toric divisor with exactly one node smoothed. Note that for any Delzant polytope, and any single vertex V, the polytope is vacuously $\{V\}$-centered. To fix a specific example, we could start with the toric 4-manifold $\mathbb{CP}^2$ whose toric divisor is a collection of three $\mathbb{CP}^1$'s intersecting at three points, and smooth this divisor at one of the intersection points. More explicitly in homogeneous coordinates $[z_0 : z_1 : z_2]$ on $\mathbb{CP}^2$, the toric divisor is given by the union of the three lines $L_0 = \{z_0 = 0\}$, $L_1 = \{z_1 = 0\}$ and $L_2 = \{z_2 = 0\}$. Let us smooth the intersection of L_1 with L_2 at $[1 : 0 : 0]$. In the affine coordinate chart where $z_0 = 1$, with coordinates (z_1, z_2), this aligns exactly with our local model in Sect. 3. After smoothing, the lines L_1 and L_2 are joined to form a conic Q, which intersects the remaining line L_0 at two points. By Theorem 3, the complement of this smoothed divisor is obtained by attaching a single 2-handle to D^*T^2 along the Legendrian lift of a curve in T^2 of slope $(1, -1)$.

In order to translate Legendrian attaching circles in S^*T^2 described as the co-normal lift of a curve in T^2 into Legendrian curves drawn in the Gompf diagram (Fig. 11), we need to understand how these two pictures get identified. As mentioned in Sect. 2.2, the Gompf handle diagram is obtained by starting with a smooth handle decomposition of T^2 with a single 0-handle, two 1-handles, and one 2-handle. This diagram is thickened by two dimensions to obtain $T^2 \times D^2$, and then the attaching curve of the 2-handle is isotoped around until it agrees with a Legendrian front diagram with induced framing $tb - 1 = 0$. On the other hand, the co-normal lift construction is more compatible with the canonical (Morse-Bott) Weinstein structure on D^*T^2 which has critical locus along the 0-section. In [2, Theorem 7.1], we prove that these two structures are Weinstein homotopic and identify the image of the Lagrangian torus giving the zero-section of D^*T^2 in the Gompf diagram. The handle decomposition on the 4-manifold induces the corresponding handle decomposition on the Lagrangian torus by intersection. In particular, we see a Legendrian (un)knot $\mathcal{K}$ in the boundary of the 4-dimensional 0-handle, which partially coincides with the attaching sphere of the 2-handle, and partially corresponds with attaching arcs for the 1-handles of the Legendrian torus. See Fig. 12.

Now consider the Legendrian co-normal lift to S^*T^2 of the circle in T^2 which is the boundary of the 2-dimensional 0-handle, with the inward co-orientation. This is a Legendrian push-off of $\mathcal{K}$ in $\partial B^4 = \partial D^*D^2$, because a small positive Reeb flow applied to $\mathcal{K} = \partial D^2$ yields the co-normal lift of a concentric circle close to ∂D^2. We will perform an isotopy to the curves in our torus corresponding to attaching

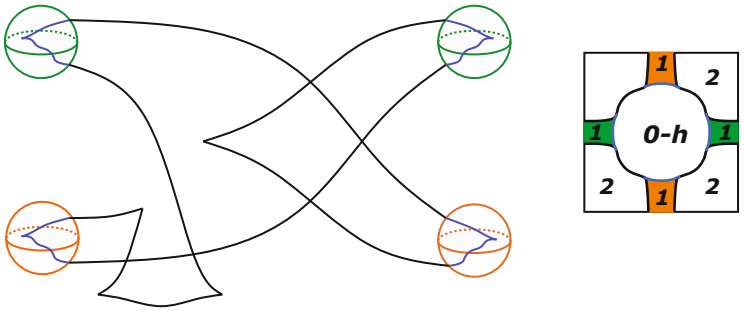


Fig. 12 Left: The Legendrian unknot $\mathcal{K}$ in the boundary of the 4-dimensional 0-handle of the Gompf diagram, indicating the intersection of this boundary with the Lagrangian torus (0-section of D^*T^2). The black portions coincide with segments of the attaching circle of the 2-handle, and the blue portions give the attaching arcs of the 2-dimensional 1-handles of the torus. Right: The corresponding decomposition of the Lagrangian torus

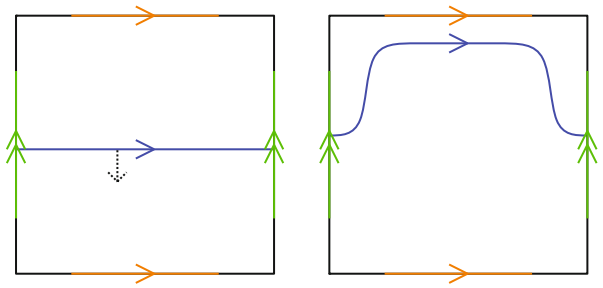


Fig. 13 The circle with $(1, 0)$-slope on T^2 with the dotted arrow indicating the co-orientation of the curve, and an isotoped version which runs parallel to the boundary of the 0-handle of T^2 except where it passes through a 1-handle

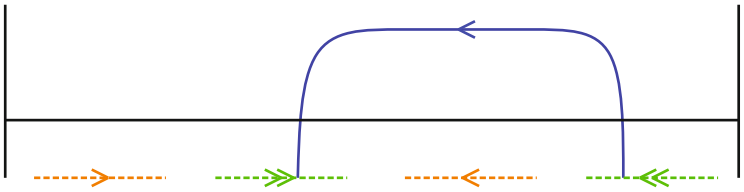


Fig. 14 The curve in $J^1(S^1)$ which is identified with the curve $(1, 0)$ on T^2 as in Fig. 13

spheres of the additional 2-handles so that these curves agree with parallel copies of such circles except where they enter the 1-handles. Circles which are further inward will be pushed off more in the *positive* Reeb direction. Note that every Legendrian circle has a standard neighborhood which is contactomorphic to a neighborhood of the zero section in $J^1(S^1)$ with the contact form $dz - ydx$ where x is the coordinate on S^1. We will translate the diagram on the torus to a front projection diagram of $J^1(S^1)$, where we think of the S^1 as the Legendrian boundary of the 0-handle of

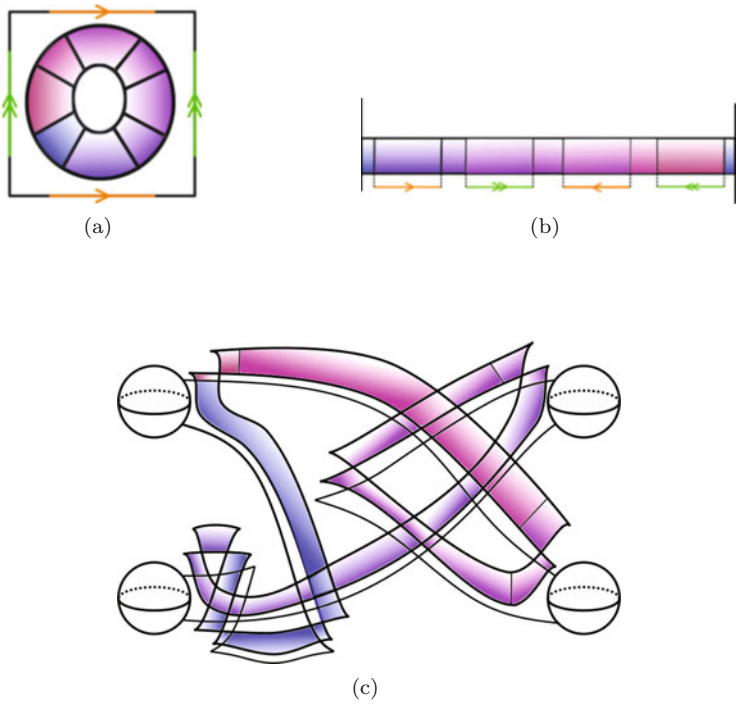


Fig. 15 Transferring curves from T^2 to $J^1(S^1)$ and then to the Gompf diagram. (**a**) The neighborhood U of the co-normal lift of a curve $\gamma \subset T^2$. (**b**) Identification of U with neighborhood of the zero section of $J^1(S^1)$. (**c**) The front projection of U in the Gompf diagram for D^*T^2

the T^2 0-section. Since the Reeb direction is the positive ∂_z direction in $J^1(S^1)$, circles which are further inwards in the torus (pushed further by the Reeb flow) will correspond to curves which are pushed upwards more in the $J^1(S^1)$ diagram. See an example of this procedure in Figs. 13 and 14. Finally, once we have our diagram in the 1-jet space, we can satellite the diagram onto the image of S^1 in the Gompf diagram. The images of the co-normal lifts of curves contained in a neighborhood U of the curve γ that coincides with the attaching circle of the 2-handle and the attaching arcs of the 2 dimensional 1 handles of the torus are illustrated in Fig. 15. As indicated by the shading gradient, curves which lie further inward towards the center of the square correspond to curves at greater z-height values in the jet-space. In turn, curves at a higher z-height in the jet space will correspond to higher Reeb pushoffs of the complicated looking Legendrian unknot of Fig. 15c.

Initially, let us apply this procedure in the simple example where we are attaching a 2-handle along the co-normal lift of a circle in the torus with slope $(1, 0)$. The cotangent projection of this model is presented on the left in Fig. 13. Push the curve $(1, 0)$ to the upper side of the square, so it lies close to the boundary, and then cut the

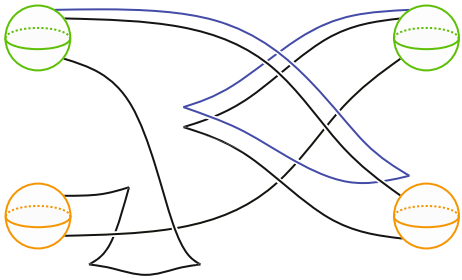


Fig. 16 The Legendrian handle diagram of the complement of the toric divisor smoothed in one node. That is, $T^*T^2 \cup \Lambda_{(1,0)}$

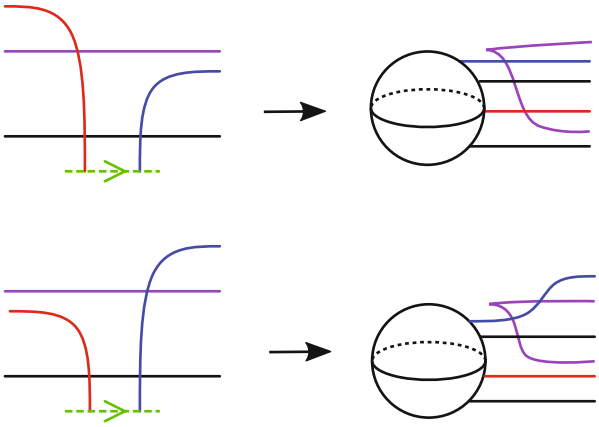


Fig. 17 Mapping the red, blue and purple curves from $J^1(S^1)$ to T^*T^2. The relative heights of these curves with respect to the black curve are preserved. The purple curve follows a Reeb push off of the cusping curve in the 1-handle's attaching region (see Fig. 12) while the red and blue curves pass through the 1-handle

rectangle at the bottom left vertex to map to $J^1(S^1)$, obtaining Fig. 14. Satelliting this onto the Legendrian unknot in Fig. 12, we obtain Fig. 16.

In general, when satelliting, we need to be somewhat careful with the behavior of the curves near the 1-handle. If the curves pass above the attaching region of a 1-handle without entering the 1-handle, they will follow an upward Reeb push-off of the cusps that appear inside the 1-handle attaching balls in Fig. 12. Note that we will typically push these cusps out of the attaching regions of the 1-handles by a Legendrian isotopy. If the curves pass through the 1-handle in the torus, they will pass through the corresponding 1-handle in the 4-dimensional handlebody. See Fig. 17 for the conventions in a more complicated example.

Although our initial explicit example asked us to attach a handle along a curve of slope $(1, -1)$, we can see that in fact the resulting Weinstein manifold is equivalent to using the $(1, 0)$ slope. Figure 18 shows the series of 1 handle slides, Reidemeister and Gompf moves that take the Legendrian lift of the resulting diagram obtained by attaching along a $(1, -1)$ curve to the diagram corresponding to attaching along a $(1, 0)$ curve.

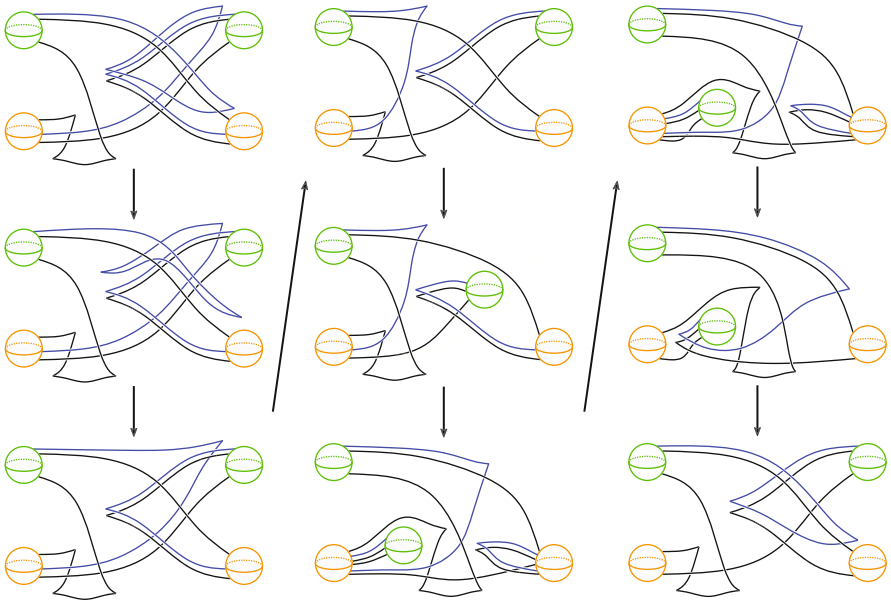


Fig. 18 In columns from left to right, a series of Reidemeister and Gompf moves and 1-handle slides that take the Legendrian lift of a $(1, -1)$ curve to the Legendrian lift of a $(1, 0)$ curve. (1) Reidemeister III, (2) Reidemeister II and I, (3) Reidemeister II, (4) Reidemeister II, (5) slide green 1-handle over orange 1-handle, (6) Gompf move 5, (7) Gompf move 4, and (8) Reidemeister II

Proposition 4 *For any toric* 4*-manifold, let* Σ *be the result of smoothing the toric divisor at a single node. Then the complement of a small regular neighborhood of* Σ *is a Weinstein domain which is Weinstein homotopic to the domain represented by the handle diagram of Fig. 16.*

Proof First note that for a single node, the centered hypothesis is automatically satisfied. By Theorem 3 such a complement is a Weinstein domain obtained by attaching a single 2-handle to D^*T^2 along the Legendrian curve given by a co-oriented co-normal lift of a curve of slope (a, b) in the torus. We verify that the resulting Weinstein domain does not depend on the choice of the slope up to Weinstein homotopy. We can see that $T^*T^2 \cup \Lambda_{(a,b)}$ is Weinstein homotopic to $T^*T^2 \cup \Lambda_{(1,0)}$, by performing 1-handle slides on $T^*T^2 \cup \Lambda_{(a,b)}$ similar to Fig. 18, to take a $\Lambda_{(a,b)}$ curve to $\Lambda_{(a,b\pm a)}$ or to $\Lambda_{(a\pm b,b)}$. Using the Euclidean algorithm, with an appropriate choice of 1-handle slides, one can start with $T^*T^2 \cup \Lambda_{(a,b)}$, for any pair $a, b \in \mathbb{Z}$ that are relatively prime, and end with $T^*T^2 \cup \Lambda_{(1,0)}$. That the Weinstein domains are symplectomorphic can also be proved using toric arguments (see [2, Proposition 5.5]). □

The diagram where we attach along the $(1, 0)$ curve is preferable to the one where we attach along slope $(1, -1)$ (or a more complicated (a, b) curve) because it is easier to simplify. Starting with Fig. 16, we can perform Reidemeister moves,

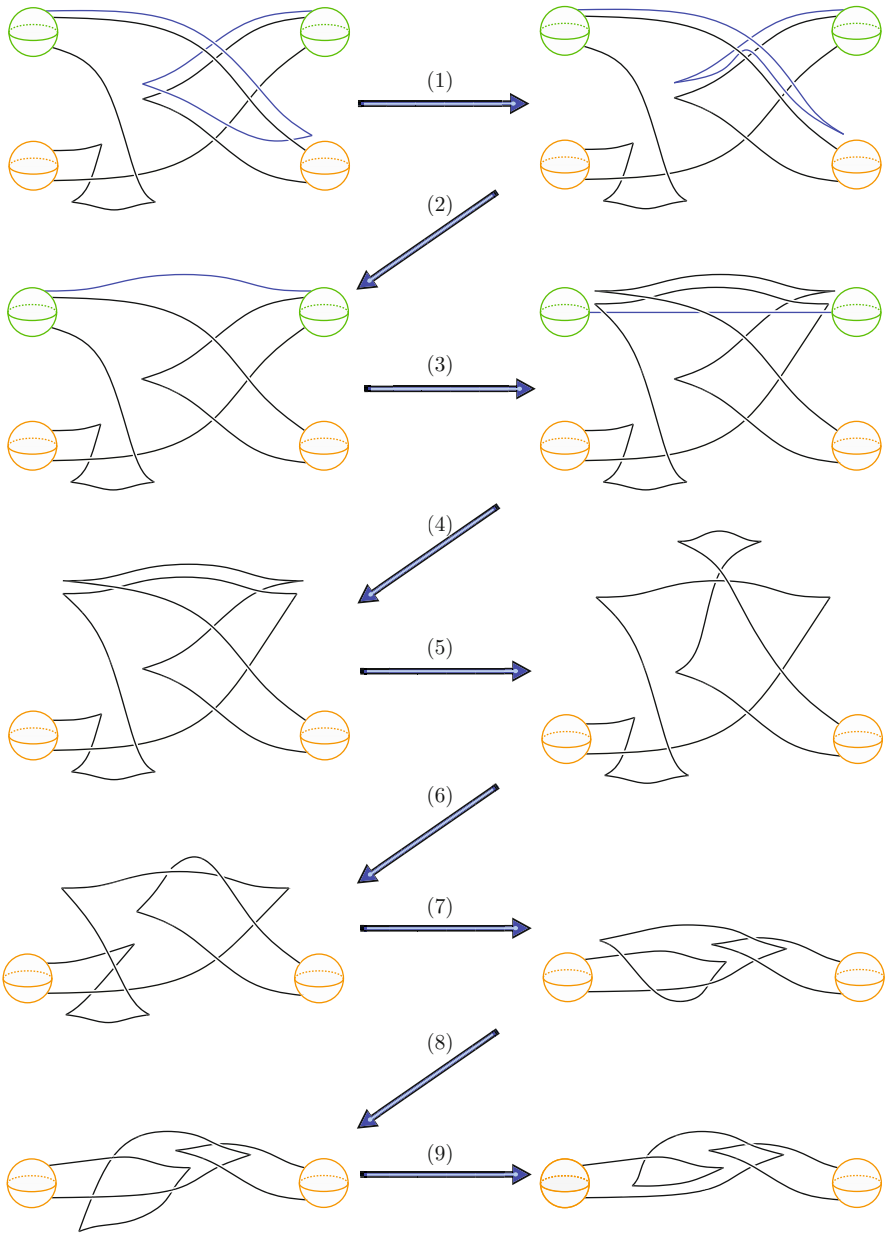


Fig. 19 A series of Reidemeister moves and handle slides simplifying the Weinstein handle diagram of the complement of the toric divisor smoothed in one node. (1) Reidemeister III, (2) Reidemeister II and I, (3) 2-handle slide, (4) Handle cancellation, (5) Reidemeister III, (6) Reidemeister I and II, (7) Reidemeister III, I and II, (8) Gompf move 6, (9) Reidemeister II

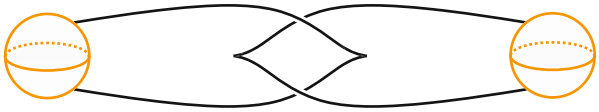


Fig. 20 The diagram of the complement of the toric divisor smoothed in one node after simplifications, in particular after applying a single Reidemeister I to the last diagram in Fig. 19

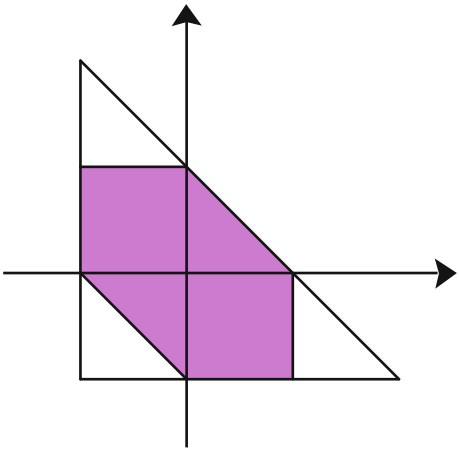


Fig. 21 The Delzant polytope of monotone $\mathbb{CP}^2\#3\overline{\mathbb{CP}}^2$

Gompf moves, handle cancellations and handle slides. We choose such a simplifying sequence in Fig. 19 to obtain the diagram illustrated in Fig. 20.

5 A More Complicated Example: Smoothing a Toric Divisor in $\mathbb{CP}^2\#3\overline{\mathbb{CP}}^2$

Consider $\mathbb{CP}^2$ blown up three times with the same size of the blow ups. The corresponding Delzant polytope is a hexagon illustrated in Fig. 21. Observe that the sizes of the blow-ups have been chosen precisely to make the polytope centered with respect to all of its vertices. The inward normals of the six corners are: $(1, 0)$, $(0, -1)$, $(-1, -1)$, $(-1, 0)$, $(0, 1)$, $(1, 1)$. By Theorem 3, if we smooth all six singularities that map under the moment map to the vertices of the hexagon polytope, the complement of the smoothed divisor is T^*T^2 with 2-handles attached along

$$\Lambda_{(1,0)}, \Lambda_{(0,-1)}, \Lambda_{(-1,-1)}, \Lambda_{(-1,0)}, \Lambda_{(0,1)}, \Lambda_{(1,1)}$$

as shown in Fig. 22.

Figure 22 shows the curves $\Gamma = (1, 0), (0, -1), (-1, -1), (-1, 0), (0, 1), (1, 1)$ on T^2. As in the previous section, we isotope the curves giving a Legendrian isotopy

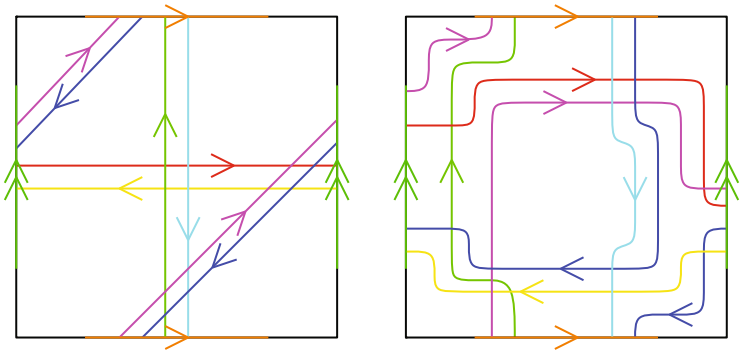


Fig. 22 The curves $(1, 0), (0, -1), (-1, -1), (-1, 0), (0, 1), (1, 1)$ on T^2 before and after isotoping them to agree with concentric circles co-oriented inward (except where they pass through the 1-handles)

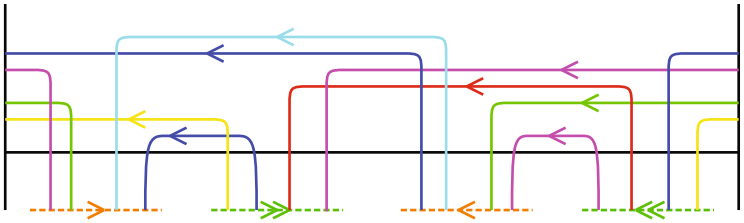


Fig. 23 The curves in $J^1(S^1)$ corresponding to the curves $(1, 0), (0, -1), (-1, -1), (-1, 0), (0, 1), (1, 1)$

of their co-normal lifts, in order to have them run parallel to the boundary of the 0-handle in T^2 except possibly at 1-handles. This allows us to identify them in $J^1(S^1)$. We isotope each curve so that at the end, the co-orientation points inward. Since there are multiple curves, we choose an isotopy which minimizes the number of crossings.

Once these curves agree (except where they enter the 1-handles) with parallel copies of suitable concentric circles which are positive Reeb push-offs of boundary of the 0-handle of the T^2 0-section, we identify a neighborhood of boundary of the square (containing all our isotoped curves) with a neighborhood of the zero-section of $J^1(S^1)$. This identifies all our curves in T^2 with curves in $J^1(S^1)$ as in Fig. 23. We then satellite the image of the curves in $J^1(S^1)$ onto the image of S^1 in the Gompf diagram of T^*T^2, using the conventions described in the previous section and illustrated in Figs. 15 and 17 to maintain the relative positions of the curves. The result is Fig. 24.

Performing a sequence of handle slides and Legendrian isotopy yields a simpler Weinstein handle diagram shown in Fig. 25. The Weinstein handle diagram allows us to easily compute the homology of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2}\setminus\nu(\widetilde{\Sigma_1})$, where $\widetilde{\Sigma_1}$ is the smoothed toric divisor and $\nu(\widetilde{\Sigma_1})$ is the neighborhood of $\widetilde{\Sigma_1}$. The handle structure determines

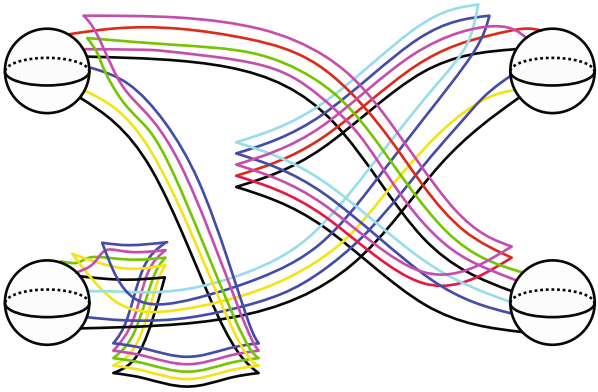


Fig. 24 A Legendrian handle diagram obtained by simplifying the Legendrian handlbody diagram of the complement of the toric divisor of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2}$ smoothed in six nodes

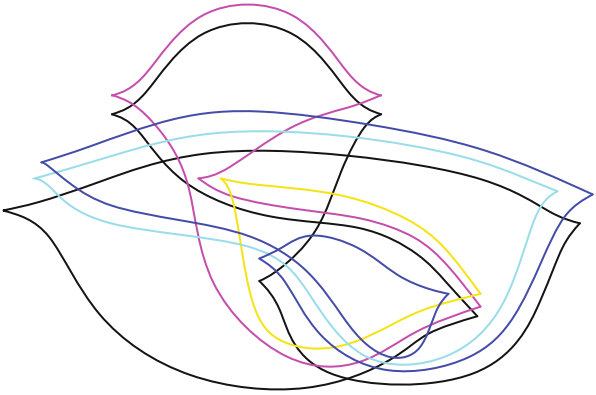


Fig. 25 A Legendrian link obtained by simplifying the Legendrian handle diagram of the complement of the toric divisor of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2}$ smoothed in six nodes

the cellular/Morse homology chain complex, which allows us to determine the homology groups. A presentation for the fundamental group can also be computed from the handle decomposition where the 1-handles correspond to generators and 2-handles provide relations.

In this example we obtain:

$$\pi_1(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma}_1); \mathbb{Z}) = 0,$$

$$H_0(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma}_1); \mathbb{Z}) = \mathbb{Z},$$

$$H_1(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma}_1); \mathbb{Z}) = 0,$$

$H_2(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma}_1); \mathbb{Z}) = \mathbb{Z}^{\oplus 5}$ with intersection form $\begin{bmatrix} 0 & -2 & 1 & 0 & 0 \\ -2 & -4 & -1 & 1 & 1 \\ 1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 1 & -2 & -1 \\ 0 & 1 & 0 & -1 & -3 \end{bmatrix}$,

$H_i(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma}_1); \mathbb{Z}) = 0$ for $i = 3, 4$.

Using the same toric diagram, instead of smoothing all six nodes, we can choose to smooth only a subset of them. The symplectic topology of the corresponding complement will differ (with fewer handles attached). We consider here an example yielding a Weinstein manifold of interest. Let us smooth three alternating nodes in $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2}$. The complement of the smoothed divisor given by T^*T^2 with 2-handles attached along $\Lambda_{(1,0)}$, $\Lambda_{(-1,-1)}$, $\Lambda_{(0,1)}$ is pictured in Fig. 26. Performing a sequence of handle slides and Legendrian isotopy on Fig. 26 yields a simpler Weinstein handle diagram shown in Fig. 27 consisting of a Legendrian (2,4) torus knot.

In this picture we can see two Lagrangian spheres in distinct homology classes in the complement of this smoothed divisor. Again, the homology of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma}_2)$, where $\widetilde{\Sigma}_2$ is the smoothed toric divisor and $\nu(\widetilde{\Sigma}_2)$ is the neighborhood of $\widetilde{\Sigma}_2$ can be computed from the Weinstein handle diagram. In particular,

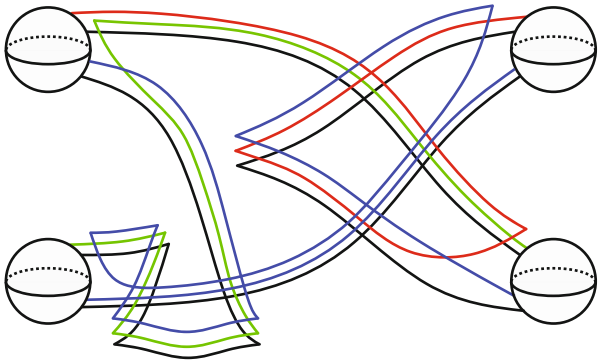


Fig. 26 A Legendrian handle diagram obtained by simplifying the Legendrian handlbody diagram of the complement of the toric divisor of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2}$ smoothed in three nodes

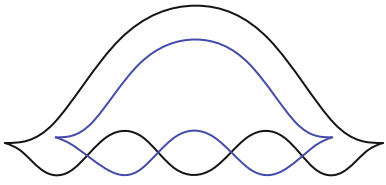


Fig. 27 A Legendrian (2,4)-torus link obtained from Fig. 26 by a sequence of handleslides and Legendrian isotopy

$$\pi_1(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2}); \mathbb{Z}) = 0,$$
$$H_0(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2}); \mathbb{Z}) = \mathbb{Z},$$
$$H_1(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2}); \mathbb{Z}) = 0,$$
$$H_2(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2}); \mathbb{Z}) = \mathbb{Z} \oplus \mathbb{Z} \text{ with intersection form } \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix},$$
$$H_i(\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2}); \mathbb{Z}) = 0 \text{ for } i = 3, 4.$$

The Weinstein diagram allows us to also compute the DGA of $\Lambda \in S^3$, $\mathcal{A}(\Lambda)$ (see [13] for a good survey on how to calculate the Legendrian DGA). In this case is generated by a, b, x, y, z, w and marked points t_1 and t_2, as shown in Fig. 28. The generators have gradings

$$|a| = |b| = 1, |x| = |y| = |z| = |w| = 0$$

with differential

$$\partial x = \partial y = \partial z = \partial w = 0$$
$$\partial a = xyzw + xw + xy + zw + 1 + t_1^{-1}$$
$$\partial ab = wzyx + yx + wz + wx + 1 + t_2.$$

The DGA of Λ is an interesting invariant to consider because the Wrapped Fukaya chain complex of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2})$ is A_∞-quasi isomorphic to the DGA $\mathcal{A}(\Lambda)$ [4, 9, 10]. Note that $\mathcal{A}(\Lambda)$ has no negatively graded Reeb chords and therefore the degree-0 Legendrian contact homology of $\mathcal{A}(\Lambda)$, denoted by $LCH_0(\Lambda)$, is finitely generated. Furthermore, it is expected that $LCH_0(\Lambda)$ is Morita equivalent to a commutative ring R such that $\hat{X} = Spec(R)$ where $\hat{X}$ is the mirror of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2} \setminus \nu(\widetilde{\Sigma_2})$. See [6, 11] for other examples of such a phenomenon.

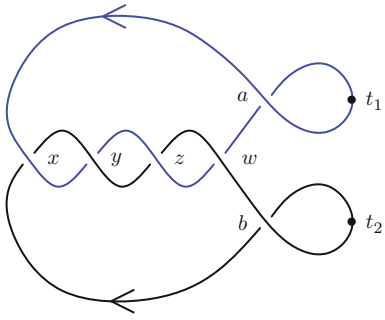


Fig. 28 A diagram of the complement of the toric divisor of $\mathbb{CP}^2\#3\overline{\mathbb{CP}^2}$ smoothed in three nodes in the Lagrangian projection with labelled Reeb chords and marked points generating the Chekanov-Eliashberg DGA

Acknowledgments This project was initiated at the 2019 Research Collaboration Conference for Women in Symplectic and Contact Geometry and Topology (WiSCon) that took place on July 22–26, 2019 at ICERM. The authors would like to extend their gratitude to the WiSCon organizers and the hosting institution ICERM and their staff for their hospitality. The authors would like to thank Lenhard Ng for useful conversations. BA and AM would like to thank the Oberwolfach Research Institute for Mathematics for hosting them during the completion of this project. AG was partially supported by Wallenberg grant no. KAW 2016-0440 and the Fondation Mathématique Jacques Hadamard. AM is partially supported by Ministry of Education and Science of Republic of Serbia, project ON174034. LS is supported by NSF grant no. DMS 1904074. AW is supported by the Engineering and Physical Sciences Research Council [EP/L015234/1], the EPSRC Centre for Doctoral Training in Geometry and Number Theory (The London School of Geometry and Number Theory), University College London. OC-S is supported by NSF Graduate Research Fellowship under grant no. DGE-1644868.

References

1. M. Abouzaid, *Morse homology, tropical geometry, and homological mirror symmetry for toric varieties*, Selecta Math. (N.S.) **15** (2009), no. 2, 189–270. MR 2529936
2. B. Acu, O. Capovilla-Searle, A. Gadbled, A. Marinković, E. Murphy, L. Starkston, and A. Wu, *Weinstein handlebodies for complements of smoothed toric divisors*, arXiv:2012.08666v1 (2020).
3. D. Auroux, *Mirror symmetry and T-duality in the complement of an anticanonical divisor*, J. Gökova Geom. Topol. GGT **1** (2007), 51–91. MR 2386535
4. F. Bourgeois, T. Ekholm, and Y. Eliashberg, *Effect of Legendrian surgery*, Geom. Topol. **16** (2012), no. 1, 301–389, With an appendix by Sheel Ganatra and Maksim Maydanskiy. MR 2916289
5. B. Chantraine, G. Dimitroglou Rizell, P. Ghiggini, and R. Golovko, *Geometric generation of the wrapped Fukaya category of Weinstein manifolds and sectors*, arXiv:1712.09126 (2017).
6. R. Casals and E. Murphy, *Legendrian fronts for affine varieties*, Duke Math. J. **168** (2019), no. 2, 225–323.
7. F. Ding and H. Geiges, *Handle moves in contact surgery diagrams*, J. Topol. **2** (2009), no. 1, 105–122. MR 2499439
8. S. K. Donaldson, *Symplectic submanifolds and almost-complex geometry*, J. Differential Geom. **44** (1996), no. 4, 666–705. MR 1438190
9. T. Ekholm, *Holomorphic curves for Legendrian surgery*, arXiv:1906.07228 (2019).
10. Tobias Ekholm and YankıLekili, *Duality between Lagrangian and Legendrian invariants*, arXiv:1701.01284 (2017).
11. Tolga Etgü and YankıLekili, *Fukaya categories of plumbings and multiplicative preprojective algebras*, Quantum Topol. **10** (2019), no. 4, 777–813. MR 4033516
12. T. Ekholm and L. Ng, *Legendrian contact homology in the boundary of a subcritical Weinstein 4-manifold*, J. Differential Geom. **101** (2015), no. 1, 67–157. MR 3356070
13. John B. Etnyre and Lenhard Ng, *Legendrian contact homology in* $\mathbb{R}^3$, arXiv preprint arXiv:1811.10966 (2019).
14. Y. Eliashberg and L. Polterovich, *Local Lagrangian* 2*-knots are trivial*, Ann. of Math. (2) **144** (1996), no. 1, 61–76. MR 1405943
15. John B. Etnyre, *Chapter 3 - legendrian and transversal knots*, Handbook of Knot Theory (William Menasco and Morwen Thistlethwaite, eds.), Elsevier Science, Amsterdam, 2005, pp. 105–185.
16. M. Gross, P. Hacking, and S. Keel, *Mirror symmetry for log Calabi-Yau surfaces I*, Publ. Math. Inst. Hautes Études Sci. **122** (2015), 65–168. MR 3415066

17. Emmanuel Giroux, *Géométrie de contact: de la dimension trois vers les dimensions supérieures*, Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002), Higher Ed. Press, Beijing, 2002, pp. 405–414. MR 1957051
18. _____, *Remarks on Donaldson's symplectic submanifolds*, Pure Appl. Math. Q. **13** (2017), no. 3, 369–388. MR 3882202
19. R. E. Gompf, *Handlebody construction of Stein surfaces*, Ann. of Math. (2) **148** (1998), no. 2, 619–693. MR 1668563
20. S. Ganatra, J. Pardon, and V. Shende, *Sectorial descent for wrapped Fukaya categories*, arXiv:1809.03427 (2018).
21. _____, *Covariantly functorial wrapped Floer theory on Liouville sectors*, V. Publ.math.IHES (2019).
22. Paul Hacking and Ailsa Keating, *Homological mirror symmetry for log Calabi-Yau surfaces*, arXiv:2005.05010 (2020).
23. D. McDuff and D. Salamon, *Introduction to symplectic topology*, Oxford Mathematical Monographs, Oxford University Press, 2017.
24. A. Weinstein, *Contact surgery and symplectic handlebodies*, Hokkaido Math. J. **20** (1991), no. 2, 241–251. MR 1114405
25. C. Wendl, *Strongly fillable contact manifolds and J-holomorphic foliations*, Duke Math. J. **151** (2010), no. 3, 337–384. MR 2605865

Constructions of Lagrangian Cobordisms

Sarah Blackwell, Noémie Legout, Caitlin Leverson, Maÿlis Limouzineau, Ziva Myer, Yu Pan, Samantha Pezzimenti, Lara Simone Suárez , and Lisa Traynor

1 Introduction

A contact manifold is an odd-dimensional manifold Y^{2n+1} together with a maximally non-integrable hyperplane distribution ξ. In a contact manifold, *Legendrian submanifolds* play a central role. These are the maximal integral submanifolds of

S. Blackwell
Department of Mathematics, University of Georgia, Athens, GA, USA
e-mail: seblackwell@uga.edu

N. Legout
Department of Mathematics, Uppsala University, Uppsala, Sweden
e-mail: noemie.legout@math.uu.se

C. Leverson
Bard College, Annandale-on-Hudson, NY, USA
e-mail: cleverson@bard.edu

M. Limouzineau
Cologne University, Cologne, Germany

Z. Myer · L. Traynor (✉)
Department of Mathematics, Bryn Mawr College, Bryn Mawr, PA, USA
e-mail: zmyer@brynmawr.edu; ltraynor@brynmawr.edu

Y. Pan
Center of Applied Mathematics, Tianjin University, Tianjin, Tianjin, China

S. Pezzimenti
Penn State Brandywine, Media, PA, USA
e-mail: spezzimenti@psu.edu

L. S. Suárez
Ruhr-Universität Bochum, Bochum, Germany
e-mail: lara.suarezlopez@rub.de

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_5

ξ: Λ^n such that $T_p\Lambda \subset \xi$, for all $p \in \Lambda$. In general, Legendrian submanifolds are plentiful and easy to construct. In this article we will restrict our attention to the contact manifold $\mathbb{R}^3$ with its standard contact structure $\xi = \ker\alpha$, where $\alpha = dz - ydx$. In this setting, every smooth knot or link has an infinite number of non-equivalent Legendrian representatives. More background on Legendrian knots is given in Sect. 2.

The even-dimensional siblings of contact manifolds are symplectic manifolds. These are even-dimensional manifolds M^{2n} equipped with a closed, non-degenerate 2-form ω. In symplectic manifolds, *Lagrangian submanifolds* play a central role. Lagrangian submanifolds are the maximal dimensional submanifolds where ω vanishes on the tangent spaces: L^n such that $\omega|_L = 0$. When the symplectic manifold is exact, $\omega = d\lambda$, it is important to understand the more restrictive subset of *exact* Lagrangians: these are submanifolds where $\lambda|_L$ is an exact 1-form. Geometrically, L exact means that for any closed curve $\gamma \subset L$, $\int_\gamma \lambda = 0$. In this article, we will restrict our attention to a symplectic manifold that is symplectomorphic to $\mathbb{R}^4$ with its standard symplectic structure $\omega_0 = \sum dx_i \wedge dy_i$. In contrast to Legendrians, Lagrangians are scarce. For example, in $\mathbb{R}^4$ with its standard symplectic structure, the torus is the only closed surface that will admit a Lagrangian embedding into $\mathbb{R}^4$. A famous theorem of Gromov [37] states that there are *no* closed, exact Lagrangian submanifolds of $\mathbb{R}^4$.

There has been a great deal of recent interest in a certain class of non-closed, exact Lagrangian submanifolds, known as *Lagrangian cobordisms*. These Lagrangian submanifolds live in the symplectization of a contact manifold and have cylindrical ends over Legendrians. In this article, we will focus on exact, orientable Lagrangian cobordisms from the Legendrian Λ_- to the Legendrian Λ_+ that live in the symplectization of $\mathbb{R}^3$; this symplectization is $\mathbb{R} \times \mathbb{R}^3$ equipped with the exact symplectic form $\omega = d(e^t\alpha)$, where t is the coordinate on $\mathbb{R}$ and $\alpha = dz - ydx$ is the standard contact form on $\mathbb{R}^3$. See Fig. 5 for a schematic picture of a Lagrangian cobordism and Definition 1 for a formal definition. Such Lagrangian cobordisms were first introduced in Symplectic Field Theory (SFT) [25]: in relative SFT, we get a category whose objects are Legendrians and whose morphisms are Lagrangian cobordisms. *Lagrangian fillings* occur when $\Lambda_- = \emptyset$ and are key objects in the Fukaya category, which is an important invariant of symplectic four-manifolds. A *Lagrangian cap* occurs when $\Lambda_+ = \emptyset$.

A basic question tied to understanding the general existence and behavior of Lagrangian submanifolds is to understand the existence of Lagrangian cobordisms: *Given two Legendrians $\Lambda_\pm$, when does there exist a Lagrangian cobordism from Λ_- to Λ_+?* There are known to be a number of obstructions to this relation on Legendrian submanifolds coming from both classical and non-classical invariants of the Legendrians $\Lambda_\pm$. Some of these obstructions are described in Sect. 2.3. To complement the obstructions, there are some known constructions. For example, it is well known [7, 23, 26] that there exists a Lagrangian cobordism between Legendrians $\Lambda_\pm$ that differ by Legendrian isotopy. In addition, by Ekholm et al. [23], Chantraine [8], it is known that there exists a Lagrangian cobordism from Λ_- to Λ_+ if Λ_- can be obtained from Λ_+ by a "pinch" move or if $\Lambda_+ = \Lambda_- \cup U$, where

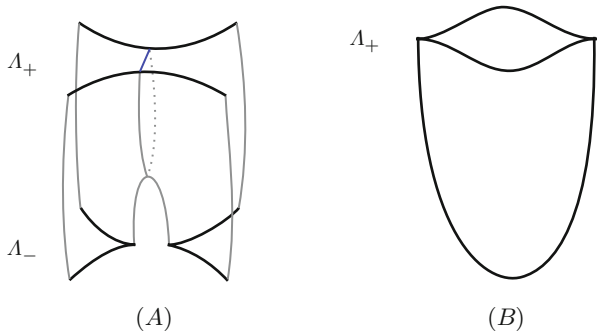


Fig. 1 (**a**) The pinch move on Λ_+ produces a Lagrangian saddle; (**b**) Λ_+ obtained by introducing an unknotted component to Λ_- corresponds to the Lagrangian cobordism having a local min

U denotes a Legendrian unknot with maximal Thurston-Bennequin number of -1 that is contained in the complement of a ball containing Λ_-. Topologically, between these slices, the cobordism changes by a saddle move (1-handle) and the addition of a local minimum (0-handle); see Fig. 1. It is important to notice that there is *not* an elementary move corresponding to a local maximum (2-handle) move. By stacking these individual cobordisms obtained from isotopy, saddles, and minimums, one obtains what is commonly referred to as a *decomposable* Lagrangian cobordism. Through these moves, it is easy to construct Lagrangian cobordisms and fillings; see an example in Fig. 7.

Towards understanding the existence of Lagrangians, it is natural to ask: *Does there exist a Lagrangian cobordism from Λ_- to Λ_+ if and only if there exists a decomposable Lagrangian cobordism from Λ_- to Λ_+?* We know the answer to this question is "No": by studying the "movies" of the not necessarily Legendrian slices of a Lagrangian, Sauvaget, Murphy, and Lin [42, 59] have shown that there exists a genus two Lagrangian cap of the Legendrian unknot with Thurston-Bennequin number equal to -3 and rotation number 0. The *Lagrangian diagram moves* used by Lin [42] to construct a Lagrangian cap are described in Sect. 3.3. The necessity of a local maximum when $\Lambda_+ \neq \emptyset$ is not currently understood.

To formulate some precise motivating questions, we will use *ribbon cobordism* to denote a $2n$-dimensional manifold that can be built from k-handles with $k \leq n$. This idea of restricting the handle index is well known in symplectic topology: Eliashberg [15, 51] has shown that any $2n$-dimensional Stein manifold admits a handle decomposition with handles of dimension at most n, and thus any $2n$-dimensional Stein cobordism between closed, $(2n-1)$-dimensional contact manifolds must be ribbon. Working in the relative setting with submanifolds and using the handle decomposition from the "height" function given by the $\mathbb{R}$ coordinate on $\mathbb{R} \times \mathbb{R}^3$, we see that all decomposable 2-dimensional Lagrangian cobordisms between 1-dimensional Legendrian submanifolds are ribbon cobordisms. We are led to the following natural questions.

Motivating Questions 1 Suppose $\Lambda_+ \neq \emptyset$ and there exists a connected Lagrangian cobordism L from Λ_- to Λ_+. Then:

1. Does there exist a decomposable Lagrangian cobordism from Λ_- to Λ_+?
2. Does there exist a ribbon Lagrangian cobordism from Λ_- to Λ_+?
3. Is L Lagrangian isotopic to a ribbon and/or decomposable Lagrangian cobordism?

There are some results known about Motivating Question 3 for the special case of the simplest Legendrian unknot. If U denotes the Legendrian unknot with Thurston-Bennequin number -1, it is known that every (exact) Lagrangian filling is orientable [55], and there is a unique (exact, orientable) Lagrangian filling of U up to compactly supported Hamiltonian isotopy [27]. Moreover, any Lagrangian cobordism from U to U is Lagrangian isotopic, via a compactly supported Hamiltonian isotopy, to one in a countable collection given by the trace of a Legendrian isotopy induced by a rotation [11].

Motivating Questions (1) and (2) are closely related and have deep ties to important questions in topology. Observe that a "yes" answer to (1) implies a "yes" to (2): if the existence of a Lagrangian cobordism implies the existence of a decomposable Lagrangian cobordism, then we also know the existence of a ribbon cobordism. Also note that when Λ_+ is topologically a slice knot and $\Lambda_- = \emptyset$, (2) is a symplectic version of the topological Slice-Ribbon conjecture: is every Lagrangian slice disk a ribbon disk? Cornwell, Ng, and Sivek conjecture that the answer to Motivating Question (1) and (3) is "No": using the theory of satellites, we know that there is a Lagrangian concordance between $\Lambda_\pm$ shown in Fig. 2, and in [17, Conjecture 3.3] it is conjectured that the concordance between the pair is not decomposable.

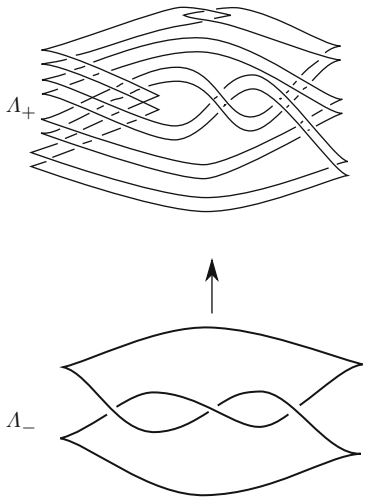


Fig. 2 There is a Lagrangian concordance between these Legendrian knots that is conjectured to be non-decomposable. Here Λ_- is a Legendrian trefoil and Λ_+ is a Legendrian Whitehead double of $m(9_{46})$

Very recently, Roberta Guadagni has discovered additional combinatorial moves that can be used to construct a "movie," meaning a sequence of slice pictures, of a Lagrangian cobordism; Fig. 9 illustrates one of these tangle moves. With one of Guadagni's moves, it is possible to construct a movie of a Lagrangian cobordism between the Legendrians pictured in Fig. 2; see Fig. 10. Guadagni's moves are "geometric": they are developed through proofs similar to those used in the satellite procedure, and thus the handle attachments involved in the cobordism are not obvious. In particular, at this point it is not known if Guadagni's tangle moves are independent from the decomposable moves.

This survey article is organized as follows. In Sect. 2, we provide some background on Legendrians and Lagrangians, formally define Lagrangian cobordisms, and summarize known obstructions to the existence of Lagrangian cobordisms. In Sect. 3, we describe three "combinatorial" ways to construct Lagrangian cobordisms, and in Sect. 4, we describe more abstract "geometric" ways to construct Lagrangian concordances and cobordisms through satellites. Then in Sect. 5, we describe some potential pathways—through the theory of rulings, Heegaard-Floer homology, and contact surgery—to potentially show the existence of Legendrians that are Lagrangian cobordant but are not related by a decomposable Lagrangian cobordism.

2 Background

2.1 Legendrian Knots and Links

In this section, we give a very brief introduction to Legendrian submanifolds in $\mathbb{R}^3$ and their invariants. More details can be found, for example, in the survey paper [28].

In $\mathbb{R}^3$, the **standard contact structure** ξ is a 2-dimensional plane field given by the kernel of the 1-form $\alpha = dz - ydx$. In $(\mathbb{R}^3, \xi = \ker\alpha)$, a **Legendrian knot** is a knot in $\mathbb{R}^3$ that is tangent to ξ everywhere. A useful way to visualize a Legendrian knot is to project it from $\mathbb{R}^3$ to $\mathbb{R}^2$. There are two useful projections: the **Lagrangian projection**

$$\begin{aligned} \pi_L : \quad \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x, y, z) &\mapsto (x, y), \end{aligned}$$

as well as the **front projection**

$$\begin{aligned} \pi_F : \quad \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x, y, z) &\mapsto (x, z). \end{aligned}$$

An example of a Legendrian trefoil is shown in Fig. 3.

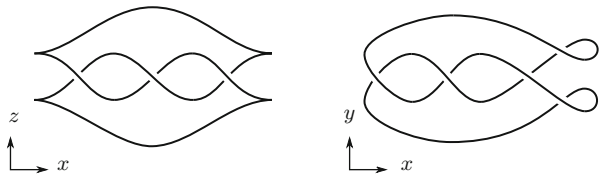


Fig. 3 The front projection (left) and the Lagrangian projection (right) of a Legendrian trefoil

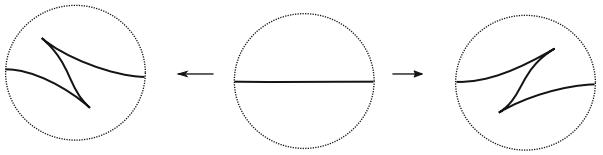


Fig. 4 Two ways to stabilize a Legendrian knot in the front projection

Legendrian submanifolds are equivalent if they can be connected by a 1-parameter family of Legendrian submanifolds. In fact, for each topological knot type there are infinitely many different Legendrian knots. Indeed, we can stabilize a Legendrian knot (as shown in Fig. 4) to get another Legendrian knot of the same topological knot type. We can see that these are not Legendrian equivalent using Legendrian invariants.

Two useful classical invariants of Legendrian knots Λ are the Thurston-Bennequin number $tb(\Lambda)$ and the rotation number $r(\Lambda)$. They can be computed easily from front projections. Given the front projection of a Legendrian knot or link Λ, the **Thurston-Bennequin number** is

$$tb(\Lambda) = \text{writhe}(\pi_F(\Lambda)) - \#(\text{right cusps}),$$

where the writhe is the number of crossings counted with sign. Once the Legendrian knot is equipped with an orientation, the **rotation number** is

$$r(\Lambda) = \frac{1}{2}\Big(\#(\text{down cusps}) - \#(\text{up cusps})\Big).$$

One can use these two invariants to see that stabilizations change the Legendrian knot type.

In future sections, we will not assume that our Legendrians $\Lambda_\pm$ come equipped with an orientation. In our Motivating Questions described in Sect. 1, our Lagrangian cobordisms are always orientable, so the existence of a Lagrangian cobordism from Λ_- to Λ_+ will induce orientations on $\Lambda_\pm$.

There are many powerful non-classical invariants that can be assigned to a Legendrian knot. Although this will not be a focus of this paper, we will give a brief description of some of these invariants. One important invariant stems from *normal rulings*, defined independently by Chekanov and Pushkar [54] and Fuchs

[30]. A count of normal rulings leads to *ruling polynomials* [54]; more details will be discussed in Sect. 5.1. Through the closely related theory of generating families, one can also associate invariant polynomials that record the dimensions of *generating family homology groups* [32, 40, 58, 61]. In addition, through the theory of pseudo-holomorphic curves, one can associate to a Legendrian Λ a *differential graded algebra (DGA)*, $\mathcal{A}(\Lambda)$ [13, 24]. An *augmentation* is a DGA map from $\mathcal{A}(\Lambda)$ to a field. The count of augmentations is closely related to the count of ruling polynomials [30, 47, 50]. Augmentations can be used to construct finite-dimensional *linearized contact homology groups* [13], which are often known to be isomorphic to the generating family homology groups [32]. In addition, there are invariants for Legendrian knots coming from *Heegaard Floer Homology* [43, 52].

2.2 Lagrangian Cobordisms

Lagrangian cobordisms between Legendrian submanifolds always have "cylindrical ends" over the Legendrians, but other conditions vary: sometimes it is specified that the Lagrangian is exact, is embedded (or immersed), is orientable, or has a fixed Maslov class. In this paper, a Lagrangian cobordism is always exact, embedded, and orientable.

Definition 1 Let $\Lambda_\pm$ be two Legendrian knots or links in $(\mathbb{R}^3, \xi = \ker\alpha)$. A **Lagrangian cobordism** L **from** Λ_- **to** Λ_+ is an embedded, orientable Lagrangian surface in the symplectization $(\mathbb{R}\times\mathbb{R}^3, d(e^t\alpha))$ such that for some $N > 0$,

1. $L \cap ([-N, N]\times\mathbb{R}^3)$ is compact,
2. $L \cap ((N,\infty)\times\mathbb{R}^3) = (N,\infty)\times\Lambda_+$,
3. $L \cap ((-\infty,-N)\times\mathbb{R}^3) = (-\infty,-N)\times\Lambda_-$, and
4. there exists a function $f : L \to \mathbb{R}$ and constant numbers $\mathfrak{c}_\pm$ such that $e^t\alpha|_{TL} = df$, where $f|_{(-\infty,-N)\times\Lambda_-} = \mathfrak{c}_-$, and $f|_{(N,\infty)\times\Lambda_+} = \mathfrak{c}_+$.

A **Lagrangian filling of** Λ_+ is a Lagrangian cobordism with $\Lambda_- = \emptyset$; a **Lagrangian cap of** Λ_- is a Lagrangian cobordism with $\Lambda_+ = \emptyset$. A **Lagrangian concordance** occurs when $\Lambda_\pm$ are knots and L has genus 0.

Figure 5 is a schematic representation of a Lagrangian cobordism.

Remark 1 In condition (4) of Definition 1, the fact that $\Lambda_\pm$ are Legendrian will guarantee that $f_\pm$ will be locally constant. Using this, it follows that any genus zero Lagrangian surface that is cylindrical over Legendrian knots will be exact. When $\Lambda_\pm$ have multiple components, one needs to check that the constant does not vary: this condition guarantees the exactness of the Lagrangian cobordism obtained by "gluing" together Lagrangian cobordisms [10].

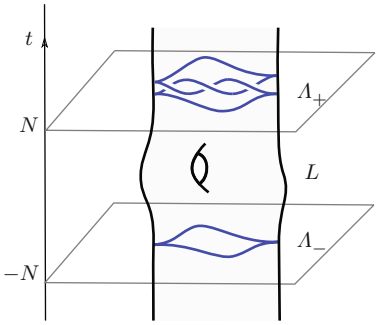


Fig. 5 A Lagrangian cobordism from Λ_- to Λ_+

Remark 2 In contrast to topological cobordisms, Lagrangian cobordisms form a non-symmetric relationship on Legendrian knots [9]. In this article we will always denote the direction of increasing $\mathbb{R}_t$ coordinate by an arrow.

2.3 Obstructions to Lagrangian Cobordisms

The focus of this paper is on constructing Lagrangian cobordisms between two given Legendrians $\Lambda_\pm$. In the smooth world, any two knots are related by a smooth cobordism, but in this more restrictive Lagrangian world, there are a number of obstructions that are important to keep in mind when trying to explicitly construct Lagrangian cobordisms. Here we mention a few that come from classical and non-classical invariants of the Legendrians $\Lambda_\pm$.

Obstructions

1. If there exists a Lagrangian cobordism of genus g between Λ_- and Λ_+, then there must exist a smooth cobordism of genus g between the smooth knot types of Λ_- and Λ_+. Thus any obstruction of a smooth genus g cobordism between Λ_- and Λ_+ would obstruct a Lagrangian genus g cobordism.
2. Since there are no closed, exact Lagrangian surfaces [37], if there exists a Lagrangian cap (respectively, filling) for Λ, then there cannot exist a Lagrangian filling (respectively, cap) of Λ.
3. As shown in [7], if there exists a Lagrangian cobordism L from Λ_- to Λ_+, then

$$r(\Lambda_-) = r(\Lambda_+) \quad \text{and} \quad tb(\Lambda_+) - tb(\Lambda_-) = -\chi(L).$$

In particular, if a Legendrian knot Λ admits a Lagrangian filling or cap, then $r(\Lambda) = 0$. Also, combining this equality on tb and the slice-Bennequin inequality [56], we see that, when Λ is a single component knot, if there exists a Lagrangian cap L of Λ, then $tb(\Lambda) \leq -1$ and $g(L) \geq 1$.

4. If there exists a Maslov 0 ([21]) Lagrangian cobordism Σ from Λ_- to Λ_+, and Λ_- has an augmentation, then
   (a) $\#Aug(\Lambda_+; \mathbb{F}_2) \geq \#Aug(\Lambda_-; \mathbb{F}_2)$, where $\mathbb{F}_2$ is the finite field of two elements, and $\#Aug(\Lambda; \mathbb{F}_2)$ denotes the number of augmentations of Λ to $\mathbb{F}_2$ up to DGA homotopy [6, 53], and
   (b) the ruling polynomials $R_{\Lambda_\pm}(z)$ (see Sect. 5.1 for definitions) satisfy
   $$R_{\Lambda_-}(q^{1/2} - q^{-1/2}) \leq q^{-\chi(\Sigma)/2} R_{\Lambda_+}(q^{1/2} - q^{-1/2}),$$
   for any q that is a power of a prime number [53].
5. If Λ admits a Maslov 0 Lagrangian filling L, and if ϵ_L denotes the augmentation of Λ induced by L, then $LCH^k_{\epsilon_L}(\Lambda) \cong H_{n-k}(L)$, which is known as the Ekholm-Seidel isomorphism [20], and whose proof was completed by Dimitroglou Rizell in [19]. More generally, if there is a cobordism from Λ_- to Λ_+, and if Λ_- admits an augmentation, then [12] provides several long exact sequences relating the homology of the cobordism and the Legendrian contact (co)homologies of its Legendrian ends. A version of this isomorphism and these long exact sequences using generating families are given in [58].
6. If Λ admits an augmentation, Λ does not admit a Lagrangian cap, as the augmentation implies the non-acyclicity of the DGA $\mathcal{A}(\Lambda)$ [22, Theorem 5.5], and from [18, Corollary 1.9] if a Legendrian admits a Lagrangian cap then its DGA $\mathcal{A}(\Lambda)$ (with $\mathbb{F}_2$ coefficients) is acyclic.

There are additional obstructions, obtained through Heegaard Floer Theory, that can be used to obstruct Lagrangian concordances and cobordisms [1, 3, 35]. Some of these will be discussed more in Sect. 5.3.

Remark 3 Observe that the Obstructions 4 and 6 assume that the bottom Λ_- has an augmentation, and stabilized knots will never have an augmentation. It would be nice to have more obstructions when Λ_- is a stabilized knot. This might be possible using the theory of "satellites" described in Sect. 4.1: it is possible for the satellite of a stabilized Legendrian to admit an augmentation. See Sect. 4.3 for more discussions in this direction.

3 Combinatorial Constructions of Lagrangian Cobordisms

A convenient way of visualizing topological cobordisms is through "movies": a sequence of pictures that represent slices of the Lagrangian. In this section, we describe three known combinatorial ways to construct Lagrangian cobordisms through such an approach.

3.1 Decomposable Moves

It is well known that if Λ_- and Λ_+ are Legendrian isotopic, then there exists a Lagrangian cobordism from Λ_- to Λ_+; see, for example, [7, 23, 26]. Isotopy, together with two types of handle moves, form the basis for decomposable Lagrangian cobordisms.

Theorem 1 ([5, 23]) *If the front diagrams of two Legendrian links Λ_- and Λ_+ are related by any of the following moves, there is a Lagrangian cobordism L from Λ_- to Λ_+.*

Isotopy *There is a Legendrian isotopy between Λ_- and Λ_+; see Fig. 6a–c for Reidemeister Moves I–III.*

1**-handle** *The front diagram of Λ_- can be obtained from the front diagram of Λ_+ by "pinching" two oppositely-oriented strands; see Fig. 6d. We will also refer to this move as a "Pinch Move."*

0**-handle** *The front diagram of Λ_- can be obtained from the front diagram of Λ_+ by deleting a component of Λ_+ that is the front diagram of a standard Legendrian unknot U with maximal Thurston-Bennequin number of -1 as long as there exist disjoint disks $D_U, D_{U^c} \subset \mathbb{R}^2_{xz}$ containing the xz-projection of U and the other components of Λ_+, respectively. Such an "unknot filling" can be seen in Fig. 6e.*

Definition 2 A Lagrangian cobordism L from Λ_- to Λ_+ is called **elementary** if it arises from isotopy, a single 0-handle, or a single 1-handle. A Lagrangian cobordism L from Λ_- to Λ_+ is **decomposable** if it is obtained by stacking elementary Lagrangian cobordisms.

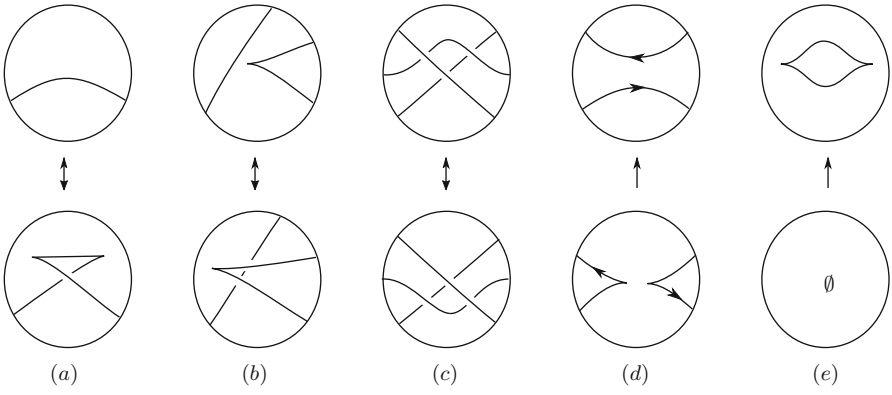


Fig. 6 Decomposable moves in terms of front projections. Arrows indicate the direction of increasing $\mathbb{R}_t$ coordinate in the symplectization. The move in (**b**) only shows the Reidemeister II move in the left cusp case, but there is an analogous move for the right cusp case

Observe that there is *not* an elementary move corresponding to a 2-handle (maximum). Also note that the elementary 1-handle (saddle) move can be used to connect two components or to split one component into two.

Decomposable cobordisms are particularly convenient as they are easy to describe in a combinatorial fashion, through a list of embedded Legendrian curves,

$$\Lambda_- = \Lambda_0 \to \Lambda_1 \to \cdots \to \Lambda_n = \Lambda_+,$$

where the front projection of the Legendrian Λ_{i+1} is related to that of Λ_i by isotopy or one of the 0-handle or 1-handle moves.

Example 1 One can construct a Lagrangian filling of a positive Legendrian trefoil with maximal Thurston-Bennequin number using the series of moves shown in Fig. 7: a 0-handle, followed by three Reidemeister I moves, followed by two 1-handles (or pinch moves). This gives a genus 1 (orientable, exact) Lagrangian filling of this Legendrian trefoil. Since we are assuming that Lagrangian fillings and caps are always exact, this implies that this trefoil cannot admit a Lagrangian cap; see Sect. 2.3 Obstruction 2.

Example 2 Using elementary moves, one can also construct a Lagrangian concordance from the unknot with $tb = -1$ to a Legendrian representative of the knot $m(9_{46})$, as shown on Fig. 8.

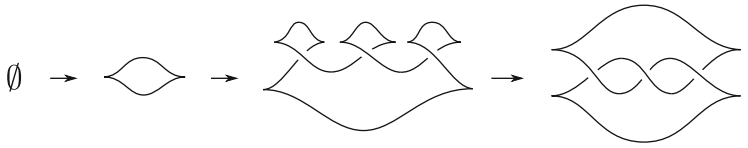


Fig. 7 A decomposable Lagrangian filling of a Legendrian trefoil

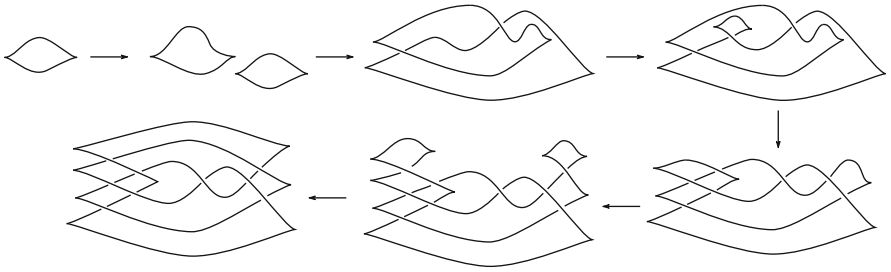


Fig. 8 A decomposable Lagrangian cobordism from a Legendrian unknot to a Legendrian $m(9_{46})$

3.2 *Guadagni Moves*

Very recently, Roberta Guadagni has discovered a new "tangle" move; see Fig. 9. This is not a local move: there are some global requirements. In particular, this move cannot be applied if all components of the tangle are contained in the same component of Λ_-: the component of Λ_- containing the blue strand must be different than the components containing the other strands of the tangle.

Example 3 With Guadagni's tangle move, it is possible to construct a Lagrangian cobordism between the Legendrians pictured in Fig. 2; see Fig. 10. However, at this point it is not known if Guadagni's tangle move is independent of the decomposable moves.

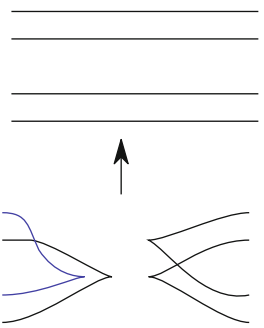


Fig. 9 Under some global conditions, there exists a Lagrangian cobordism between these tangles

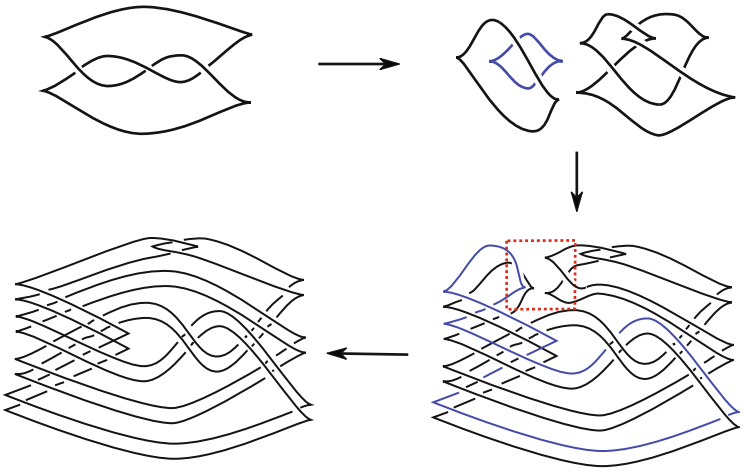


Fig. 10 A movie, using a Guadagni move, of an (orientable, exact) Lagrangian cobordism from the trefoil to the Whitehead double of $m(9_{46})$ in Fig. 2

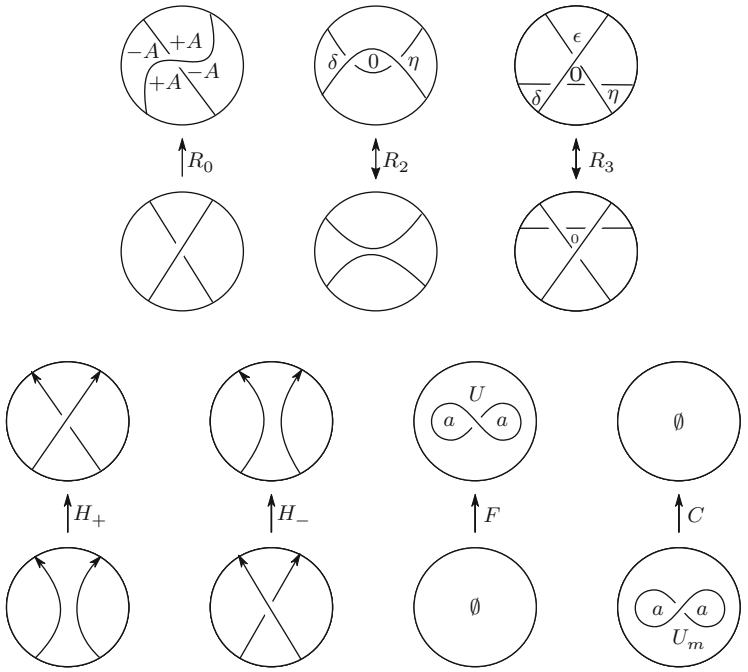


Fig. 11 The Lagrangian diagram moves. The labels in the R_0 move represent the change of area through the move, while other labels $0, \epsilon, \delta, \eta, a$ indicate the area of the corresponding regions; here 0 represents a positive area that is smaller than either the area ϵ, the area δ or the area η

3.3 Lagrangian Diagram Moves

As shown in Sect. 3.1, decomposable cobordisms are constructed from 0-handles and some 1-handles (saddles) but no 2-handles (caps). Based on the work of Sauvaget [59], Lin [42] constructs a genus two cap of a twice stabilized unknot, and thus gives the first explicit example of a non-decomposable Lagrangian cobordism. The construction describes time-slices of a Lagrangian cobordism through a list of moves on "decorated Lagrangian diagrams."

A **decorated Lagrangian diagram** is a curve in the xy-plane with the compact regions decorated by a positive number, which is the **area** of the region. Figure 11 shows some examples: in the illustration of the F move, U is a Lagrangian projection of the Legendrian unknot with maximal Thurston-Bennequin number; in the illustration of the C move, U_m *is* a decorated Lagrangian diagram, but is *not* the Lagrangian projection of a Legendrian knot.

Theorem 2 ([42]) *Let $\Lambda_\pm$ be Legendrian links and $D_\pm$ be their corresponding decorated Lagrangian projections. If one can create a sequence of decorated Lagrangian diagrams*

$$D_- = D_0 \to D_1 \to \cdots \to D_n = D_+$$

such that each diagram D_{i+1} can be obtained from D_i by the following combinatorial moves, then there is a compact Lagrangian submanifold in $\mathbb{R} \times \mathbb{R}^3$ with boundary $\Lambda_- \cup \Lambda_+$, where $\Lambda_\pm \subset \{\pm N\} \times \mathbb{R}^3$, for some $N > 0$.

1. *R_0: a planar isotopy that changes areas by the amount $\pm A$, for $A > 0$. This operation can only be done in the direction specified.*
2. *R_2: a Reidemeister II move. One can either introduce or eliminate two crossings assuming some area conditions are satisfied: it is possible to introduce or remove two crossings as long as the area of the inner region, denoted by 0 in the diagram, is less than either the area δ or the area η. One can also do this move with the lower strand passing under the upper strand.*
3. *R_3: a Reidemeister III move. One can perform a Reidemeister III move as long as the area of the inner region, denoted by 0 in the diagram, is less than either the area ϵ, the area δ or the area η. The fixed center crossing can be reversed. Additionally, the moving strand can also occur as an overstrand.*
4. *H_+: a handle attachment that creates a positive crossing in the diagram.*
5. *H_-: a handle attachment that removes a negative crossing in the diagram.*
6. *F: a filling that creates the diagram U, which is the Lagrangian projection of an unknot with maximal Thurston-Bennequin number.*
7. *C: a cap that eliminates the diagram U_m, which is the topological mirror of U.*

*These moves are called **Lagrangian diagram moves**. Moreover, the constructed Lagrangian will be exact if, in addition,*

(E1) Each move results in a diagram with all components having a total signed area equal to 0. The signed area of a region is determined by the sum of the signed heights of its Reeb chords.

(E2) If a handle attachment merges two components of a link, the components being merged must be vertically split, meaning that the images of the xy-projections of these components are contained in disjoint disks.

Remark 4

1. For condition (E2), the H_- can never be applied to merge components, and H_+ can only be applied if the components being merged are vertically split.
2. A main distinction between the Lagrangian diagram moves and the decomposable moves is that each diagram D_i in the middle of the sequence is not necessarily the Lagrangian projection of a Legendrian link. They are just the xy-projection of some time t_i-slice of the cobordism. Thus the Lagrangian diagram moves are more flexible than the decomposable moves. However, keeping track of the areas is an added complication.

Example 4 Figure 12 illustrates the construction of a Lagrangian torus using the Lagrangian diagram moves. This torus fails to be exact since condition (E1) is violated. Figure 13 gives another construction of a Lagrangian torus. This time, all components have signed area 0, but now condition (E2) is violated.

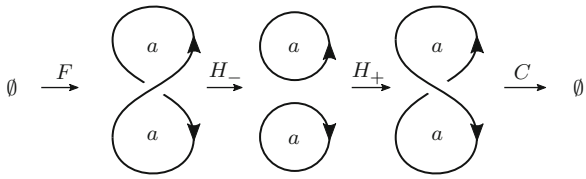


Fig. 12 A (non-exact) Lagrangian torus constructed using the Lagrangian diagram moves. The middle figure violates (E1)

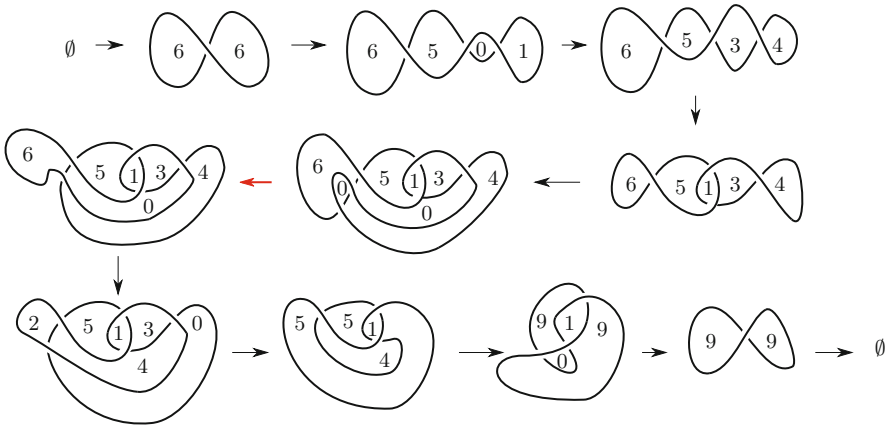


Fig. 13 A (non-exact) Lagrangian torus constructed using the Lagrangian diagram moves. These figures satisfy (E1) but (E2) is violated in the step labelled by a red arrow

4 Geometrical Constructions of Lagrangian Cobordisms

An important general way to know of the existence of Lagrangian cobordisms without using the constructions described in Sect. 3 comes through the satellite operation. In this section, we review the satellite construction and then state results from [17, 38] about the existence of a Lagrangian concordance/cobordism from Λ_- to Λ_+ implying the existence of a Lagrangian concordance/cobordism between corresponding satellites.

4.1 The Legendrian Satellite Construction

We begin by reviewing the construction of a Legendrian satellite; see also [48, Appendix] and [17, Sect. 2.2]. To construct a Legendrian satellite, begin by identifying the open solid torus $S^1 \times \mathbb{R}^2$ with the 1-jet space of the circle, $J^1 S^1 \cong$

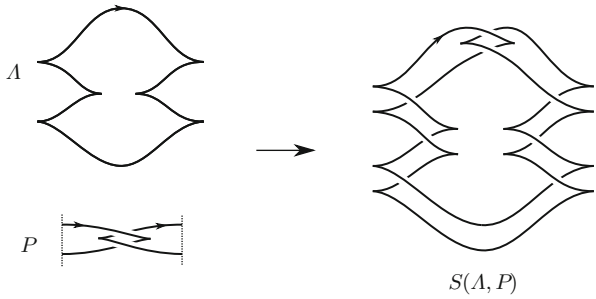


Fig. 14 An example of a Legendrian satellite

$T^*S^1 \times \mathbb{R}$, equipped with the contact form $\alpha = dz - ydx$, where x, y are the coordinates in T^*S^1 and z is the coordinate in $\mathbb{R}$. Similar to the situation for $\mathbb{R}^3 \cong J^1\mathbb{R}$, we can recover a Legendrian knot in J^1S^1 from its front projection in $S^1_x \times \mathbb{R}_z$, which is typically drawn by representing S^1 as an interval with its endpoints identified.

Given an oriented Legendrian **companion** knot $\Lambda \subset \mathbb{R}^3$ and a oriented Legendrian **pattern** knot $P \subset J^1(S^1)$, the Legendrian neighborhood theorem says that Λ has a standard neighborhood $N(\Lambda)$ such that there is a contactomorphism $\kappa : J^1(S^1) \to N(\Lambda)$. The **Legendrian satellite**, $S(\Lambda, P)$, is then the image $\kappa(P)$. The front projection of $S(\Lambda, P)$ is as shown in Fig. 14. In particular, suppose that the front projection of the pattern P intersects the vertical line at the boundary of the S^1 interval n times. We then make an n-copy of Λ by using n-disjoint copies of Λ that all differ by small translations in the z-direction. Take a point on the front projection of Λ that is oriented from left to right, cut the front of the n-copy open along the n-copy at that point, and insert the front diagram of P. The orientation on the satellite $S(\Lambda, P)$ is induced by the orientation on P.

Remark 5 The satellite operation often makes Legendrian knots "nicer"; for example, in Fig. 14, the companion Λ is stabilized and does not admit an augmentation or a normal ruling. However, the satellite $S(\Lambda, P)$ does admit a normal ruling and an augmentation.

4.2 Lagrangian Cobordisms for Satellites

In [17, Theorem 2.4], Cornwell, Ng, and Sivek, show that Lagrangian *concordance* is preserved by the Legendrian satellite operation.

Theorem 3 ([17]) *Suppose $P \subset J^1S^1$ is a Legendrian knot. If there exists a Lagrangian concordance L from a Legendrian knot Λ_- to a Legendrian knot Λ_+, then there exists a Lagrangian concordance L_P from $S(\Lambda_-, P)$ to $S(\Lambda_+, P)$.*

In particular, as shown in Fig. 8, there is a Lagrangian concordance from Λ_-, which is the Legendrian unknot with $tb = -1$, to Λ_+, which is the Legendrian

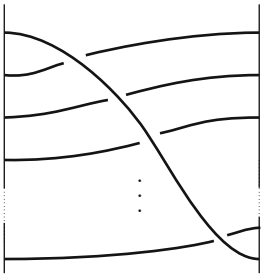


Fig. 15 For an n-stranded tangle, repeating this basic tangle n times produces a full twist

$m(9_{46})$ with maximal $tb = -1$. Using the Legendrian "clasp" tangle P as shown in Fig. 14—which produces the Legendrian Whitehead double—we can conclude that there exists a Lagrangian concordance from $S(\Lambda_-, P)$ to $S(\Lambda_+, P)$. In fact, $S(\Lambda_-, P)$ is the positive trefoil with $tb = 1$. Thus Theorem 3 implies that there exists a Lagrangian concordance between the Legendrian knots in Fig. 2.

Conjecture 1 ([17, Conjecture 3.3]) The Lagrangian concordance from $S(\Lambda_-, P)$ to $S(\Lambda_+, P)$ built through the satellite construction is not decomposable.

Theorem 3 has been extended to higher genus cobordisms by Guadagni, Sabloff, and Yacavone in [38]. To state their theorem, we need to first introduce the notion of "twisting" and then closing a tangle $T \subset J^1[0, 1]$. Given a Legendrian tangle $T \subset J^1[0, 1]$, ΔT is the tangle obtained by adding the tangle T and the full twist tangle Δ, which is illustrated in Fig. 15; the tangle $\Delta^t T$ can be thought of as T followed by t full twists. Given a Legendrian tangle $T \subset J^1[0, 1]$, $\overline{T} \subset J^1(S^1)$ will denote the associated closure to a Legendrian link.

Theorem 4 ([38]) *Suppose $T \subset J^1[0, 1]$ is a Legendrian tangle whose closure $\overline{T} \subset J^1(S^1)$ is a Legendrian knot. If there exists a Lagrangian cobordism L from Λ_- to Λ_+ of genus $g(L)$, then there exists a Lagrangian cobordism L_T from $S(\Lambda_-, \overline{\Delta^{2g(L)+1}T})$ to $S(\Lambda_+, \overline{\Delta T})$.*

In fact, Theorem 4 can be generalized to use the closure of different tangles T_- and T_+ that are Lagrangian cobordant; for details, see [38].

Remark 6 It is natural to wonder if, along the lines of Conjecture 1, this higher genus satellite procedure can create additional *candidates* for Legendrians that can be connected by a Lagrangian cobordism but not by a decomposable Lagrangian cobordism. In [38, Theorem 1.5], it is shown that if the cobordism L from Λ_- to Λ_+ is decomposable and the handles in the decomposition satisfy conditions known as "Property A", then the corresponding satellites $S(\Lambda_-, \Delta^{2g(L)+1}P)$ and $S(\Lambda_+, \Delta P)$ will also be connected by a decomposable Lagrangian cobordism. In particular, if there exists a decomposable cobordism L that does *not* satisfy Property A and is not isotopic to a cobordism that satisfies Property A, then the satellite construction would lead to a higher genus candidate that generalizes Conjecture 1.

4.3 Obstructions to Cobordisms Through Satellites

In Sect. 2.3, some known obstructions to the existence of a Lagrangian cobordism were mentioned. As mentioned in Remark 3, a number of these obstructions require Λ_- to admit an augmentation, and thus in particular Λ_- must be non-stabilized. However, as mentioned in Remark 5, it is possible for the satellite of a Legendrian Λ to admit an augmentation even if Λ does not. So the contrapositive of Theorem 3 provides a potential strategy for further obstructions to the existence of a Lagrangian cobordism when Λ_- does not admit an augmentation. For example, motivated by Obstruction 4 in Sect. 2.3, one can ask: *Can a count of augmentations give an obstruction to the existence of a Lagrangian concordance from $S(\Lambda_-, P)$ to $S(\Lambda_+, P)$ and thereby obstruct the existence of a Lagrangian concordance from Λ_- to Λ_+?* In fact, this augmentation count will not likely provide a further obstruction: a simple computation shows that when Λ is stabilized enough, the number of augmentations of $S(\Lambda, P)$ only depends on the Legendrian pattern P. If trying to pursue this path to obtain further obstructions to Lagrangian cobordisms, it is useful to keep in mind the following result of Ng that shows the DGA of the satellite of a Legendrian Λ might only remember the underlying knot type of Λ.

Theorem 5 ([49]) *Suppose Λ_1 and Λ_2 are stabilized Legendrian knots that are of the same topological knot type and have the same Thurston-Bennequin and rotation numbers. For a Legendrian pattern P whose front intersects a vertical line by two points, the DGAs of $S(\Lambda_1, P)$ and $S(\Lambda_2, P)$ are equivalent.*

5 Candidates for Non-decomposable Lagrangian Cobordisms

Now that we have developed some ways to construct a Lagrangian cobordism through combinatorial moves and satellites, we state some theorems that show *if* a Lagrangian cobordism does exist, then it cannot be decomposable: this addresses Motivating Question 1. While we discuss these theorems, it is useful to keep in mind the known obstructions to Lagrangian cobordisms that were mentioned in Sect. 2.3.

5.1 Candidates for Non-decomposable Lagrangian Cobordisms from Normal Rulings

One simple way to show that two Legendrians $\Lambda_\pm$ cannot be connected by a decomposable Lagrangian cobordism comes from a count of "combinatorial" rulings. Roughly, a **normal ruling** of a Legendrian Λ is a "decomposition" of the front projection into pairs of paths from left cusps to right cusps such that

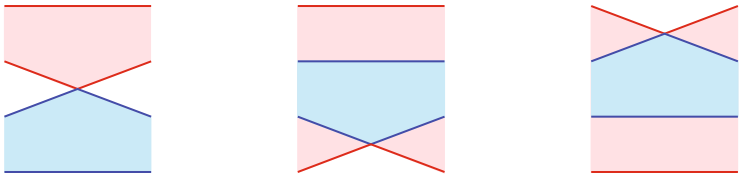


Fig. 16 Normal rulings near a switch

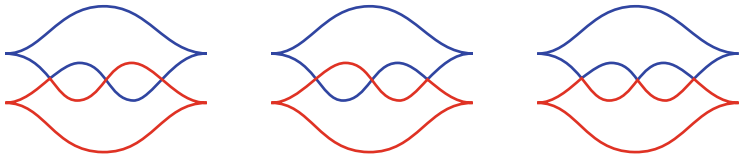


Fig. 17 All normal rulings of this max tb positive Legendrian trefoil

1. each pair of paths starts from a common left cusp and ends at a common right cusp, has no further intersections, and bounds a topological disk whose boundary is smooth everywhere other than at the cusps and certain crossings called **switches**, and
2. near a switch, the pair of paths must be arranged as in one of the diagrams in Fig. 16; observe that near the switch, vertical slices of the associated disks are either disjoint or the slices of one are contained in the slices of the other.

Formal definitions of normal rulings can be found in, for example, [54] and [30].

As an illustration, all normal rulings of a particular Legendrian trefoil are shown in Fig. 17.

For each normal ruling R, let $s(R)$ and $d(R)$ be the number of switches and number of disks, respectively. By Pushkar and Chekanov [54], the **ruling polynomial**

$$R_\Lambda(z) = \sum_R z^{s(R)-d(R)},$$

where the sum is over all the normal rulings, is an invariant of Λ under Legendrian isotopy. Normal rulings and augmentations are closely related even though they are defined in very different ways [30, 31, 50, 57].

We have the following obstruction to decomposable cobordisms in terms of normal rulings.

Theorem 6 *If Λ_- has m normal rulings and Λ_+ has n normal rulings with $m > n$, then there is no decomposable Lagrangian cobordism from Λ_- to Λ_+.*

Proof One can compare the number of normal rulings of the two ends for the decomposable moves, as shown in Fig. 18. Thus any normal ruling of Λ_- induces a normal ruling of Λ_+. Different normal rulings of Λ_- induce different normal rulings

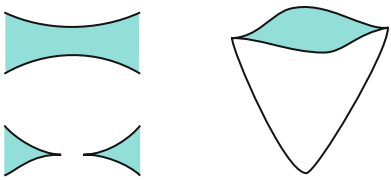


Fig. 18 Comparison of normal rulings for decomposable moves

of Λ_+. Therefore the number of normal rulings of Λ_+ is bigger than or equal to the number of normal rulings of Λ_-. □

Here is a strategy to show the existence of Legendrians that can be connected by a Lagrangian cobordism but not by one that is decomposable.

Strategy 1 *Choose Legendrians $\Lambda_\pm$ such that:*

1. *Λ_+ has fewer graded normal rulings than Λ_-, and*
2. *it is possible to construct, via a combination of the combinatorial constructions from Sect. 3 or the satellite construction from Sect. 4, a Lagrangian cobordism from Λ_- to Λ_+.*

Remark 7 If $\Lambda_\pm$ admit normal rulings, they will admit augmentations [31, 57]. From Sect. 2.3 Obstruction 4(b), we then know that if there is a Lagrangian cobordism from Λ_- to Λ_+, their ruling polynomials satisfy

$$R_{\Lambda_-}(q^{1/2} - q^{-1/2}) \leq q^{-\chi(\Sigma)/2} R_{\Lambda_+}(q^{1/2} - q^{-1/2}),$$

for any q that is a power of a prime number. Satisfying condition (1) in Strategy 1 means that the polynomial on the right side of the inequality has fewer terms than the polynomial on the left side of the inequality. If following this approach, it may be helpful to start by first finding a pair of positive integer coefficient polynomials that satisfy this inequality and condition (1) at the same time. One can start with checking the ruling polynomials of small crossing number Legendrian knots in [14].

5.2 Candidates for Non-decomposable Lagrangian Concordances from Topology

Observe that any decomposable Lagrangian concordance will be a smooth ribbon concordance. Thus it is potentially possible to use known obstructions to ribbon concordances to find examples of smooth knots whose Legendrian representatives cannot be connected by a decomposable Lagrangian concordance: constructing a Lagrangian concordance between very stabilized Legendrian representatives of these knot types, via the combinatorial techniques of Sect. 3 or geometric techniques of Sect. 4, will give an example of an exact Lagrangian concordance between knots that cannot be connected by a decomposable Lagrangian concordance.

For example, it is known [36, 41, 62] that the only knot that admits a ribbon concordance to the unknot is the unknot itself. This has as a corollary the following obstruction to a decomposable Lagrangian concordance.

Theorem 7 ([17, Theorem 3.2]) *If Λ_- is topologically non-trivial and Λ_+ is topologically an unknot, then there is no decomposable Lagrangian concordance from Λ_- to Λ_+.*

Example 5 To illustrate this theorem, here is a possible low crossing number Legendrian knot to examine as Λ_-. Consider the topological knot 6_1 which is slice and ribbon. Its maximum tb Legendrian representative Λ_{6_1} (see Fig. 19) has $tb = -5$ and $r = 0$. The DGA of this Legendrian $\mathcal{A}(\Lambda_{6_1})$ admits an augmentation, and thus Λ_{6_1} does not admit a Lagrangian cap; see Obstruction 6. in Sect. 2.3. Since we are trying to construct a Legendrian Λ_- that could be Lagrangian concordant to a stabilized unknot, which might have a Lagrangian cap, we will add some stabilizations that will prevent augmentations and thereby allow the possibility of a Lagrangian cap. If we now add a positive and a negative stabilization to Λ_{6_1}, we get a knot $\Lambda_{6_1}^{\pm}$ with $tb = -7$ and $r = 0$, which has no augmentation and is still topologically the knot 6_1. If, by a sequence of moves in Sect. 3, one can construct a concordance from $\Lambda_{6_1}^{\pm}$ to the $tb = -7$ stabilized unknot, then by Theorem 7 this Lagrangian concordance will not be decomposable; see Fig. 20. In fact, one can stabilize Λ_{6_1} *as many times as we wish* resulting in $tb(\Lambda_-) = t$ and $r(\Lambda_-) = r$ and try, using the combinatorial constructions of Sect. 3, to construct a Lagrangian concordance to Λ_+, where Λ_+ is a Legendrian unknot with $tb(\Lambda_+) = t$ and $r(\Lambda_+) = r$. *If possible*, such a construction would prove the existence of a non-decomposable Lagrangian concordance.

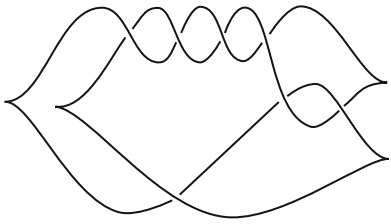


Fig. 19 Front diagram of Λ_{6_1}

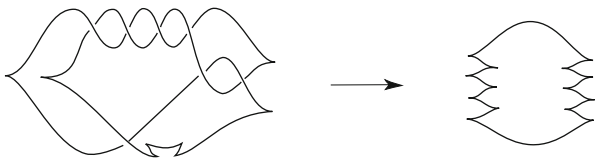


Fig. 20 Any Lagrangian concordance from the doubly stabilized Λ_{6_1} to the $tb = -7$, $r = 0$ Legendrian unknot would necessarily be non-decomposable

There are additional results from topology that give obstructions to the existence of ribbon concordances. For example, as shown by Gilmer [34] and generalized by Friedl and Powell [29], if K_- is ribbon concordant to K_+, then the Alexander polynomial of K_- divides the Alexander polynomial of K_+. We can invoke these results in a strategy to show the existence of non-decomposable Lagrangian concordances.

Strategy 2

1. *Use results from smooth topology to find examples of smooth knots $K_\pm$ such that K_- is not ribbon concordant to K_+.*
2. *For any pair of Legendrian representatives $\Lambda_\pm$ of the knot type $K_\pm$, even highly stabilized, use a combination of the combinatorial moves described in Sect. 3 to construct a Lagrangian concordance from Λ_- to Λ_+.*

The example with the knot 6_1 given above is a concrete example to try to apply this strategy with $K_- = 6_1$ and K_+ being an unknot. A possible example when K_+ is non-trivial is the following.

Example 6 Let K_- be the connect sum of the right- and left-handed trefoils, $K_- = T_r \# T_l$, and let K_+ be the connect sum of the figure-8 knot, F_8, with itself, $K_+ = F_8 \# F_8$. These knots are concordant but there is no ribbon concordance from K_- to K_+, as first shown by Gordon [36]. Choose Legendrian representatives $\Lambda_\pm$ of $K_\pm$ such that $tb(\Lambda_-) = tb(\Lambda_+)$ and $r(\Lambda_-) = r(\Lambda_+)$; note that $\Lambda_\pm$ can be very stabilized. *If* we can construct a Lagrangian concordance from Λ_- to Λ_+, via the combinatorial moves of Sect. 3, then we will have shown the existence of a pair of Legendrians that are (exactly, orientablility) Lagrangian concordant but cannot be connected by a decomposable Lagrangian concordance.

Remark 8 Some known obstructions to ribbon concordance are, in fact, obstructions to generalizations of ribbon concordance, namely **strong homotopy ribbon concordance** and **homotopy ribbon concordance**. A strong homotopy ribbon concordance is one whose complement is ribbon, i.e., can be built with only 1-handles and 2-handles. A homotopy ribbon concordance from K_- to K_+ is a concordance where the induced map on π_1 of the complement of K_- (resp. K_+) injects (resp. surjects) into π_1 of the complement of the concordance. Gordon [36] showed that

$$\begin{aligned} \text{ribbon concordant} &\implies \text{strong homotopy ribbon concordant} \\ &\implies \text{homotopy ribbon concordant.} \end{aligned}$$

There have been a number of recent results obstructing (homotopy or strong homotopy) ribbon concordances from Heegaard-Floer and Khovanov homology [39, 41, 45, 62]; these results play an important role in Strategy 2.

5.3 Candidates for Non-decomposable Lagrangian Cobordisms from GRID Invariants

Some candidates for non-decomposable Lagrangian cobordisms of higher genus come from knot Floer homology. Using the grid formulation of knot Floer homology [52], Ozsváth, Szabó, and Thurston defined Legendrian invariants of a Legendrian link $\Lambda \subset \mathbb{R}^3$, called GRID invariants, which are elements in the *hat* flavor of knot Floer homology of $\Lambda \subset -S^3$:

$$\widehat{\lambda}^+(\Lambda), \widehat{\lambda}^-(\Lambda) \in \widehat{HFK}(-S^3, \Lambda).$$

For more background, see [44, 52].

Baldwin et al. [1] have shown that these GRID invariants can be used to obstruct the existence of decomposable Lagrangian cobordisms.

Theorem 8 ([1, Theorem 1.2]) *Suppose that $\Lambda_\pm$ are Legendrian links in $\mathbb{R}^3$ such that either*

1. $\widehat{\lambda}^+(\Lambda_+) = 0$ *and* $\widehat{\lambda}^+(\Lambda_-) \neq 0$, *or*
2. $\widehat{\lambda}^-(\Lambda_+) = 0$ *and* $\widehat{\lambda}^-(\Lambda_-) \neq 0$.

Then there is no decomposable Lagrangian cobordism from Λ_- to Λ_+.

Remark 9 By Baldwin et al. [4], in the standard contact manifold $\mathbb{R}^3$, the GRID invariants agree with the LOSS invariant [43]. The LOSS invariant is functorial on Lagrangian concordances by Baldwin and Sivek [2, 3]. Thus Theorem 8 would also obstruct the existence of general Lagrangian concordances and not only the decomposable ones. To find non-decomposable cobordisms using obstructions from [1], we should focus on non-zero genus cobordisms.

Using the facts that the GRID invariants are non-zero for the $tb = -1$ Legendrian unknot and that $\widehat{\lambda}^+(\Lambda_+)$ (resp. $\widehat{\lambda}^-(\Lambda_+)$) vanish for positively (negatively) stabilized Legendrian links, Theorem 8 gives the following corollary.

Corollary 1 ([1, Corollaries 1.3, 1.4])

1. *If $\Lambda \subset \mathbb{R}^3$ is a Legendrian link such that $\widehat{\lambda}^+(\Lambda) = 0$ or $\widehat{\lambda}^-(\Lambda) = 0$, then there is no decomposable Lagrangian filling of Λ.*
2. *Suppose $\Lambda_\pm$ are Legendrian links such that either*

   (a) $\widehat{\lambda}^+(\Lambda_-) \neq 0$ and Λ_+ is the positive stabilization of a Legendrian link, or
   (b) $\widehat{\lambda}^-(\Lambda_-) \neq 0$ and Λ_+ is the negative stabilization of a Legendrian link.

   Then there is no decomposable Lagrangian cobordism from Λ_- to Λ_+.

This provides another strategy to show the existence of Legendrians $\Lambda_\pm$ that are Lagrangian conbordant but cannot be connected by a decomposable Lagrangian cobordism.

Strategy 3

1. *Find Legendrians $\Lambda_\pm$ satisfying the GRID invariants conditions of Corollary 1 and Theorem 8 such that there are no known obstructions, as described in Sect. 2.3, to the existence of a Lagrangian cobordism from Λ_- to Λ_+.*
2. *Use a combination of the combinatorial moves described in Sect. 3 to construct a Lagrangian cobordism from Λ_- to Λ_+.*

Example 7 Concrete examples mentioned in [1, Sect. 4.1] can be used for Strategy 3. Let Λ_0, Λ_1 be the Legendrian $m(10_{132})$ knots and Legendrian $m(12n_{200})$ knots shown in [46, Figs. 2 and 3]. Modify them with a pattern shown in [1, Fig. 13] to get Λ_0' and Λ_1', which are of knot type $m(12n_{199})$ and $m(14n_{5047})$ (or its mirror), respectively. For $i, j = 0, 1$ we have $tb(\Lambda_i') = tb(\Lambda_i) + 2$ and $r(\Lambda_i') = r(\Lambda_i)$. There is no decomposable Lagrangian cobordism from

1. Λ_0 to Λ_1', or
2. Λ_1 to Λ_0'.

If we can construct, using the combinatorial techniques of Sect. 3, a Lagrangian cobordism (necessarily of genus 1) from Λ_0 to Λ_1' or from Λ_1 to Λ_0', then we will have found a non-decomposable Lagrangian cobordism.

Example 8 In [1, Sect. 4.3], the authors provide an infinite family of pairs of Legendrian knots where there does not exist a decomposable Lagrangian cobordism between them.

Remark 10 In Strategies 2 and 3, we emphasized the construction of Lagrangian cobordisms using the combinatorial techniques of Sect. 3. It would be interesting to know if the geometric constructions of Sect. 4 could also be used to show the existence of a Lagrangian concordance/cobordism from the theory of normal rulings, topology, or grid invariants, that are known to not be decomposable.

5.4 Non-decomposable Candidates Through Surgery

An additional strategy to show the existence of a non-decomposable Lagrangian filling comes from understanding properties of the contact manifold that is obtained from surgery on the Legendrian knot. In particular, Conway, Etnyre, and Tosun [16] have detected a relationship between Lagrangian fillings of a Legendrian and symplectic fillings of the contact manifold obtained by performing a particular type of surgery on the Legendrian.

Theorem 9 ([16, Theorem 1.1]) *There is a Lagrangian disk filling of Λ_+ if and only if the contact $+1$-surgery on $\Lambda_+ \subset \mathbb{R}^3 \subset S^3$ produces a contact manifold that is strongly symplectically fillable. If Λ_+ has a decomposable Lagrangian filling, then the filling can be taken to be Stein.*

In fact, [16] also shows that a filling will be a Stein filling if and only if Λ_+ bounds a *regular* Lagrangian disk: a Lagrangian disk is regular if there is a Liouville vector field that is tangent to the disk. Any decomposable Lagrangian filling is regular.

We now see another strategy to construct a non-decomposable Lagrangian filling.

Strategy 4 *Find a Legendrian Λ such that the $+1$-surgery on Λ produces a contact manifold that is strongly symplectically fillable but does not admit a Stein filling.*

An issue with this approach is a lack of examples: there are very few manifolds which carry strongly fillable but not Stein fillable contact structures. The main examples are the $1/n$ surgeries on the positive and negative trefoils; see works by Ghiggini [33] and Tosun [60]. However it is not obvious whether any of these contact structures are a contact $+1$-surgery on a Legendrian knot in S^3.

6 Conclusion

The desire to understand the flexibility and rigidity of Lagrangian submanifolds has led to a great deal of interesting research in symplectic topology. Similarly, trying to understand constructions of and obstructions to Lagrangian cobordisms has led to many interesting results. At this point, we have few concrete answers to the Motivating Questions stated in our Introduction. In particular, regarding Motivating Question (1), there are presently many candidates for Legendrians $\Lambda_\pm$ that can be connected by a Lagrangian cobordism but not by a decomposable Lagrangian cobordism: by understanding all the obstructions to Lagrangian cobordisms, one can come up with some good candidates. When trying and failing to construct a Lagrangian cobordism between a given pair, one may gain intuition for additional obstructions to Lagrangian cobordisms that are waiting to be discovered.

Acknowledgments This project was initiated at the workshop Women in Symplectic and Contact Geometry and Topology (WiSCoN) that took place at ICERM in July 2019. The authors thank the NSF-HRD 1500481 - AWM ADVANCE grant for funding this workshop. Leverson was supported by NSF postdoctoral fellowship DMS-1703356. Limouzineau was supported by DFG postdoctoral fellowship, 281071066 CRC/TRR 191. We thank Emmy Murphy for suggesting and encouraging us to work on this project. In addition, we thank John Etnyre, Roberta Guadagni, Tye Lidman, Lenny Ng, Josh Sabloff, and Bülent Tosun for useful conversations related to this project.

References

1. Baldwin, J.A., Lidman, T., Wong, C.-M.M.: Lagrangian cobordisms and Legendrian invariants in knot floer homology. Michigan Math. J. 1–31, (2021). https://doi.org/10.1307/mmj/20195786
2. Baldwin, J.A., Sivek, S.: Invariants of Legendrian and transverse knots in monopole knot homology. J. Symplectic Geom. **16**(4), 959–1000 (2018). https://doi.org/10.4310/JSG.2018.v16.n4.a3

3. Baldwin, J.A., Sivek, S.: On the equivalence of contact invariants in sutured Floer homology theories. Geom. Topol. 25(3), 1087–1164 (2021). https://doi.org/10.2140/gt.2021.25.1087
4. Baldwin, J.A., Vela-Vick, D.S., Vértesi, V.: On the equivalence of Legendrian and transverse invariants in knot Floer homology. Geom. Topol. **17**(2), 925–974 (2013). https://doi.org/10.2140/gt.2013.17.925
5. Bourgeois, F., Sabloff, J., Traynor, L.: Lagrangian cobordisms via generating families: Construction and geography. Algebr. Geom. Topol. **15**(4), 2439–2477 (2015). https://doi.org/10.2140/agt.2015.15.2439
6. Capovilla-Searle, O., Legout, N., Limouzineau, M., Murphy, E., Pan, Y., Traynor, L.: Obstructions to reversing Lagrangian surgery in Lagrangian fillings. In preparation (2021).
7. Chantraine, B.: Lagrangian concordance of Legendrian knots. Algebr. Geom. Topol. **10**(1), 63–85 (2010). https://doi.org/10.2140/agt.2010.10.63
8. Chantraine, B.: Some non-collarable slices of Lagrangian surfaces. Bull. Lond. Math. Soc. **44**(5), 981–987 (2012). https://doi.org/10.1112/blms/bds026
9. Chantraine, B.: Lagrangian concordance is not a symmetric relation. Quantum Topol. **6**(3), 451–474 (2015). https://doi.org/10.4171/QT/68
10. Chantraine, B.: A note on exact Lagrangian cobordisms with disconnected Legendrian ends. Proc. Amer. Math. Soc. **143**(3), 1325–1331 (2015). https://doi.org/10.1090/S0002-9939-2014-12302-1
11. Chantraine, B., Dimitroglou Rizell, G., Ghiggini, P., Golovko, R.: Floer homology and Lagrangian concordance. In: Proceedings of the Gökova Geometry-Topology Conference 2014, pp. 76–113. Gökova Geometry/Topology Conference (GGT), Gökova (2015)
12. Chantraine, B., Dimitroglou Rizell, G., Ghiggini, P., Golovko, R.: Floer theory for Lagrangian cobordisms. J. Differential Geom. **114**(3), 393–465 (2020). https://doi.org/10.4310/jdg/1583377213
13. Chekanov, Y.V.: Invariants of Legendrian knots. In: Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002), pp. 385–394. Higher Ed. Press, Beijing (2002)
14. Chongchitmate, W., Ng, L.: An atlas of Legendrian knots. Exp. Math. **22**(1), 26–37 (2013). https://doi.org/10.1080/10586458.2013.750221
15. Cieliebak, K., Eliashberg, Y.: From Stein to Weinstein and back, *American Mathematical Society Colloquium Publications*, vol. 59. American Mathematical Society, Providence, RI (2012). https://doi.org/10.1090/coll/059. Symplectic geometry of affine complex manifolds
16. Conway, J., Etnyre, J.B., Tosun, B.: Symplectic Fillings, Contact Surgeries, and Lagrangian Disks. Int. Math. Res. Not. **2021**(8), 6020–6050 (2019). https://doi.org/10.1093/imrn/rny291
17. Cornwell, C., Ng, L., Sivek, S.: Obstructions to Lagrangian concordance. Algebr. Geom. Topol. **16**(2), 797–824 (2016). https://doi.org/10.2140/agt.2016.16.797
18. Dimitroglou Rizell, G.: Exact Lagrangian caps and non-uniruled Lagrangian submanifolds. Ark. Mat. **53**(1), 37–64 (2015). https://doi.org/10.1007/s11512-014-0202-y
19. Dimitroglou Rizell, G.: Lifting pseudo-holomorphic polygons to the symplectisation of $P \times \mathbb{R}$ and applications. Quantum Topol. **7**(1), 29–105 (2016). https://doi.org/10.4171/QT/73
20. Ekholm, T.: Rational SFT, linearized Legendrian contact homology, and Lagrangian Floer cohomology. In: Perspectives in analysis, geometry, and topology, *Progr. Math.*, vol. 296, pp. 109–145. Birkhäuser/Springer, New York (2012). https://doi.org/10.1007/978-0-8176-8277-4_6
21. Ekholm, T., Etnyre, J., Sullivan, M.: Non-isotopic Legendrian submanifolds in $\mathbb{R}^{2n+1}$. J. Differential Geom. **71**(1), 85–128 (2005). http://projecteuclid.org/euclid.jdg/1143644313
22. Ekholm, T., Etnyre, J.B., Sabloff, J.M.: A duality exact sequence for Legendrian contact homology. Duke Math. J. **150**(1), 1–75 (2009). https://doi.org/10.1215/00127094-2009-046
23. Ekholm, T., Honda, K., Kálmán, T.: Legendrian knots and exact Lagrangian cobordisms. J. Eur. Math. Soc. (JEMS) **18**(11), 2627–2689 (2016). https://doi.org/10.4171/JEMS/650
24. Eliashberg, Y.: Invariants in contact topology. In: Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998), Extra Vol. II, pp. 327–338 (1998)

25. Eliashberg, Y., Givental, A., Hofer, H.: Introduction to symplectic field theory. Special Volume, Part II, pp. 560–673 (2000). https://doi.org/10.1007/978-3-0346-0425-3_4. GAFA 2000 (Tel Aviv, 1999)
26. Eliashberg, Y., Gromov, M.: Lagrangian intersection theory: finite-dimensional approach. In: Geometry of differential equations, *Amer. Math. Soc. Transl. Ser. 2*, vol. 186, pp. 27–118. Amer. Math. Soc., Providence, RI (1998). https://doi.org/10.1090/trans2/186/02
27. Eliashberg, Y., Polterovich, L.: Local Lagrangian 2-knots are trivial. Ann. of Math. (2) **144**(1), 61–76 (1996). https://doi.org/10.2307/2118583
28. Etnyre, J.B.: Legendrian and transversal knots. In: Handbook of knot theory, pp. 105–185. Elsevier B. V., Amsterdam (2005). https://doi.org/10.1016/B978-044451452-3/50004-6
29. Friedl, S., Powell, M.: Homotopy ribbon concordance and Alexander polynomials. Arch. Math. (Basel) **115**(6), 717–725 (2020). https://doi.org/10.1007/s00013-020-01517-5
30. Fuchs, D.: Chekanov-Eliashberg invariant of Legendrian knots: existence of augmentations. J. Geom. Phys. **47**(1), 43–65 (2003). https://doi.org/10.1016/S0393-0440(01)00013-4
31. Fuchs, D., Ishkhanov, T.: Invariants of Legendrian knots and decompositions of front diagrams. Mosc. Math. J. **4**(3), 707–717, 783 (2004). https://doi.org/10.17323/1609-4514-2004-4-3-707-717
32. Fuchs, D., Rutherford, D.: Generating families and Legendrian contact homology in the standard contact space. J. Topol. **4**(1), 190–226 (2011). https://doi.org/10.1112/jtopol/jtq033
33. Ghiggini, P.: Strongly fillable contact 3-manifolds without Stein fillings. Geom. Topol. **9**, 1677–1687 (2005). https://doi.org/10.2140/gt.2005.9.1677
34. Gilmer, P.M.: Ribbon concordance and a partial order on S-equivalence classes. Topology Appl. **18**(2–3), 313–324 (1984). https://doi.org/10.1016/0166-8641(84)90016-6
35. Golla, M., Juhász, A.: Functoriality of the EH class and the LOSS invariant under Lagrangian concordances. Algebr. Geom. Topol. **19**(7), 3683–3699 (2019). https://doi.org/10.2140/agt.2019.19.3683
36. Gordon, C.M.: Ribbon concordance of knots in the 3-sphere. Math. Ann. **257**(2), 157–170 (1981). https://doi.org/10.1007/BF01458281
37. Gromov, M.: Pseudo holomorphic curves in symplectic manifolds. Invent. Math. **82**(2), 307–347 (1985). https://doi.org/10.1007/BF01388806
38. Guadagni, R., Sabloff, J.M., Yacavone, M.: Legendrian satellites and decomposable cobordisms. arXiv preprint arXiv:2103.03340 (2021)
39. Gujral, O.S., Levine, A.S.: Khovanov homology and cobordisms between split links. arXiv preprint arXiv:2009.03406 (2020)
40. Jordan, J., Traynor, L.: Generating family invariants for Legendrian links of unknots. Algebr. Geom. Topol. **6**, 895–933 (2006). https://doi.org/10.2140/agt.2006.6.895
41. Levine, A.S., Zemke, I.: Khovanov homology and ribbon concordances. Bull. Lond. Math. Soc. **51**(6), 1099–1103 (2019). https://doi.org/10.1112/blms.12303
42. Lin, F.: Exact Lagrangian caps of Legendrian knots. J. Symplectic Geom. **14**(1), 269–295 (2016). https://doi.org/10.4310/JSG.2016.v14.n1.a10
43. Lisca, P., Ozsváth, P., Stipsicz, A.I., Szabó, Z.: Heegaard Floer invariants of Legendrian knots in contact three-manifolds. J. Eur. Math. Soc. (JEMS) **11**(6), 1307–1363 (2009). https://doi.org/10.4171/JEMS/183
44. Manolescu, C., Ozsváth, P., Sarkar, S.: A combinatorial description of knot Floer homology. Ann. of Math. (2) **169**(2), 633–660 (2009). https://doi.org/10.4007/annals.2009.169.633
45. Miller, M., Zemke, I.: Knot Floer homology and strongly homotopy-ribbon concordances. Math. Res. Lett. **28**(3), 849–861 (2021).
46. Ng, L., Ozsváth, P., Thurston, D.: Transverse knots distinguished by knot Floer homology. J. Symplectic Geom. **6**(4), 461–490 (2008). URL http://projecteuclid.org/euclid.jsg/1232029299
47. Ng, L., Rutherford, D.: Satellites of Legendrian knots and representations of the Chekanov-Eliashberg algebra. Algebr. Geom. Topol. **13**(5), 3047–3097 (2013). https://doi.org/10.2140/agt.2013.13.3047
48. Ng, L., Traynor, L.: Legendrian solid-torus links. J. Symplectic Geom. **2**(3), 411–443 (2004). URL http://projecteuclid.org/euclid.jsg/1118755327

49. Ng, L.L.: Invariants of Legendrian links. ProQuest LLC, Ann Arbor, MI (2001). http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:0803248. Thesis (Ph.D.)–Massachusetts Institute of Technology
50. Ng, L.L., Sabloff, J.M.: The correspondence between augmentations and rulings for Legendrian knots. Pacific J. Math. **224**(1), 141–150 (2006). https://doi.org/10.2140/pjm.2006.224.141
51. Oancea, A.: *From Stein to Weinstein and back. Symplectic geometry of affine complex manifolds* [book review of MR3012475]. Bull. Amer. Math. Soc. (N.S.) **52**(3), 521–530 (2015). https://doi.org/10.1090/S0273-0979-2015-01487-4
52. Ozsváth, P., Szabó, Z., Thurston, D.: Legendrian knots, transverse knots and combinatorial Floer homology. Geom. Topol. **12**(2), 941–980 (2008). https://doi.org/10.2140/gt.2008.12.941
53. Pan, Y.: The augmentation category map induced by exact Lagrangian cobordisms. Algebr. Geom. Topol. **17**(3), 1813–1870 (2017). https://doi.org/10.2140/agt.2017.17.1813
54. Pushkar, P.E., Chekanov, Y.V.: Combinatorics of fronts of Legendrian links, and Arnold's 4-conjectures. Uspekhi Mat. Nauk **60**(1(361)), 99–154 (2005). https://doi.org/10.1070/RM2005v060n01ABEH000808
55. Ritter, A.F.: Novikov-symplectic cohomology and exact Lagrangian embeddings. Geom. Topol. **13**(2), 943–978 (2009). https://doi.org/10.2140/gt.2009.13.943
56. Rudolph, L.: The slice genus and the Thurston-Bennequin invariant of a knot. Proc. Amer. Math. Soc. **125**(10), 3049–3050 (1997). https://doi.org/10.1090/S0002-9939-97-04258-5
57. Sabloff, J.M.: Augmentations and rulings of Legendrian knots. Int. Math. Res. Not. (19), 1157–1180 (2005). https://doi.org/10.1155/IMRN.2005.1157
58. Sabloff, J.M., Traynor, L.: Obstructions to Lagrangian cobordisms between Legendrians via generating families. Algebr. Geom. Topol. **13**(5), 2733–2797 (2013). https://doi.org/10.2140/agt.2013.13.2733
59. Sauvaget, D.: Curiosités lagrangiennes en dimension 4. Ann. Inst. Fourier (Grenoble) **54**(6), 1997–2020 (2005) (2004). URL http://aif.cedram.org/item?id=AIF_2004__54_6_1997_0
60. Tosun, B.: Tight small Seifert fibered manifolds with $e_0 = -2$. Algebr. Geom. Topol. **20**(1), 1–27 (2020). https://doi.org/10.2140/agt.2020.20.1
61. Traynor, L.: Generating function polynomials for Legendrian links. Geom. Topol. **5**, 719–760 (2001). https://doi.org/10.2140/gt.2001.5.719
62. Zemke, I.: Knot Floer homology obstructs ribbon concordance. Ann. of Math. (2) **190**(3), 931–947 (2019). https://doi.org/10.4007/annals.2019.190.3.5

On Khovanov Homology and Related Invariants

Carmen Caprau, Nicolle González, Christine Ruey Shan Lee, Adam M. Lowrance, Radmila Sazdanović, and Melissa Zhang

AMS MSC2010 Codes: 57M27, 57M25

1 Introduction

The discovery of the Jones polynomial [18] has invigorated low-dimensional topology by introducing a plethora of link and 3-manifold invariants. Efforts to study

C. Caprau
California State University, Fresno, CA, USA
e-mail: ccaprau@csufresno.edu
https://sites.google.com/mail.fresnostate.edu/ccaprau/

N. González
UCLA, Los Angeles, CA, USA
e-mail: nicolle@math.ucla.edu
https://sites.google.com/view/nicolle-gonzalez/

C. R. S. Lee
University of South Alabama, Mobile, AL, USA
e-mail: crslee@southalabama.edu
https://sites.google.com/a/southalabama.edu/crslee/home

A. M. Lowrance
Vassar College, Poughkeepsie, NY, USA
e-mail: adlowrance@vassar.edu
http://pages.vassar.edu/adamlowrance/

R. Sazdanović
North Carolina State University, Raleigh, NC, USA
e-mail: rsazdan@ncsu.edu
https://sazdanovic.wordpress.ncsu.edu

M. Zhang (✉)
University of Georgia, Athens, GA, USA
e-mail: melissa.zhang@uga.edu
https://www.melissa-zhang.com

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_6

these quantum invariants have yielded powerful new link invariants, in the form of homology theories, through categorification. In this article, we focus on the impact of the most influential homology theory arising from quantum invariants: Khovanov homology [22]. Our goal is to sample some recent applications of Khovanov-type theories to smooth low-dimensional topology. By bringing together the various ideas and constructions, we hope to facilitate new applications.

In Sect. 2, we curate a survey of recent developments in knot concordance, mutation detection, unknotting, and the categorification of knot polynomials. Note that our overview will focus on Lee's spectral sequence, Rasmussen's s-invariant, and generalizations of these constructions. We will exclude results linking Khovanov homology to knot Floer homology or Heegaard Floer homology, for which the readers may consult the resources [3, 37, 45]. We also exclude applications toward low-dimensional contact and symplectic geometry.

Following the survey, we give two new applications. In Sect. 3, we extend Levine-Zemke's [30] ribbon concordance obstruction from Khovanov homology to $\mathfrak{sl}(n)$ foam homology for $n \geq 4$, as well as to universal $\mathfrak{sl}(2)$ and $\mathfrak{sl}(3)$ foam homology theories. More generally, we show that a ribbon concordance between links induces injective maps on link homologies defined via webs and foams modulo relations. Kang provides a different approach in [19, Theorem 1], where it is shown that a ribbon concordance induces injective maps on link homology theories that are multiplicative link TQFTs and which are either associative or Khovanov-like. Our proof relies mainly on the fact that all of the homology theories considered in Sect. 3 satisfy certain cutting neck and sphere relations in the category of dotted cobordisms, without the need to provide new definitions or develop special techniques.

In Sect. 4, we use spectral sequences coming from Khovanov homology to bound the alternation number, as well as the Turaev genus of a knot in S^3.

We hope that this article provides a convenient reference to those entering this area of research and sparks interest in the subject.

2 A Survey of Applications of Khovanov Homology

2.1 Rasmussen's s-Invariant

Possibly the most well-known application of the original Khovanov homology [22] lies in Rasmussen's [47] concordance invariant s, which comes from a spectral sequence arising from a filtration on the Lee complex. The Lee spectral sequence is a key ingredient of the proof that the Khovanov homology of an alternating knot is thin [29, Theorem 1.2]. Rasmussen shows that s induces a homomorphism from the concordance group to the integers. Therefore, it provides a slice obstruction. In fact, s gives lower bounds on the slice genus of a knot. As an example of an application, he uses this to give a strikingly short proof of the Milnor conjecture [47, Corollary

1], which was previously proven by Kronheimer and Mrowka using gauge theory [28, Corollary 1.2].

Many others have since modeled the algebraic construction of Rasmussen's invariant to produce more concordance invariants, many of which are generalizations of the s-invariant to $\mathfrak{sl}(n)$ Khovanov-Rozansky homology [31, 34, 35, 56] or to the universal $\mathfrak{sl}(2)$ homology [10].

In 2012, Lipshitz and Sarkar introduced a stable homotopy type for Khovanov homology [32]. They define a refinement of s for each stable cohomology operation, and show that the refinement corresponding to Sq^2 is stronger than s in [33, Section 5], using Cotton Seed's computations (see also [50]).

2.2 Mutants

Mutant knots are notoriously difficult to distinguish using knot invariants. It has been shown that for a knot, mutation preserves the signature, the Alexander polynomial, the volume (if the mutants in question are hyperbolic), and the Jones polynomials [13, Corollary 6, 7, and 8], [44, Theorem 2], [48, Corollary 1.4]. It is an open question whether Khovanov homology is invariant under mutation on knots. While there exist mutant *links* with distinct Khovanov homologies (see [54, Theorem 3]), it has been shown that odd Khovanov homology and Khovanov homology with $\mathbb{F}_2$ coefficients are invariant under mutation; for details, we refer the reader to [8, Theorem 1] and [55, Theorem 1.1], respectively.

There has been some recent indication that Khovanov-type theories may be used to distinguish mutants. For example, a prominent open problem was resolved when Piccirillo showed that the Conway knot is not slice [46, Theorem 1.1], using the s-invariant defined by Rasmussen from the Lee spectral sequence. Lobb-Watson's [36] filtered invariant is able to detect mutants in the presence of an involution. In a different direction, one may also consider generalized mutations along genus 2 surfaces from which (Conway) mutation may be recovered [16, Section 2.6]. It has been shown that Khovanov homology distinguishes a pair of generalized mutants, while the signature, HOMFLY-PT polynomial, Jones polynomials, and Kauffman polynomial are the same [16, Proposition 1.6].

2.3 Ribbon Concordance

Motivated by Gordon's conjecture [17, Conjecture 1.1] that ribbon concordance gives a partial ordering on knots in S^3, there has been great interest in studying the behavior of knot invariants under ribbon concordance. Notably, in 2019, Zemke [57, Theorem 1.1] showed that knot Floer homology obstructs ribbon concordance. This led to an exciting series of papers extending this result to various homology-type invariants for knots. Within the realm of Khovanov-type invariants,

Levine-Zemke [30, Theorem 1] extended the result to the original Khovanov homology, Kang [19, Theorem 1] extended the result to a setup that includes Khovanov-Rozansky homologies [26], knot Floer homologies and other theories, and Sarkar [49] defined the notion of ribbon distance [49, Section 3] and derived bounds on this from Khovanov-Lee homology [49, Theorem 1.1, Corollary 1.2].

2.4 Unknotting and Unlinking via Spectral Sequences

Besides the s-invariant and its relationship to the slice genus, one can also relate spectral sequences from Khovanov homology to other link invariants. Alishahi and Dowlin [2] proved that the page at which the Lee spectral sequence collapses can be used to give a lower bound on the unknotting number of the knot [2, Theorem 1.2]. A consequence of this bound is that the Knight Move conjecture holds for all knots with unknotting number at most two [2, Corollary 1.4]. Alishahi also proved a similar lower bound for the unknotting number using the Bar-Natan spectral sequence coming from the characteristic two Khovanov homology [1, Theorem 1.2]. In another direction, Batson and Seed [6] constructed a spectral sequence starting with the Khovanov homology of a link and converging to the Khovanov homology of the disjoint union of its components. The page at which this spectral sequence collapses yields a lower bound on the link splitting number of the link.

2.5 $\mathfrak{sl}(n)$ Homology and HOMFLY-PT Homology

For each n, the $\mathfrak{sl}(n)$ link invariant is a certain one-variable specialization of the HOMFLY-PT polynomial. In [26], Khovanov and Rozansky gave a categorification of the $\mathfrak{sl}(n)$ polynomial using matrix factorizations. Moreover, using matrix factorizations with a different potential, Khovanov and Rozansky [27] constructed a categorification of the HOMFLY-PT polynomial. For the $\mathfrak{sl}(3)$ link invariant, Khovanov [23] constructed another categorification using trivalent webs and foams between such webs. This was later generalized to the universal $\mathfrak{sl}(3)$ homology by Mackaay and Vaz [40]. An approach to the universal $\mathfrak{sl}(2)$ homology theory was constructed by Caprau [9], using a combination of ideas from [4] and [23]. In [39], Mackaay, Stošić, and Vaz gave a topological categorification of the $\mathfrak{sl}(n)$ polynomial, for all $n \geq 4$, via webs and a special type of foams. For specific details on the versions of $\mathfrak{sl}(n)$ homologies that are used in this paper, we refer the reader to Sect. 3. A potential topological application of $\mathfrak{sl}(n)$ and HOMFLY-PT homologies is that they would be better able to distinguish mutant knots, due in part to the fact that the corresponding decategorifications can detect mutants (see [42, Theorem 3], [43, Section 1.3]).

3 Link Homologies and Ribbon Concordance

Let K_0 and K_1 be knots in S^3. A *concordance* $C \subset S^3 \times [0,1]$ from K_0 to K_1 is a smooth embedding of the annulus $f : S^1 \times [0,1] \to S^3 \times [0,1]$ such that $f(S^1 \times \{0\}) = L_0 \times \{0\}$ and $f(S^1 \times \{1\}) = L_1 \times \{1\}$. In this case, we say that the knots K_0 and K_1 are *concordant*. For k-component links L_0 and L_1, a concordance is a disjoint union of k knot concordances between the components of L_0 and the components of L_1.

By a small isotopy of $S^3 \times [0,1]$, the concordance C may be adjusted so that the restriction to C of the projection $S^3 \times [0,1] \to [0,1]$ is a Morse function. If this Morse function has only critical points of index 0 (local minima) and 1 (saddle points) (that is, if it has no critical points of index 2, i.e. local maxima), then C is called a *ribbon concordance*. In this case, we say that L_0 is *ribbon concordant* to L_1.

Denote by $\overline{C}$ the mirror image of C and regard it as a concordance from L_1 to L_0. Then $\overline{C} \circ C$ is the concordance from L_0 to itself obtained by concatenating C and $\overline{C}$. Zemke [57] proved that the concordance $\overline{C} \circ C$ can be obtained by taking the identity concordance $L_0 \times [0,1]$ and "tubing in" unknotted, unlinked 2-spheres $S_1, \ldots, S_n$ using "tubes" $T_1, \ldots, T_n$. The tubes are annuli embedded in $S^3 \times [0,1]$, joining $L_0 \times [0,1]$ with the spheres $S_1, \ldots, S_n$. Specifically, Zemke [57, Section 3] explained that the concordance $\overline{C} \circ C$ can be described, up to isotopy, by the following movie presentation:

- n births of disjoint unknots $U_1, \ldots, U_n$, each of which being disjoint from the link L_0;
- n saddles represented by bands $B_1, \ldots, B_n$, such that B_i connects U_i with L_0;
- n saddles represented by bands $\overline{B}_1, \ldots, \overline{B}_n$, where each $\overline{B}_i$ is respectively the mirror image (dual) of B_i;
- n deaths, deleting $U_1, \ldots, U_n$.

The embedded annuli T_i are obtained by concatenating the second and third movie frames above, by joining the bands B_i together with their respective dual bands, $\overline{B}_i$. The births and deaths of the unknots $U_1, \ldots, U_n$ determine n unknotted, unlinked 2-spheres $S_1, \ldots, S_n$. The annuli T_i are the boundaries of some three-dimensional 1-handles h_i, and each handle h_i intersects the surface $L_0 \times [0,1]$ and the sphere S_i in some disks D_i and D_i', respectively. Then, the concordance $\overline{C} \circ C$ can be thought of as the following union:

$$\overline{C} \circ C = \big((L_0 \times [0,1]) \setminus (D_1 \cup \cdots \cup D_n)\big) \cup (T_1 \cup \cdots \cup T_n) \cup \big((S_1 \setminus D_1') \cup \cdots \cup (S_n \setminus D_n')\big).$$

The goal of this section is to use the above result by Zemke [57] to show that a ribbon concordance between two links induces an injective map on the $\mathfrak{sl}(n)$ link homology theories using foams, for all $n \geq 2$. That is, we want to show that the main result proved by Levine and Zemke in [30, Theorem 1] can be generalized to universal Khovanov homology, as well as to higher rank link

homologies. The proofs of the following statements are similar in nature to the proofs of the analogous statements provided in [30].

Here we are considering $\mathfrak{sl}(n)$ foam homologies, which we will denote by $\mathcal{H}_n$. For $n = 2$ and $n = 3$, we are working with the corresponding *universal* theories (see Remark 1). The universal theory categorifying the $\mathfrak{sl}(2)$ link polynomial corresponds to a Frobenius system of rank two associated to the ring $\mathcal{A}_2 = \mathbb{Z}[X, h, t]/(X^2 - hX - t)$, where h and t are formal parameters. The homology of the unknot is the ring $\mathcal{A}_2$, and the homology of the empty link is the ground ring $\mathbb{Z}[X, h, t]$. To obtain a homology theory that is purely functorial with respect to link cobordisms, Caprau [9] worked with singular cobordisms and with the ground ring $\mathbb{Z}[i][X, h, t]$, where $i^2 = -1$. In this paper, we work with the universal $\mathfrak{sl}(2)$ theory developed in [9]. Similarly, the universal $\mathfrak{sl}(3)$ foam theory, introduced by Mackaay and Vaz in [40], corresponds to a Frobenius system of rank three associated to the ring $\mathcal{A}_3 = \mathbb{Z}[X, a, b, c]/(X^3 - aX^2 - bX - c)$, where a, b, and c are formal parameters. For $n \geq 4$, we consider the homology theory introduced by Mackaay, Stošić, and Vaz in [39], which corresponds to the ring $\mathcal{A}_n = \mathbb{Q}[X]/(X^n)$. The foams in [39] are more complicated than those for the cases of $n = 2$ and 3, as these foams have additional types of singularities and their evaluation makes use of the Kapustin-Li formula [20].

Remark 1 These foam theories are termed 'universal' (see [24]) in contrast to the 'ordinary' $\mathfrak{sl}(2)$ (resp. $\mathfrak{sl}(3)$) homology, for which $h = t = 0$ (resp. $a = b = c = 0$). These specializations were constructed first by Mackaay and Vaz [40], generalizing Khovanov's construction of $\mathfrak{sl}(3)$ foam homology in [23]. For more recent developments in foam theories related to $\mathfrak{sl}(2)$ homology, see [25]. Note that prior to the construction of these foam theories, Khovanov and Rozansky developed $\mathfrak{sl}(n)$ homologies via matrix factorizations [26, 27]; as this method is not immediately compatible with Zemke's topological arguments, we focus solely on the available foam theories.

These homology theories use foams modulo local relations, as pioneered by Bar-Natan [4] in his approach to local Khovanov homology for tangles. In each case of the $\mathfrak{sl}(n)$ homology theory considered here for a fixed value of $n \geq 2$, one associates to a link diagram a formal chain complex in a certain abelian category $\mathrm{Kom}(\mathbf{Foam}_n)$, whose objects are column vectors of closed 1-manifolds in the plane, and whose morphisms are matrices of dotted foams in $\mathbb{R}^2 \times [0, 1]$, which are considered up to boundary-preserving isotopies, and modulo local relations.

For our purposes, for each $\mathfrak{sl}(n)$ foam homology theory for $n \geq 2$, we will only need the local relations involving smooth, oriented surfaces and (1+1)-cobordisms in $\mathbb{R}^2 \times [0, 1]$ marked with dots. Specifically, we will employ the sphere relations (S_n) and the cutting neck relation (CN_n), for fixed $n \geq 2$. Figure 1 shows these relations for the foam homology theory for $n \geq 4$, and the paragraph after the figure describes the relations for $n = 2, 3$. In this figure, a letter i on a surface means that the surface is marked with i dots. Recall that in terms of the 2-dimensional TQFT associated with the corresponding Frobenius extension and the resulting $\mathfrak{sl}(n)$ homology theory for links, a dot on a surface corresponds to the endomorphism of

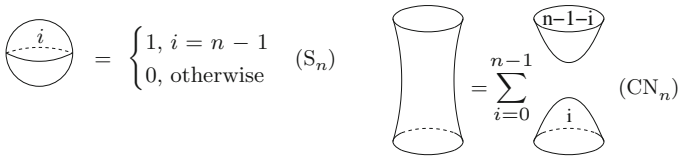


Fig. 1 The local relations (S_n) and (CN_n) for $n \geq 4$. The label 'i' (resp. '$n-1-i$') on a surface indicates the presence of i (resp. $n-1-i$) dots on that surface

the ring $\mathcal{A}_n$ that is multiplication by X. The sphere relations (S_n) are the geometric counterparts of the evaluations of the counit map $\epsilon : \mathcal{A}_n \to R$ on the generators $1, X, \ldots, X^{n-1}$, where R is the ground ring. Moreover, the cutting neck relation (CN_n), for each $n \geq 2$, is the geometric representation of the formula for $\Delta(1)$, where $\Delta : \mathcal{A}_n \to \mathcal{A}_n \otimes_R \mathcal{A}_n$ is the comultiplication map corresponding to the Frobenius system defining the 2-dimensional TQFT. Specifically, for $n \geq 4$, the dot relations (S_n) correspond to $\epsilon(1) = \epsilon(X) = \cdots = \epsilon(X^{n-2}) = 0$ and $\epsilon(X^{n-1}) = 1$, while the cutting neck relation (CN_n) corresponds to $\Delta(1) = \sum_{i=0}^{n-1} X^i \otimes X^{n-i-1}$.

For the universal $\mathfrak{sl}(2)$ foam theory [9], the dot relations (S_2) correspond to $\epsilon(1) = 0, \epsilon(X) = 1$, and the cutting neck relation (CN_2) is the geometric representation for $\Delta(1) = 1 \otimes X + X \otimes 1 - h\ 1 \otimes 1$. Moreover, for the universal $\mathfrak{sl}(3)$ foam theory (as constructed in [40]), the dot relations (S_3) are the geometric representations of the evaluations $\epsilon(1) = \epsilon(X) = 0, \epsilon(X^2) = -1$, and the cutting neck relation (CN_3) corresponds to

$$-\Delta(1) = 1 \otimes X^2 + X \otimes X + X^2 \otimes 1 - a(1 \otimes X + X \otimes 1) - b\ 1 \otimes 1.$$

We denote by $\mathcal{T}_n$, for $n \geq 2$, the tautological functors in the above homology theories (see Remark 2). Recall that these functors are multiplicative with respect to disjoint unions of objects, as well as with respect to disjoint unions of morphisms, in the geometric categories $\mathbf{Foam}_n$, for $n \geq 2$. It was proved in [9, Theorem 2] that the universal $\mathfrak{sl}(2)$ homology theory satisfies the functoriality property with respect to smooth, oriented link (and tangle) cobordisms without sign ambiguity. These cobordisms are equivalent up to boundary-preserving isotopy. Clark [12, Theorem 1.3] also proved that Khovanov's $\mathfrak{sl}(3)$ homology theory is properly functorial. Moreover, it was explained in [40, Sections 2.1–2.3] that the universal $\mathfrak{sl}(3)$ homology theory is functorial at least up to a minus sign (that is, up to multiplication by a unit in $\mathbb{Z}$). Finally, recall that the $\mathfrak{sl}(n)$ homology theory, for $n \geq 4$, is functorial (at least) up to multiplication by a non-zero rational number, as shown in [39, Proposition 8.5]. Note that for the purpose of this paper, it suffices that a certain $\mathfrak{sl}(n)$ foam homology theory is functorial up to multiplication by a unit in the ground ring.

Remark 2 For details on tautological functors and the universal construction, see [7]. Friendly examples can be found in [4, Section 9] and [23].

For the remainder of this section, embedded link cobordisms in $\mathbb{R}^3 \times [0, 1]$ may possibly be decorated with dots.

Lemma 1 *Let $F \subset \mathbb{R}^3 \times [0, 1]$ be an embedded cobordism from a link L_0 to a link L_1. Let S be an unknotted 2-sphere in $\mathbb{R}^3 \times [0, 1]$ and unlinked from F, and denote by $S^{(k)}$ the sphere S marked with k dots. Then,*

(a) $\mathcal{H}_2(F \cup S) = 0$ *and* $\mathcal{H}_2(F \cup S^{(1)}) = \mathcal{H}_2(F)$.
(b) $\mathcal{H}_3(F \cup S) = 0 = \mathcal{H}_3(F \cup S^{(1)})$ *and* $\mathcal{H}_3(F \cup S^{(2)}) = -\mathcal{H}_3(F)$.
(c) $\mathcal{H}_n(F \cup S) = \mathcal{H}_n(F \cup S^{(1)}) = \cdots = \mathcal{H}_n(F \cup S^{(n-2)}) = 0$, *and* $\mathcal{H}_n(F \cup S^{(n-1)}) = \mathcal{H}_n(F)$, *where* $n \geq 4$.

Proof If necessary, we may perform an ambient isotopy of $\mathbb{R}^3 \times [0, 1]$ so that the unknotted 2-sphere S lies in a slice $\mathbb{R}^3 \times \{t\}$, for some $t \in [0, 1]$, and that the intersection of F with $\mathbb{R}^3 \times \{t\}$ is a $(1 + 1)$-cobordism. Then, the result in part (a) follows from the sphere relations (S_2) and the properties of the functors $\mathcal{T}_2$ and $\mathcal{H}_2$. Remember that by relations (S_2), a sphere with no dot evaluates to zero, while a sphere with one dot evaluates to 1. Similarly, the sphere relations (S_3) and separately (S_n), for $n \geq 4$, together with the application of the functors $\mathcal{T}_3$ and $\mathcal{T}_n$, for $n \geq 4$, (along with the fact that $\mathcal{H}_3$ and $\mathcal{H}_n$ are functors) yield the equalities in parts (b) and (c). We note that the negative sign in part (b) arises due to the convention for the relations (S_3), in which a sphere with two dots evaluates to -1.

Lemma 2 *Let $F \subset \mathbb{R}^3 \times [0, 1]$ be an embedded cobordism from a link $L_0 \subset \mathbb{R}^3 \times \{0\}$ to a link $L_1 \subset \mathbb{R}^3 \times \{1\}$. Let $\gamma : [0, 1] \to \mathbb{R}^3 \times [0, 1]$ be a smoothly embedded arc with endpoints on F and otherwise disjoint from F, and let T be the boundary of an embedded tubular neighborhood of γ (that is, T is an annulus). Let F' be the result of removing the disk neighborhoods of $\partial\gamma$ from F and attaching T. Denote by $F^{(i,j)}$ the cobordism obtained from F' by surgery along a compressing disk of T, so that there are i dots on the disk in F' bounded by the circle where T was attached to F around $\gamma(0)$, and there are j dots on the disk bounded by the circle where T was attached to F around $\gamma(1)$.*

Then,

(a) $\mathcal{H}_2(F') = \mathcal{H}_2(F^{(1,0)}) + \mathcal{H}_2(F^{(0,1)}) - h\,\mathcal{H}_2(F)$.
(b) $-\mathcal{H}_3(F') = \mathcal{H}_3(F^{(2,0)}) + \mathcal{H}_3(F^{(1,1)}) + \mathcal{H}_3(F^{(0,2)}) - a[\mathcal{H}_3(F^{(1,0)}) + \mathcal{H}_3(F^{(0,1)})] - b\mathcal{H}_3(F)$.
(c) $\mathcal{H}_n(F') = \sum_{i=0}^{n-1} \mathcal{H}_n(F^{(i,n-1-i)})$, *where* $n \geq 4$.

Proof The proof is similar to that of Lemma 1, only that now we make use of the cutting neck relations. We also encourage the reader to compare the equalities in parts (a), (b) and (c) with the evaluations for $\Delta(1)$ for the universal $\mathfrak{sl}(2)$ foam theory, the universal $\mathfrak{sl}(3)$ foam theory, and respectively, the $\mathfrak{sl}(n)$ (for $n \geq 4$) foam homology theory. We perform first an isotopy of $\mathbb{R}^3 \times [0, 1]$ so that T lies in a small ball contained in a slice $\mathbb{R}^3 \times \{t\}$, for some $t \in [0, 1]$, and the intersections of F' and $F^{(i,j)}$ with the ball can be identified with the pictures depicted in the cutting

neck relations. The cutting neck relations imply that the morphisms in $\mathbf{Foam}_n$ corresponding to the cobordisms in the statement of the lemma (where $n = 2$ in part (a), $n = 3$ in part (b), and $n \geq 4$ in part (c)) satisfy the skein relations in the statement. Then, the claimed identities on the homology groups follow at once from these, and from the properties of the tautological functors $\mathcal{T}_n$, and since $\mathcal{H}_n$ is a functor, for each $n \geq 2$.

Proposition 1 *Let $D \subset \mathbb{R}^3 \times [0,1]$ be an embedded cobordism from a link $L_0 \subset \mathbb{R}^3 \times \{0\}$ to a link $L_1 \subset \mathbb{R}^3 \times \{1\}$. Suppose S is an unknotted 2-sphere in $\mathbb{R}^3 \times [0,1]$ and unlinked from D. Let γ be a smoothly embedded arc with one endpoint on D and the other on S, and otherwise disjoint from $D \cup S$, and let T be the boundary of an embedded tubular neighborhood of γ (that is, T is an annulus). Let D' be the result of removing the neighborhood of $\partial\gamma$ from $D \cup S$ and attaching T.*

Then $\mathcal{H}_n(D') = \mathcal{H}_n(D)$, for all $n \geq 2$.

Proof The proof follows from Lemmas 1 and 2. We apply first Lemma 2 to the cobordism $F := D \cup S$, with $F' := D'$. Then note that $F^{(i,j)} = D^{(i)} \cup S^{(j)}$, where $D^{(i)}$ is the cobordism D marked with i dots, and $S^{(j)}$ is the 2-sphere S marked with j dots. So, we have

$$\begin{aligned}\mathcal{H}_2(D') &= \mathcal{H}_2(D^{(1)} \cup S) + \mathcal{H}_2(D \cup S^{(1)}) - h\,\mathcal{H}_2(D \cup S)\\ &= 0 + \mathcal{H}_2(D) - h \cdot 0\\ &= \mathcal{H}_2(D),\end{aligned}$$

where the second equality holds due to part (a) in Lemma 1. Similarly, using part (b) from Lemma 2, we get,

$$\begin{aligned}-\mathcal{H}_3(D') = {} & \mathcal{H}_3(D^{(2)} \cup S) + \mathcal{H}_3(D^{(1)} \cup S^{(1)}) + \mathcal{H}_3(D \cup S^{(2)})\\ & -a[\mathcal{H}_3(D^{(1)} \cup S) + \mathcal{H}_3(D \cup S^{(1)})] - b\mathcal{H}_3(D \cup S).\end{aligned}$$

Using part (b) from Lemma 1, we see that only the third term above, $\mathcal{H}_3(D \cup S^{(2)})$, survives and equals to $-\mathcal{H}_3(D)$. Hence, $\mathcal{H}_3(D') = \mathcal{H}_3(D)$, as desired.

Moreover, the following equalities follow from parts (c) of the previous two lemmas:

$$\mathcal{H}_n(D') = \sum_{i=0}^{n-1} \mathcal{H}_n(D^{(i)} \cup S^{(n-1-i)}) = \mathcal{H}_n(D \cup S^{(n-1)}) + 0 = \mathcal{H}_n(D).$$

Hence, the statement holds for every $n \geq 2$.

We are now ready to prove the main result of this section.

Theorem 1 *Let C be a ribbon concordance from a link L_0 to a link L_1. Then the induced maps on $\mathfrak{sl}(n)$ homologies*

$$\mathcal{H}_n(C) : \mathcal{H}_n(L_0) \to \mathcal{H}_n(L_1)$$

are injective, for all $n \geq 2$.

Proof Let C be a ribbon concordance from L_0 to L_1, and let $\overline{C}$ be the mirror image of C (that is, $\overline{C}$ is the reverse concordance from L_1 to L_0). Let $D := \overline{C} \circ C$. Then D is a concordance from L_0 to itself. Since for each $n \geq 2$, the foam homology theory $\mathcal{H}_n$ is a functor, we have that:

$$\mathcal{H}_n(D) = \mathcal{H}_n(\overline{C}) \circ \mathcal{H}_n(C), \text{ for each } n \geq 2.$$

By the discussion at the beginning of this section, we know that the concordance D can be obtained by taking the identity concordance $L_0 \times [0, 1]$ and "tubing in" unknotted, unlinked 2-spheres $S_1, \ldots, S_n$ using embedded annuli $T_1, \ldots, T_n$. These annuli are the boundaries of embedded 3-dimensional 1-handles $h_1, \ldots, h_i$ in $\mathbb{R}^3 \times [0, 1]$, where each h_i connects $L_0 \times [0, 1]$ with S_i and is disjoint from S_j, for $j \neq i$. Then, by Proposition 1 and the functoriality properties of the corresponding foam homology theories, we get:

$$\mathcal{H}_2(D) = \mathcal{H}_2(L_0 \times [0, 1]) = \mathrm{id}_{\mathcal{H}_2(L_0)},$$

$$\mathcal{H}_3(D) = \pm\mathcal{H}_3(L_0 \times [0, 1]) = \pm\mathrm{id}_{\mathcal{H}_3(L_0)}, \text{ and}$$

$$\mathcal{H}_n(D) = q\,\mathcal{H}_n(L_0 \times [0, 1]) = q\,\mathrm{id}_{\mathcal{H}_n(L_0)}, \text{ for all } n \geq 4,$$

where $q \in \mathbb{Q}^*$. Therefore,

$$\mathcal{H}_2(\overline{C}) \circ \mathcal{H}_2(C) = \mathrm{id}_{\mathcal{H}_2(L_0)}, \ \mathcal{H}_3(\overline{C}) \circ \mathcal{H}_3(C) = \pm\mathrm{id}_{\mathcal{H}_3(L_0)}, \text{ and}$$

$$\mathcal{H}_n(\overline{C}) \circ \mathcal{H}_n(C) = q\,\mathrm{id}_{\mathcal{H}_n(L_0)}, \text{ for some } q \in \mathbb{Q}^*.$$

In all of the above cases, the composition $\mathcal{H}_n(\overline{C}) \circ \mathcal{H}_n(C)$ is a bijective function, for each $n \geq 2$. Hence for each $n \geq 2$, $\mathcal{H}_n(C)$ is an injective map and $\mathcal{H}_n(\overline{C})$ is surjective.

Remark 3 To our knowledge, it is not known whether the $\mathfrak{sl}(n)$ foam homology theory for $n \geq 4$ is purely functorial with respect to link cobordisms. But it is known that it is functorial up to multiplication by a non-zero rational number (see [39, Proposition 8.5]). This is the reason for using a $q \in \mathbb{Q}^*$ in the above proof, for the case of $n \geq 4$. Similarly, the universal $\mathfrak{sl}(3)$ link homology is known to be functorial at least up to a unit in $\mathbb{Z}$ (see [40, Proposition 2.8]), therefore the $\pm$ sign in the above proof for the case of $n = 3$.

As a consequence of Theorem 1, we obtain that the homology theories $\mathcal{H}_n$, for all $n \geq 2$, give obstructions to ribbon concordance. For any concordance C between links and any $n \geq 2$, the map $\mathcal{H}_n(C)$ preserves both the quantum and homological

grading. Then the proof of the theorem implies that for any bigrading (i, j) and $n \geq 2$, $\mathcal{H}_n^{i,j}(L_0)$ embeds in $\mathcal{H}_n^{i,j}(L_1)$ as a direct summand.

4 Gordian Distance and Spectral Sequences in Khovanov Homology

Lee [29, Section 4] defined an endomorphism of the Khovanov homology of a knot with coefficients in $\mathbb{Q}$, and Rasmussen [47, Theorem 2.1] showed that Lee's endomorphism gives rise to a spectral sequence, called the Lee spectral sequence, whose E_1 page is isomorphic to the Khovanov homology of the knot. Shumakovitch [51, Theorem 4.1.A] defined a version of Lee's spectral sequence with coefficients in the finite field $\mathbb{F}_p$ of order p, for an odd prime p. We refer to the above spectral sequences as the Lee spectral sequence with R coefficients, where R is either $\mathbb{Q}$ or $\mathbb{F}_p$ for an odd prime p. A spectral sequence *collapses* at the kth page if $E_{k-1} \neq E_k$ and $E_k = E_m$ for all $m \geq k$. When $R = \mathbb{Q}$ or $\mathbb{F}_p$, define $\mathrm{pg}_{\mathrm{Lee}}(K; R)$ to be the page at which the Lee spectral sequence with R coefficients collapses. Similarly, Bar-Natan [4, Subsection 9.3] defined a variant of Khovanov homology with $\mathbb{F}_2$ coefficients. Turner [53, Theorem 3.2] showed that Bar-Natan's variant gives rise to a spectral sequence similar in spirit to the Lee spectral sequence. Define $\mathrm{pg}_{\mathrm{BN}}(K)$ to be the page at which the Bar-Natan spectral sequence collapses.

The *Gordian distance* $d(K_1, K_2)$ between two knots K_1 and K_2 is the minimum number of crossing changes necessary to transform K_1 into K_2. The most famous Gordian distance is the *unknotting number* $u(K)$ of a knot K, which is the Gordian distance between K and the unknot. Kawauchi [21, Definition 1.2] similarly defined the *alternation number* $\mathrm{alt}(K)$ of a knot K to be the minimum Gordian distance between K and the set of alternating knots. The Khovanov homology $Kh(K; R)$ of a knot over R is *homologically thin* if there is an integer s such that $Kh^{i,j}(K; R) = 0$, for $j - 2i \neq s \pm 1$; that is, $Kh(K; R)$ is homologically thin if $Kh(K; R)$ is supported entirely in two adjacent diagonals $j - 2i = s \pm 1$. Define $d_{\mathrm{thin}}(K; R)$ to be the minimum Gordian distance between K and the set of knots that have thin Khovanov homology over R. Because every alternating link has thin Khovanov homology over R, for all rings R that we consider, it follows that $d_{\mathrm{thin}}(K; R) \leq \mathrm{alt}(K)$.

This section is organized as follows: the results in Sect. 4.1 are followed by examples in Sect. 4.2, which illuminate the proofs provided in Sect. 4.3.

4.1 Results

For any real number x, define $\lceil x \rceil$ to be the ceiling of x; that is, $\lceil x \rceil$ is the least integer that is greater than or equal to x. The next two results relate $d_{\mathrm{thin}}(K; R)$

and alt(K) with the pages $\mathrm{pg}_{\mathrm{Lee}}(K; R)$, $\mathrm{pg}_{\mathrm{BN}}(K)$ at which the Lee and Bar-Natan spectral sequences collapse.

Theorem 2 *Let K be a knot, and let R be $\mathbb{Q}$ or $\mathbb{F}_p$, where p is an odd prime. Then*

$$\mathrm{pg}_{\mathrm{Lee}}(K; R) \leq \left\lceil \frac{d_{\mathrm{thin}}(K; R) + 3}{2} \right\rceil \leq \left\lceil \frac{\mathrm{alt}(K) + 3}{2} \right\rceil. \tag{1}$$

Theorem 3 *Let K be a knot. Then*

$$\mathrm{pg}_{\mathrm{BN}}(K) \leq d_{\mathrm{thin}}(K; \mathbb{F}_2) + 2 \leq \mathrm{alt}(K) + 2. \tag{2}$$

The Turaev genus of a knot is an invariant that measures how far a knot is from being alternating in a different way than the alternation number, and it is defined as follows. Each crossing  in a knot diagram D has an A-resolution . and a B-resolution . The *all-A Kauffman state* of D is the collection of simple closed curves obtained by choosing an A-resolution for each crossing, and similarly the *all-B Kauffman state* of D is the collection of simple closed curves obtained by choosing a B-resolution for each crossing.

The knot diagram D has a Turaev surface of genus

$$g_T(D) = \frac{1}{2}(2 + c(D) - s_A(D) - s_B(D)),$$

where $c(D)$ is the number of crossings in D, and $s_A(D)$ and $s_B(D)$ are the number of components in the all-A and, respectively, all-B Kauffman states of D. The *Turaev genus* $g_T(K)$ of a knot K is defined as follows:

$$g_T(K) = \min\{g_T(D) \mid D \text{ is a diagram of } K\}.$$

It is known that a knot is alternating if and only if its Turaev genus is zero [52, Lemma 2]. The next result is a version of Theorems 2 and 3.

Theorem 4 *Let $R = \mathbb{Q}$ or $\mathbb{F}_p$ for an odd prime p. For any knot K,*

$$2\,\mathrm{pg}_{\mathrm{Lee}}(K; R) \leq g_T(K) + 4 \qquad \text{and} \qquad \mathrm{pg}_{\mathrm{BN}}(K) \leq g_T(K) + 2.$$

There are knots with arbitrarily large Turaev genus and alternation number one [38, Proposition 4.2]. Also, there are knots with Turaev genus one that are conjectured to have arbitrarily large alternation number, and the existence of such knots would show that Theorem 4 does not immediately follow from Theorems 2 and 3.

We first give examples of how Theorems 2 and 3 can be used, and then we prove each result.

4.2 Examples

Either side of the inequalities in Theorems 2 and 3 can provide insight into the other. Example 1 gives a family of knots all of whose alternation numbers are one, but whose Khovanov homology becomes more and more complicated in terms of width. Despite having complicated Khovanov homology, Theorem 2 implies that the Lee spectral sequence for this family of knots collapses at or before the second page, and Theorem 3 implies that the Bar-Natan spectral sequence collapses at or before the third page.

We remark that Alishahi and Dowlin [2, Theorem 1.3] proved that if the unknotting number of a (nontrivial) knot is one or two, then the Lee spectral sequence collapses at the second page. However, many knots in Example 1 have unknotting number greater than two, and thus the results from [2] cannot be used for those knots.

Examples 2, 3, and 4 describe knots where the page at which the relevant spectral sequence collapses gives a nontrivial lower bound on the alternation number of the knot.

Before describing the examples in detail, we remind the reader of some of the properties of the Lee and Bar-Natan spectral sequences. The map on the E_r page of the Lee spectral sequence increases the homological grading by one and the polynomial grading by $4r$. Similarly, the map on the E_r page of the Bar-Natan spectral sequence increases the homological grading by one and the polynomial grading by $2r$. Khovanov homology with $\mathbb{F}_2$ coefficients splits as a direct sum of two copies of the reduced Khovanov homology with $\mathbb{F}_2$ coefficients; that is, $Kh^{i,j}(K;\mathbb{F}_2) \cong \widetilde{Kh}^{i,j-1}(K;\mathbb{F}_2) \oplus \widetilde{Kh}^{i,j+1}(K;\mathbb{F}_2)$. The Bar-Natan spectral sequence has this same behavior of splitting into two copies; see [53, Subsection 3.3] for details.

Example 1 For any pair of positive integers m and n, de los Angeles Hernandez [15, Section 3] constructed the hyperbolic knot $K(m,n)$ whose diagram is depicted in Fig. 2 and whose alternation number is one. Therefore, Theorem 2 implies that the Lee spectral sequence of $K(m,n)$ collapses at or before the second page, and Theorem 3 implies that the Bar-Natan spectral sequence of $K(m,n)$ collapses at or before the third page.

Moreover, the width of the Khovanov homology of $K(m,n)$, that is the fewest number of adjacent $j-2i$ diagonals supporting $Kh(K(m,n))$, is $n+2$ [15, Lemma 3.2].

Recall that if the unknotting number of a (nontrivial) knot is one or two, then the Lee spectral sequence collapses at the second page [2, Theorem 1.3]. If $n+2<m$, then one can see that $K(m,n)$ has unknotting number greater than two, as follows. Dasbach and Lowrance [14, Proposition 5.3] proved that the signature of a knot K with diagram D satisfies the inequality

$$s_A(D) - c_+(D) - 1 \leq \sigma(K) \leq -s_B(D) + c_-(D) + 1,$$

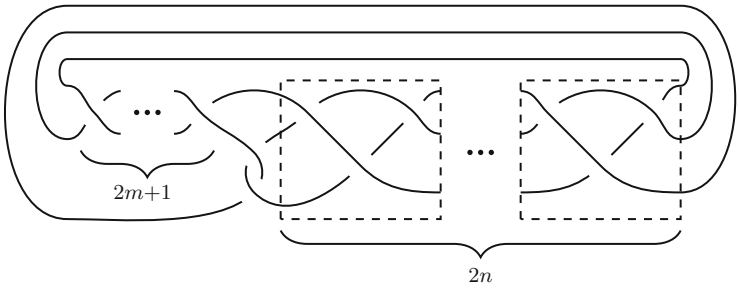
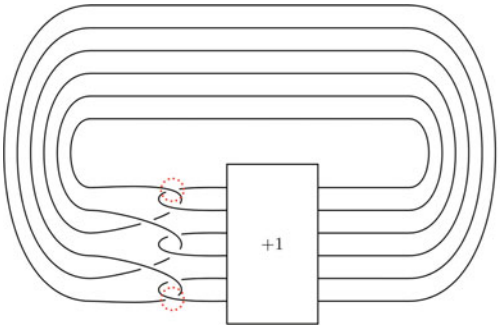


Fig. 2 A diagram of the knot $K(m, n)$



$j\backslash i$	0	1	2
5	4	1	
3	4	1	
1	1	1	
−1	1		
−3			

Fig. 3 The knot on the left has a positive full twist in the rectangle labeled +1. A portion of its Khovanov homology with $\mathbb{Q}$ coefficients is on the right. The highlighted yellow generator survives to the third page of the spectral sequence but not to the E_∞ page

where $s_A(D)$ and $s_B(D)$ are the number of components in the all-A and all-B Kauffman states, respectively, and $c_+(D)$ and $c_-(D)$ are the number of positive and negative crossings in D. Applying this inequality to the diagram of $K(m, n)$, we see that $-2m-2n \leq \sigma(K) \leq -2m+2n$. Because $|\sigma(K)| \leq 2u(K)$, if $n+2 < m$, then $u(K(m, n)) > 2$. Hence, Theorem 1.3 from [2] cannot be used for knots $K(m, n)$ with $n + 2 < m$.

In Examples 2, 3, and 4, we show the Khovanov homology of certain knots. The number in the (i, j) entry of the table in Fig. 3 is the rank of $Kh^{i,j}(K; R)$. All Khovanov homology computations for these examples are obtained using the program `JavaKh-v2` available on the Knot Atlas [5].

Example 2 Manolescu and Marengon [41, Theorem 2.1] gave an example of a knot K whose Lee spectral sequence over $\mathbb{Q}$ does not collapse at the second page. This knot K and a portion of its Khovanov homology $Kh(K, \mathbb{Q})$ appear in Fig. 3. Because $Kh^{1,1}(K; \mathbb{Q})$ is nontrivial, while $Kh^{0,-3}(K; \mathbb{Q})$ and $Kh^{2,5}(K; \mathbb{Q})$ are trivial, it follows that $\mathrm{pg}_{\mathrm{Lee}}(K; \mathbb{Q}) > 2$. Changing the two crossings of K circled in Fig. 3 transforms the knot into the figure-eight knot, and thus $\mathrm{alt}(K) \leq 2$. Using now Theorem 2, it follows that $\mathrm{alt}(K) = 2$.

Table 1 The Khovanov homology of $T_{5,6}$ with $\mathbb{F}_3$ coefficients. The highlighted yellow generator survives to the third page of the spectral sequence, but not to the E_∞ page

$j\backslash i$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
43														1	1
41													1	1	
39														1	
37												2	1		
35										2			1		
33								1		1	1				
31						1		1	2						
29						1	1		1						
27				1	1		1								
25					1										
23			1												
21	1														
19	1														

Example 3 The Lee spectral sequence for the $(5, 6)$-torus knot $T_{5,6}$ with $\mathbb{Q}$ coefficients collapses at the second page; however, this is not the case when the coefficients are $\mathbb{F}_3$. Table 1 shows the Khovanov homology of $T_{5,6}$ with $\mathbb{F}_3$ coefficients. Because $Kh^{13,43}(T_{5,6}; \mathbb{F}_3)$ is nontrivial while $Kh^{12,39}(T_{5,6}; \mathbb{F}_3)$ and $Kh^{14,47}(T_{5,6}; \mathbb{F}_3)$ are trivial, it follows that $\mathrm{pg}_{\mathrm{Lee}}(T_{5,6}; \mathbb{F}_3) > 2$. Theorem 2 implies that $2 \leq d_{\mathrm{thin}}(T_{5,6}; \mathbb{F}_3) \leq \mathrm{alt}(T_{5,6})$.

Example 4 The Khovanov homology of $T_{7,8}$ with $\mathbb{F}_2$ coefficients is shown in Table 2. Since $i = 26$ is the maximum homological grading where $Kh^{i,j}(T_{7,8}; \mathbb{F}_2)$ is nontrivial, the summands $Kh^{26,79}(T_{7,8}; \mathbb{F}_2)$ and $Kh^{26,81}(T_{7,8}; \mathbb{F}_2)$ must be paired with the summands $Kh^{25,75}(T_{7,8}; \mathbb{F}_2)$ and $Kh^{25,77}(T_{7,8}; \mathbb{F}_2)$ on the third page of Bar-Natan spectral sequence. Consequently, the summands $Kh^{25,79}(T_{7,8}; \mathbb{F}_2)$ and $Kh^{25,81}(T_{7,8}; \mathbb{F}_2)$ must be paired with the summands $Kh^{24,71}(T_{7,8}; \mathbb{F}_2)$ and $Kh^{24,73}(T_{7,8}; \mathbb{F}_2)$ on the fourth page of the Bar-Natan spectral sequence. Therefore $\mathrm{pg}_{\mathrm{BN}}(T_{7,8}) \geq 4$, and thus Theorem 3 implies that $2 \leq d_{\mathrm{thin}}(T_{7,8}; \mathbb{F}_2) \leq \mathrm{alt}(T_{7,8})$.

Table 2 The Khovanov homology $Kh(T_{7,8}; \mathbb{F}_2)$ of $T_{7,8}$. The highlighted yellow generators survive to the E_3 page, and the highlighted red generators survive to the E_4 page of the Bar-Natan spectral sequence

$j\backslash i$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
81																										1	1
79																										1	1
77																							1	1	1	1	
75																							2	2	1	1	
73																					2	3	2	2	1		
71																			1	2	4	5	2	1	1		
69																	1	1	2	5	3	2	1				
67																1	3	5	4	4	2						
65															1	4	3	5	4	1	1						
63													1	3	3	5	3	1	1								
61												1	2	6	4	2	2										
59									1	1	3	3	3	2													
57									3	3	3	3															
55							1	1	2	3	1	1															
53						1	1	2	2		1																
51						1	2	1	1																		
49				1	1		1																				
47			1	1	1																						
45			1																								
43	1																										
41	1																										

4.3 Proofs

The Lee and Bar-Natan spectral sequences both arise as spectral sequences of filtered complexes. The filtration comes from adding the Khovanov differential to different boundary maps that increase the polynomial/quantum grading. The Lee and Bar-Natan spectral sequences arise from maps $d_{\text{Lee}} : CKh^{i,j}(D; R) \to CKh^{i+1,j+4}(D; R)$ and $d_{\text{BN}} : CKh^{i,j}(D; \mathbb{F}_2) \to CKh^{i+1,j+2}(D; \mathbb{F}_2)$, respectively. For any knot diagram D, the homology of $(CKh(D; R), d + d_{\text{Lee}})$ is isomorphic to $R \oplus R$ situated in homological grading zero, and similarly, the homology of $(CKh(D; \mathbb{F}_2), d + d_{\text{BN}})$ is isomorphic to $\mathbb{F}_2 \oplus \mathbb{F}_2$ situated in homological grading zero.

Bar-Natan [4, Subsection 9.3] constructed a deformation of Khovanov homology using coefficients in $\mathbb{F}_2[h]$ for a formal variable h instead of $\mathbb{F}_2$ and using the differential $d + hd_{\text{BN}}$ instead of the usual Khovanov differential d. Turner later viewed the Bar-Natan construction through the lens of spectral sequences as described above. Alishahi and Dowlin [2, Subsection 2.2] similarly encapsulated the Lee endomorphism as part of a deformed complex with coefficients in

$\mathbb{Q}[X,t]/(X^2=t)$, where the differential in this complex is $d+td_{\text{Lee}}$. Just as with Lee's endomorphism, one can replace $\mathbb{Q}$ with $\mathbb{F}_p$, for any odd p, and all of the results of [2] hold without changing their proofs.

An element α in the homology of Bar-Natan's complex is *h-torsion of order n* if $h^n\alpha=0$ but $h^{n-1}\alpha\neq 0$. Let $\mathfrak{u}_h(K)$ be the maximum order of any torsion element in the homology of Bar-Natan's complex. Then $\mathfrak{u}_h(K)+1=\text{pg}_{\text{BN}}(K)$ [1, Lemma 3.2].

Similarly, an element α in the deformed Lee homology over $R=\mathbb{Q}$ or $\mathbb{F}_p$, for an odd prime p, is *X-torsion of order n* (respectively *t-torsion of order m*) if $X^n\alpha=0$ but $X^{n-1}\alpha\neq 0$ (respectively $t^m\alpha=0$ but $t^{m-1}\alpha\neq 0$). Alishahi and Dowlin proved the following facts about $\mathfrak{u}_X(K;\mathbb{Q})$ and $\mathfrak{u}_t(K;\mathbb{Q})$. We observed that the proofs of these facts when $R=\mathbb{Q}$ also apply when using $\mathbb{F}_p$ coefficients. As such, we state the following for $R=\mathbb{Q}$ or $\mathbb{F}_p$, where p is an odd prime.

1. If $Kh(K;R)$ is homologically thin, then $\mathfrak{u}_X(K;R)=1$;
2. $|\mathfrak{u}_X(K_+;R)-\mathfrak{u}_X(K_-;R)|\leq 1$, where K_+ and K_- are knots differing by a single crossing change;
3. $\lceil \mathfrak{u}_X(K;R)/2\rceil=\mathfrak{u}_t(K;R)$, and
4. $\mathfrak{u}_t(K)+1=\text{pg}_{\text{Lee}}(K;R)$.

We are now in a position to prove Theorems 2, 3, and 4.

Proof of Theorem 2 Let $d_{\text{thin}}(K;R)=d$. Hence, there is a sequence of knots $K=K_0,K_1,\ldots,K_d$ such that K_{i+1} is obtained from K_i via a crossing change for all $i=0,\ldots,d-1$, and $Kh(K_d;R)$ is homologically thin. Item (1) above implies that $\mathfrak{u}_X(K_d;R)=1$, and item (2) implies that $\mathfrak{u}_X(K;R)\leq d+1$. Then item (3) implies that $\mathfrak{u}_t(K;R)=\left\lceil\frac{\mathfrak{u}_X(K;R)}{2}\right\rceil\leq\left\lceil\frac{d+1}{2}\right\rceil$. Finally, item (4) implies that $\text{pg}_{\text{Lee}}(K;R)=\mathfrak{u}_t(K;R)+1\leq\left\lceil\frac{d+3}{2}\right\rceil$, as desired. The second inequality in the theorem follows at once from the fact that $d_{\text{thin}}(K;R)\leq\text{alt}(K)$, as seen in the beginning of this section.

Proof of Theorem 3 Let $d_{\text{thin}}(K;\mathbb{F}_2)=d$. Hence there is a sequence of knots $K=K_0,K_1,\ldots,K_d$ such that K_{i+1} is obtained from K_i via a crossing change for $i=0,\ldots,d-1$, and $Kh(K_d;\mathbb{F}_2)$ is homologically thin. By Alishahi [1, Corollary 3.3], since K_i and K_{i+1} differ by a crossing change, it follows that $|\mathfrak{u}_h(K_i)-\mathfrak{u}_h(K_{i+1})|\leq 1$, and thus $\mathfrak{u}_h(K)\leq d+\mathfrak{u}_h(K_d)$. Since $Kh(K_d;\mathbb{F}_2)$ is homologically thin, $\text{pg}_{\text{BN}}(K_d)\leq 2$. But $\text{pg}_{\text{BN}}(K_d)=\mathfrak{u}_h(K_d)+1$, and thus $\mathfrak{u}_h(K_d)\leq 1$. It follows that $\mathfrak{u}_h(K)\leq d+1$, and therefore $\text{pg}_{\text{BN}}(K)\leq d+2$, as desired.

Proof of Theorem 4 The *width* $w(Kh(K;R))$ of the Khovanov homology over a ring R is defined as

$$w(Kh(K;R))=1+\frac{1}{2}\left(\max\{j-2i\mid Kh^{i,j}(K;R)\neq 0\}-\min\{j-2i\mid Kh^{i,j}(K;R)\neq 0\}\right).$$

Champanerkar, Kofman, and Stoltzfus [11, Corollary 3.1] proved that $w(Kh(K;R)) \leq g_T(K)+2$. Since the Lee differential on the E_r page increases the homological grading i by one and the polynomial grading j by $4r$, if $\mathrm{pg}_{\mathrm{Lee}}(K;R) = n$, then $w(Kh(K;R)) \geq 2n-2$. Therefore $2\,\mathrm{pg}_{\mathrm{Lee}}(K;R) \leq g_T(K)+4$, as desired. Similarly, since the Bar-Natan differential on the E_r page increases the homological grading by one and the polynomial grading by $2r$, if $\mathrm{pg}_{\mathrm{BN}}(K) = n$, then $w(Kh(K;\mathbb{F}_2)) \geq n$. Therefore, $\mathrm{pg}_{\mathrm{BN}}(K) \leq g_T(K)+2$.

Acknowledgments We would like to thank ICERM and the organizers of the Women in Symplectic and Contact Geometry and Topology workshop (WiSCon) for the opportunity to work on this project during the summer of 2019. C. Caprau was partially supported by Simons Foundation Collaboration Grant 355640. N. González was partially supported by a fellowship from the Mathematical Sciences Research Institute. C. Lee was supported in part by NSF Grant DMS 1907010. A. Lowrance was supported in part by NSF Grant DMS 1811344. R. Sazdanović was partially supported by the Simons Foundation Collaboration Grant 318086 and NSF Grant DMS 1854705.

References

1. Akram Alishahi. Unknotting number and Khovanov homology. *Pacific J. Math.*, 301(1):15–29, 2019.
2. Akram Alishahi and Nathan Dowlin. The Lee spectral sequence, unknotting number, and the Knight Move Conjecture. *Topology Appl.*, 254:29–38, 2019.
3. John A. Baldwin. On the spectral sequence from Khovanov homology to Heegaard Floer homology. *Int. Math. Res. Not. IMRN*, (15):3426–3470, 2011.
4. Dror Bar-Natan. Khovanov's homology for tangles and cobordisms. *Geom. Topol.*, 9:1443–1499, 2005.
5. Dror Bar-Natan, Scott Morrison, and et al. The Knot Atlas. http://katlas.org.
6. Joshua Batson and Cotton Seed. A link-splitting spectral sequence in Khovanov homology. *Duke Math. J.*, 164(5):801–841, 2015.
7. Christian Blanchet, Nathan Habegger, Gregor Masbaum, and Pierre Vogel. Topological quantum field theories derived from the Kauffman bracket. *Topology*, 34(4):883–928, 1995.
8. Jonathan M. Bloom. Odd Khovanov homology is mutation invariant. *Math. Res. Lett.*, 17(1):1–10, 2010.
9. Carmen Caprau. The universal $sl(2)$ cohomology via webs and foams. *Topology Appl.*, 156(9):1684–1702, 2009.
10. Carmen Caprau. On the quantum filtration of the universal $sl(2)$ foam cohomology. *J. Knot Theory Ramifications*, 21(3):1250023, 2012.
11. Abhijit Champanerkar, Ilya Kofman, and Neal Stoltzfus. Graphs on surfaces and Khovanov homology. *Algebr. Geom. Topol.*, 7:1531–1540, 2007.
12. David Clark. Functoriality for the su_3 Khovanov homology. *Algebr. Geom. Topol.*, 9(2):625–690, 2009.
13. D. Cooper and W. B. R. Lickorish. Mutations of links in genus 2 handlebodies. *Proc. Amer. Math. Soc.*, 127(1):309–314, 1999.
14. Oliver T. Dasbach and Adam M. Lowrance. Turaev genus, knot signature, and the knot homology concordance invariants. *Proc. Amer. Math. Soc.*, 139(7):2631–2645, 2011.
15. María de los Angeles Guevara Hernández. Infinite families of hyperbolic prime knots with alternation number 1 and dealternating number n. *J. Knot Theory Ramifications*, 26(10):1750055, 12, 2017.

16. Nathan M. Dunfield, Stavros Garoufalidis, Alexander Shumakovitch, and Morwen Thistlethwaite. Behavior of knot invariants under genus 2 mutation. *New York J. Math.*, 16:99–123, 2010.
17. C.McA. Gordon. Ribbon concordance of knots in the 3-sphere. *Mathematische Annalen*, 257:157–170, 1981.
18. V. F. R. Jones. Hecke algebra representations of braid groups and link polynomials. *Ann. of Math. (2)*, 126(2):335–388, 1987.
19. Sungkyung Kang. Link homology theories and ribbon concordances. *arXiv preprint arXiv:1909.06969*, 2019.
20. Anton Kapustin and Yi Li. Topological correlators in Landau-Ginzburg models with boundaries. *Adv. Theor. Math. Phys.*, 7(4):727–749, 2003.
21. Akio Kawauchi. On alternation numbers of links. *Topology Appl.*, 157(1):274–279, 2010.
22. Mikhail Khovanov. A categorification of the Jones polynomial. *Duke Math. J.*, 101(3):359–426, 2000.
23. Mikhail Khovanov. $sl(3)$ link homology. *Algebr. Geom. Topol.*, 4:1045–1081, 2004.
24. Mikhail Khovanov. Link homology and Frobenius extensions. *Fund. Math.*, 190:179–190, 2006.
25. Mikhail Khovanov and Louis-Hadrien Robert. Link homology and Frobenius extensions II. *arXiv preprint arXiv:2005.08048*, 2020.
26. Mikhail Khovanov and Lev Rozansky. Matrix factorizations and link homology. *Fund. Math.*, 199(1):1–91, 2008.
27. Mikhail Khovanov and Lev Rozansky. Matrix factorizations and link homology. II. *Geom. Topol.*, 12(3):1387–1425, 2008.
28. P. B. Kronheimer and T. S. Mrowka. Gauge theory for embedded surfaces. I. *Topology*, 32(4):773–826, 1993.
29. Eun Soo Lee. An endomorphism of the Khovanov invariant. *Adv. Math.*, 197(2):554–586, 2005.
30. Adam Simon Levine and Ian Zemke. Khovanov homology and ribbon concordances. *Bull. Lond. Math. Soc.*, 51(6):1099–1103, 2019.
31. Lukas Lewark. Rasmussen's spectral sequences and the sl_N-concordance invariants. *Adv. Math.*, 260:59–83, 2014.
32. Robert Lipshitz and Sucharit Sarkar. A Khovanov stable homotopy type. *J. Amer. Math. Soc.*, 27(4):983–1042, 2014.
33. Robert Lipshitz and Sucharit Sarkar. A Steenrod square on Khovanov homology. *J. Topol.*, 7(3):817–848, 2014.
34. Andrew Lobb. A slice genus lower bound from sl(n) Khovanov-Rozansky homology. *Adv. Math.*, 222(4):1220–1276, 2009.
35. Andrew Lobb. A note on Gornik's perturbation of Khovanov-Rozansky homology. *Algebr. Geom. Topol.*, 12(1):293–305, 2012.
36. Andrew Lobb and Liam Watson. A refinement of Khovanov homology. *Geom. Topol.*, 25:1861–1917, 2021.
37. Andrew Lobb and Raphael Zentner. On spectral sequences from Khovanov homology. *Algebr. Geom. Topol.*, 20(2):531–564, 2020.
38. Adam M. Lowrance. Alternating distances of knots and links. *Topology Appl.*, 182:53–70, 2015.
39. Marco Mackaay, Marko Stošić, and Pedro Vaz. $sl(N)$-link homology ($N \geq 4$) using foams and the Kapustin-Li formula. *Geom. Topol.*, 13(2):1075–1128, 2009.
40. Marco Mackaay and Pedro Vaz. The universal sl_3-link homology. *Algebr. Geom. Topol.*, 7:1135–1169, 2007.
41. Ciprian Manolescu and Marco Marengon. The Knight Move Conjecture is false. *Proc. Amer. Math. Soc.*, 148(1):435–439, 2020.
42. H. R. Morton. Mutant knots with symmetry. *Math. Proc. Cambridge Philos. Soc.*, 146(1):95–107, 2009.
43. H. R. Morton and H. J. Ryder. Mutants and $SU(3)_q$ invariants. In *The Epstein birthday schrift*, volume 1 of *Geom. Topol. Monogr.*, pages 365–381. Geom. Topol. Publ., Coventry, 1998.

44. H. R. Morton and P. Traczyk. The Jones polynomial of satellite links around mutants. In *Braids (Santa Cruz, CA, 1986)*, volume 78 of *Contemp. Math.*, pages 587–592. Amer. Math. Soc., Providence, RI, 1988.
45. Peter Ozsváth and Zoltán Szabó. On the Heegaard Floer homology of branched double-covers. *Adv. Math.*, 194(1):1–33, 2005.
46. Lisa Piccirillo. The Conway knot is not slice. *Ann. of Math. (2)*, 191(2):581–591, 2020.
47. Jacob Rasmussen. Khovanov homology and the slice genus. *Invent. Math.*, 182(2):419–447, 2010.
48. Daniel Ruberman. Mutation and volumes of knots in S^3. *Invent. Math.*, 90(1):189–215, 1987.
49. Sucharit Sarkar. Ribbon distance and Khovanov homology. *Algebr. Geom. Topol.*, 20(2):1041–1058, 2020.
50. Cotton Seed. Computations of the Lipshitz-Sarkar Steenrod square on Khovanov homology. *arXiv preprint arXiv:1210.1882*, 2012.
51. Alexander N. Shumakovitch. Torsion of Khovanov homology. *Fund. Math.*, 225(1):343–364, 2014.
52. V. G. Turaev. A simple proof of the Murasugi and Kauffman theorems on alternating links. *Enseign. Math. (2)*, 33(3–4):203–225, 1987.
53. Paul R. Turner. Calculating Bar-Natan's characteristic two Khovanov homology. *J. Knot Theory Ramifications*, 15(10):1335–1356, 2006.
54. Stephan M. Wehrli. Khovanov homology and Conway mutation. *arXiv preprint arXiv:math/0301312*, 2003.
55. Stephan M. Wehrli. Mutation invariance of Khovanov homology over $\mathbb{F}_2$. *Quantum Topol.*, 1(2):111–128, 2010.
56. Hao Wu. On the quantum filtration of the Khovanov-Rozansky cohomology. *Adv. Math.*, 221(1):54–139, 2009.
57. Ian Zemke. Knot Floer homology obstructs ribbon concordance. *Ann. of Math. (2)*, 190(3):931–947, 2019.

Braids, Fibered Knots, and Concordance Questions

Diana Hubbard, Keiko Kawamuro, Feride Ceren Kose, Gage Martin, Olga Plamenevskaya, Katherine Raoux, Linh Truong, and Hannah Turner

1 Introduction

In this article, we will discuss a number of results, observations and questions concerning the relation between topological properties of knots and certain surface diffeomorphisms associated to these knots. The relevant surface diffeomorphisms arise in two ways.

D. Hubbard (✉)
Department of Mathematics, Brooklyn College, Brooklyn, NY, USA
e-mail: diana.hubbard@brooklyn.cuny.edu

K. Kawamuro
Department of Mathematics, University of Iowa, Iowa City, IA, USA
e-mail: keiko-kawamuro@uiowa.edu

F. C. Kose · H. Turner
Department of Mathematics, UT Austin, Austin, TX, USA
e-mail: fkose@math.utexas.edu; hannahturner@math.utexas.edu

G. Martin
Department of Mathematics, Boston College, Chestnut Hill, MA, USA
e-mail: martaic@bc.edu

O. Plamenevskaya
Department of Mathematics, Stony Brook University, Stony Brook, NY, USA
e-mail: olga@math.stonybrook.edu

K. Raoux
Department of Mathematics, Michigan State University, East Lansing, MI, USA
e-mail: raouxkat@msu.edu

L. Truong
School of Mathematics, Institute for Advanced Study, Princeton, NJ, USA
e-mail: ltruong@math.ias.edu

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9_7

1. For a fibered knot $K \subset S^3$, we consider the fibration of the knot complement $S^3 \backslash K$ over S^1. The fiber surface and the monodromy of this fibration are uniquely determined by K (see [12, p.34], [57, Corollary 8.3]).
2. An arbitrary knot in S^3 can be (non-uniquely) represented as a braid closure, and we can study the corresponding braid monodromy. Recall that Artin's braid group B_n on n-strands [4] is the mapping class group of an n-punctured disc D_n.

Working from these different perspectives, we will explore how the fiber monodromy or braid monodromy is related to invariants of the underlying knot, such as the Seifert genus $g_3(K)$ and the slice genus $g_4(K)$. As usual, $g_3(K)$ is the smallest genus of a Seifert surface the knot K bounds in the 3-sphere, and $g_4(K)$ is the smallest genus of a smooth oriented embedded surface that the knot K bounds in the 4-ball.

Intuitively, the more full twists that a braid β contains, the more complicated its closure $\widehat{\beta} = K$ is. Let $\Delta^2 = (\sigma_1 \dots \sigma_{n-1})^n$ denote a full twist on n strands. We expect the braid closure's genus $g_3(K)$ and slice genus $g_4(K)$ to grow larger as more full twists are added to the braid. This idea is made precise in the following proposition.

Proposition 1 *Let $\beta \in B_m$ be a braid, and let K_n be the closure of the braid $\beta\Delta^{2n}$. When n becomes large, $g_3(K_n) = g_4(K_n)$ and this value grows roughly as $\frac{1}{2}nm^2$.*

Proposition 1 gives us an asymptotic understanding of the behavior of the genus and slice genus as full twists are added. As for non-asymptotic behavior, for braids of three or more strands, we use Sato's work [56] to show concatenating full twists to a braid will change the concordance class of its closure.

Proposition 2 *Let $\beta \in B_m$ be a braid, $m \geq 3$, and $\beta_k = \beta\Delta^{2k}$, $k \in \mathbb{Z}$, $k > 0$. Then all braid closures $\widehat{\beta}$, $\widehat{\beta_k}$, $k > 0$ lie in pairwise distinct concordance classes.*

We now focus on the boundary twisting of the monodromies. The amount of boundary twisting is quantified by the *fractional Dehn twist coefficient*, defined in Sect. 2.2. The idea of the fractional Dehn twist coefficient, or FDTC, first appeared in [19] in the context of essential laminations. In the context of open books and contact topology, it was developed and applied in [22, 23], and explored further by many authors, [7, 24, 30, 35]. For classical braids, a similar notion (via a somewhat different approach) was studied in [42]. The fractional Dehn twist coefficient was also studied from the braid- and knot-theoretic perspective in [13, 16, 52] among others. A generalization of the fractional Dehn twist coefficient to the case of braids in arbitrary open books, and a detailed proof that different definitions are equivalent, is given in [31].

Notation We will use notation $FT(K)$ for the fractional Dehn twist coefficient of the fibered monodromy for a fibered knot K in S^3. We use $BT(\beta)$ for the fractional Dehn twist coefficient of a braid monodromy.

We will explore how the fractional Dehn twist coefficient of the braid monodromy is related to invariants of the underlying knot, such as the Seifert genus

$g_3(K)$ and the slice genus $g_4(K)$. Given a braid representative β whose closure is a knot $K = \widehat{\beta}$, we seek lower bounds for the genus $g_3(K)$ and the slice genus $g_4(K)$ in terms of the fractional Dehn twist coefficient $BT(\beta)$.

In [32] Ito used braid foliations to give a bound for the genus of a knot K in terms of the fractional Dehn twist coefficient $BT(\beta)$ of its braid representative β. See Theorem 8 for a more precise statement. Ito's theorem serves as motivation for our work.

Theorem 1 ([32]) *Let β be a braid whose closure is isotopic to a knot $K \subset S^3$. Then*

$$|BT(\beta)| \leq g_3(K) + 2.$$

For quasipositive braids, we provide a new upper bound on $BT(\beta)$ in terms of the slice genus in the following theorem.

Theorem 2 *Let $n \geq 3$. Let $\beta \in B_n$ be a quasipositive n-braid whose braid closure $\hat{\beta}$ is a knot then*

$$BT(\beta) \leq 2g_4(\hat{\beta}) + (n-2).$$

The upper bound on $BT(\beta)$ is sharp by Corollary 2, which constructs slice knots ($g_4(K) = 0$) represented by quasipositive n-braids β with $BT(\beta) = n - 2$.

The analogue of Ito's Theorem 1 with genus $g_3(K)$ replaced by slice genus $g_4(K)$ does not hold by Corollary 2 and Proposition 10. Notably, the non-quasipositive knots of Proposition 10 satisfy the same inequality as Theorem 2, leading us to ask whether such a bound holds in general.

Question 1 For any n-strand β representing a knot K, is it always true that

$$|BT(\beta)| \leq 2g_4(K) + n - 2?$$

If not, is there a bound of the order

$$|BT(\beta)| \leq C(n)g_4(K) + D(n),$$

where $C(n)$, $D(n)$ are constant for each fixed n?

As evidence towards a positive answer to Question 1, for 3-braids representing knots of finite order in the concordance group, we provide an upper bound on $BT(\beta)$ in the following theorem.

Theorem 3 *Let K be a knot that can be represented as the closure of a braid in B_3. Suppose further that K is slice, or more generally, that K has a finite order in the concordance group. Then any 3-braid representative β of K satisfies $|BT(\beta)| \leq 1$.*

The proof of Theorem 3 relies on Murasugi's classification of three-braids [45] and work of Baldwin [5] and is given in Sect. 6. Section 6 concludes with a discussion on computational evidence for the bounds in Question 1.

One might hope to use the Ozsváth-Szabo τ-invariant [47] or the Rasmussen s-invariant [54] to construct an upper bound on the fractional Dehn twist coefficient BT of a braid. We show that this idea fails in Proposition 3, by constructing a family of braids with small τ-invariant and small s-invariant but large BT.

Proposition 3 *Let $\beta = (\Delta^2)^k \sigma_1^{-1} \sigma_2^{-(6k-1)}$ in B_3. Then $\hat{\beta}$ is a knot, $BT(\beta) \geq k-1$, $|\tau(\hat{\beta})| \leq 1$ and $|s(\hat{\beta})| \leq 2$.*

It is important to note that $BT(\beta)$ is very sensitive to braid stabilizations. Recall that the positive braid stabilization of $\beta \in B_n$ is the braid $\beta_+ = \beta\sigma_n \in B_{n+1}$, and the negative stabilization is the braid $\beta_- = \beta\sigma_n^{-1} \in B_{n+1}$. The braid closures of β, β_+ and β_- are all isotopic to the same link. Proposition 9 shows that stabilized braid representatives have fractional Dehn twist coefficient $BT(\beta_+)$ and $BT(\beta_-)$ that are bounded between -1 and 1. Thus, a stabilized braid representative for a knot K will not provide a useful lower bound for the genus $g_3(K)$ or slice genus $g_4(K)$. To our knowledge it is an open question how to determine which braid representative has largest possible fractional Dehn twist coefficient.

Finally, we explore the relationship between the fractional Dehn twist coefficient $FT(K)$ of a fibered knot K and other topological invariants of K, including its knot Floer stable equivalence class in Sect. 7 and slice genus in Sect. 8.

2 Background on Fractional Dehn Twist Coefficient and Braids

We begin with background on the braid group and the fractional Dehn twist coefficient and collect some previous results that inspired our work.

2.1 The Braid Group

Recall that Artin's braid group B_n on n-strands [4] is the mapping class group of an n-punctured disc D_n. It admits the presentation

$$B_n = \left\langle \sigma_1, \ldots, \sigma_{n-1} \,\middle|\, \begin{array}{l} \sigma_i\sigma_j = \sigma_j\sigma_i : |i-j| \geq 2 \\ \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} : 1 \leq i \leq n-2 \end{array} \right\rangle.$$

The σ_i's are usually referred to as the *standard* or Artin generators of B_n. A *n-braid* β is an element of B_n. The braid β can be considered as a proper embedding of n strands in $D^2 \times I$, so that each generator σ_i corresponds to a positive half-twist

between the i-th and $(i+1)$-th strands. When $D^2 \times \{0\}$ is glued to $D^2 \times \{1\}$ by the identity map, we obtain a link in $D^2 \times S^1 \subset S^3$ called the closure of β, denoted $\hat{\beta}$. A classical theorem of Alexander [3] allows to reverse this process (although not uniquely): every knot or link in S^3 can be represented as a closed braid. The *braid index* of a link is the minimum number n such that there exists an n-braid β whose closure represents the given link.

Thus, a classical braid can be considered from several viewpoints: (1) a braid is a "word" in standard generators σ_i of the braid group B_n; (2) the braid closure is a knot or link that can be studied by means of classical knot theory and by modern knot-homological invariants (such as knot Floer homology and Khovanov homology); (3) the braid monodromy action shows how the braid twists, and is related to ideas from geometry and dynamics of surface diffeomorphisms. There are many excellent resources for learning more about braids; for instance, see [6, 10], and [13].

2.2 Fractional Dehn Twist Coefficient

The amount of boundary twisting of a surface diffeomorphism is quantified by the *fractional Dehn twist coefficient* (FDTC), defined geometrically as follows. Let S be a compact oriented surface with connected boundary and $\chi(S) < 0$, and $\phi : S \to S$ a homeomorphism fixing the boundary ∂S pointwise. Using the Nielsen-Thurston classification, we can find a free isotopy $\Phi : S \times [0,1] \to S$ connecting ϕ to its Nielsen-Thurston representative, so that $\Phi_0 = \phi$ and Φ_1 is either periodic, reducible, or pseudo-Anosov. (Note that Φ_1 no longer fixes the boundary of S.) The fractional Dehn twist coefficient $FT(\phi)$ is defined as the winding number of the arc $\Phi(p \times [0,1])$ for a chosen basepoint $p \in \partial S$. It can be shown that $FT(\phi)$ depends only on the isotopy class of ϕ *rel* boundary, and is independent of the choice of basepoint p and isotopy Φ. Moreover, $FT(\phi)$ is always a rational number. Although we primarily focus on knots in S^3, one could also consider the FDTC for fibered knots in an arbitrary closed oriented 3-manifold Y (and for corresponding open books on Y).

The FDTC of a braid is defined by considering the braid monodromy as a boundary-fixing homeomorphism of the punctured disk, and taking the FDTC at the boundary of the disk. For surfaces with multiple boundary components, the FDTC can be defined for each component (the corresponding number measures twisting around the given component).

For a fibered knot K in S^3, there are different monodromies to consider: the monodromy of the fibration $S^3 \setminus K \to S^1$ as well as the monodromy of a braid representative β of K. These monodromies live on different surfaces and have very different fractional Dehn twist coefficients. In particular, the FDTC of the fibration is fixed, while the FDTC of the braid monodromy depends on the choice of the braid representative. As in Notation 1 we will use notation $FT(K)$ for the FDTC of the fibration monodromy and $BT(\beta)$ for the FDTC of the braid

monodromy. An example of how these two values can be different comes from the torus knots $T_{p,q}$. Since these are fibered in S^3, by work of Gabai and Kazez-Roberts, $|FT(T_{p,q})| \leq \frac{1}{2}$ (see Theorem 5 below). However, a braid representative of $T_{p,q}$ is the word $\beta_{p,q} = (\sigma_1 \cdots \sigma_{p-1})^q$ in B_p, and $|BT(\beta_{p,q})| = \frac{q}{p}$, which may be a large quantity. In general, the quantity $BT(\beta)$ can also be defined more algebraically, using braid orderings and Dehornoy's floor function, see Sect. 2.3.

We collect a few useful properties of the FDTC in the next proposition.

Proposition 4 ([30, 42]) *Let t_∂ denote the positive Dehn twist about a boundary-parallel curve. Then for any two boundary-fixing homeomorphisms ϕ, ψ of S, we have*

(a) (Quasimorphism) $|FT(\phi\psi) - FT(\phi) - FT(\psi)| \leq 1$.
(b) (Homogeneity) $FT(\phi^n) = nFT(\phi)$.
(c) (Behaviour under full twists) $FT(t_\partial \phi) = FT(\phi) + 1$.
(d) (Conjugacy invariant) $FT(\phi) = FT(\psi\phi\psi^{-1})$.

All of the above properties also hold for braids and their fractional Dehn twist coeficients BT. Additionally, if $\beta \in B_n$ is an n-braid, we have $BT(\beta) \in \{\frac{p}{q}$, where $p, q \in \mathbb{Z}, 1 \leq q \leq n\}$.

A related notion is that of a *right-veering* surface homeomorphism. As above, let S be a compact oriented surface with connected boundary, and $\phi : S \to S$ a homeomorphism fixing the boundary ∂S pointwise. Let $\alpha, \beta : [0, 1] \to S$ be arbitrary properly embedded oriented arcs with $\alpha(0) = \beta(0) = x \in \partial S$. Isotope α and β so that they intersect transversely with the fewest possible number of intersections. We say that β is to the right of α if either $(\dot{\beta}(0), \dot{\alpha}(0))$ define the orientation of S at x or $\beta = \alpha$. We now can define right-veering: we say ϕ is right-veering if for any choice of base point $x \in \partial S$ and properly embedded oriented arc α starting at x, $\phi(\alpha)$ is to the right of α. In a similar way, one defines left-veering maps. The maps that *do not* satisfy the corresponding properties are called non-right-veering resp. non-left-veering. Note that the identity map is both right-veering and left-veering (and this is the only isotopy class with both properties.)

Very roughly, one can think of the fractional Dehn twist coefficient as being a measurement of "how right-veering" a surface homeomorphism is. Indeed, when ϕ is irreducible, the FDTC of ϕ is greater than zero if and only if ϕ is right-veering: see [22, 35]. This gives a way to estimate the FDTCs:

Proposition 5 ([35, Corollary 2.6])

(1) If α is a properly embedded oriented arc starting on ∂S, then $\phi(\alpha)$ to the right of α implies $FT(\phi) \geq 0$, and $\phi(\alpha)$ to the left of α implies $FT(\phi) \leq 0$.
(2) If β is a braid with a braid word where a generator σ_i enters with positive exponents only, then $BT(\beta) \geq 0$. If there is a braid word for β where a generator σ_j enters with negative exponents only, then $BT(\beta) \leq 0$.

Part 2 of Proposition 5 follows by finding corresponding arcs: after a conjugacy, we can assume that σ_1 enters with positive exponents only; then the obvious arc

connecting the boundary of the disk to the first puncture is moved strictly to the right of itself. See also [42, Corollary 5.5, Proposition 13.1] for related statements and alternate proofs for braids.

Proposition 5 gives an easy way to show that $FT(\phi) = 0$ by finding two arcs, one moved to the right by ϕ and the other moved to the left. Similarly, $BT(\beta) = 0$ if we can find a braid word for β where one generator enters with positive exponents only, and the other with negative exponents only.

The right-veering property plays an important role in 3-dimensional contact topology: a contact structure is tight if and only if the monodromy of each compatible open book is right-veering [22]. Moreover, a contact structure supported by an open book with connected boundary and pseudo-Anosov monodromy with $FT \geq 1$ is isotopic to a perturbation of a taut foliation, and therefore is weakly symplectically fillable and has non-vanishing Heegaard Floer contact invariant (with twisted coefficients) [23]. For planar open books, the condition $FT > 1$ for every boundary component implies tightness of the supported contact structure [29]. In a similar vein, transverse braids in the standard contact S^3 must be right-veering if their Floer- or Khovanov-homological transverse invariants are non-zero [8, 9, 51, 52], and braids with $BT > 1$ have non-vanishing transverse hat-invariant in knot Floer homology [52].

A number of existing results connect the FDTC to knot invariants and topology of 3-manifolds. Gabai proved the following result concerning the genus of fibered knots with pseudo-Anosov monodromy:

Theorem 4 ([17]) *Let $K \subset S^3$ be a fibered knot. Suppose that its monodromy is either pseudo-Anosov or reducible with pseudo-Anosov near the boundary. Then either $FT(K) = 0$ or $FT(K) = 1/r$, where $2 \leq |r| \leq 4(g_3(K)) - 2$.*

Kazez and Roberts in Corollary 4.3 of [35] determined the possible FT values for the periodic case and reducible case with the monodromy periodic near the boundary. Putting together Theorem 4 with their work yields a very strict bound on the values of FT for fibered knots in S^3:

Theorem 5 ([17, 35, Theorem 4.5]) *Let $K \subset S^3$ be a fibered knot. Then $FT(K) = 0$ or $FT(K) = 1/n$, where n is an integer, $|n| \geq 2$. In particular, $-\frac{1}{2} \leq FT(K) \leq \frac{1}{2}$.*

In a different direction, Hedden and Mark [24] found an priori bound on the value of FT for any fibered knot in a fixed 3-manifold Y in terms of the dimension of the Heegaard Floer homology with $\mathbb{F} = \mathbb{Z}/2$ coefficients and the size of the torsion in singular first homology. As a corollary, they get a bound for the FDTC of classical braids, via open books on the branched double cover.

Theorem 6 ([24]) *Let Y be a closed oriented 3-manifold. Then for any fibered knot K in Y with monodromy ϕ, the FDTC satisfies*

$$|FT(\phi)| \leq \frac{1}{2}(\dim_{\mathbb{F}} \widehat{HF}(Y) - |\mathrm{Tor} H_1(Y; \mathbb{Z})|) + 1.$$

Corollary 1 ([24]) *Let β be an odd-strand braid representative of $K \subset S^3$, and let $\Sigma(K)$ denote the double-branched cover of K. Then*

$$|BT(\beta)| \leq \dim_{\mathbb{F}} \widehat{HF}(\Sigma(K)) - |H_1(\Sigma(K))| + 2.$$

Further, Hedden and Mark use the relation between the Khovanov homology of a link and the Heegaard Floer homology of its branched double cover to prove a bound on BT in terms of the rank of reduced Khovanov homology $\widetilde{Kh}$:

Theorem 7 ([24]) *Let L be a link in S^3, and let β c L with an odd number of strands. Then*

$$BT(\beta) \leq \dim_{\mathbb{F}} \widetilde{Kh}(-L) - |\det(L)| + 2.$$

2.3 Dehornoy's Braid Ordering

We will now describe another, more algebraic, approach to the fractional Dehn coefficient for braids, based on Dehornoy's braid ordering. The braid group B_m is known to be orderable, namely, there exists a *left-invariant* linear order on B_m, so that if $\beta_2 \succ \beta_1$, then $\gamma\beta_2 \succ \gamma\beta_1$ for any $\gamma \in B_m$. An ordering can be defined by considering the action of the braid monodromy $\beta \in Map(D, Q)$ on the on the "standard" punctured disk D, with the set of punctures Q on the y-axis, labeled $Q = \{p_1, p_2, \ldots p_m\}$ from bottom to top. (See Fig. 1 in Sect. 4.) Roughly, $\beta_2 \succ \beta_1$ iff β_2 twists the y-axis more to the right than β_1. These ideas can be traced back to W. Thurston and can be extended to obtain many different orderings (known as *Nielsen–Thurston* orderings).

Dehornoy [11] defined an ordering on B_m from an algebraic perspective. The algebraic definition is equivalent to the geometric one described above. We say that $\beta \succ 1$ iff the braid β admits a braid word that contains the generator σ_i but no σ_i^{-1} and no σ_j for $j < i$. (A word of this form is called σ_i-positive, which has the following geometric interpretation. Suppose we apply a σ_i-positive word β to the punctured disk and pull the image of the y-axis taut while fixing the punctures. Then the image of the y-axis under β will first diverge from y-axis at a point between the $(i-1)$'st and i'th punctures and at this point, the image will go to the right of the y-axis.) Then, for $\beta, \beta' \in B_m$ we define $\beta \succ \beta'$ if $(\beta')^{-1}\beta \succ 1$. From the algebraic perspective, checking the basic properties of the ordering is highly non-trivial; from the geometric perspective, it is not hard to see that we get a well-defined linear order, [15]. The algebraic approach becomes useful if one wants to study combinatorial braid invariants. For example, it follows from [15] that a non-right-veering braid is conjugate to a braid with a braid word where a generator σ_i enters with negative exponents only (compare with Proposition 5). The specific property of the braid word allows to relate geometry and combinatorics; this approach was used in [8, 52].

Let $\Delta = (\sigma_1\sigma_2\ldots\sigma_{n-1})(\sigma_1\sigma_2\ldots\sigma_{n-2})\ldots(\sigma_1\sigma_2)(\sigma_1) \in B_n$ be the Garside fundamental braid. Observe that $\Delta^2 = (\sigma_1\ldots\sigma_{n-1})^n$ is a full twist on n strands.

Using orderings, one defines the *Dehornoy's floor* $[\beta]_D = n$ of a braid $\beta \in B_m$ as an integer n such that $\Delta^{2n+2} \succ \beta \succ \Delta^{2n}$. The Dehornoy floor is related to the FDTC as follows [42]:

$$[\beta]_D + 1 \geq BT(\beta) \geq [\beta]_D \quad \text{and} \quad BT(\beta) = \lim_{n\to\infty} \frac{[\beta^n]_D}{n}. \tag{1}$$

In other words, the FDTC can be thought of as "homogenization" of the Dehornoy floor, where one averages over large iterates of the braid. It is important to note that while BT is an invariant of the conjugacy class of the braid, the Dehornoy floor is not.

The following theorem of Ito [32] serves as motivation for our work.

Theorem 8 ([32]) *If K is represented by an n-strand braid β, then*

$$|[\beta]_D| < \frac{4g_3(K) - 2}{n+2} + \frac{3}{2} \leq g_3(K) + 1.$$

Note that Ito uses a slightly different definition of the Dehornoy floor, so that his formulas in [32] do not have the absolute value. Ito's proof uses braid foliation techniques. It would be interesting to establish a similar bound via knot homologies.

3 Concordance Invariants and Genus Bounds

In the last two decades, a number of knot-homological invariants were introduced to study knot concordance and give bounds for the slice genus. It would be interesting to find relations between these invariants and the FDTC of fibered knots or braids. We briefly review the invariants that we need.

A number of invariants come from knot Floer homology, introduced independently by Ozsváth and Szabó in [47] and by Rasmussen in [53]. See also [43] for a survey. For the simplest version, they associate a $\mathbb{Z}$-filtered chain complex $\widehat{CFK}(K)$ to a knot K. This chain complex is a powerful knot invariant; in particular, it detects the Seifert genus $g_3(K)$ [48] and fiberedness [18, 46]. Total homology of this chain complex is of rank 1. The minimum filtration level in which the homology is supported yields an integer τ, which is a concordance invariant, [47]. A concordance invariant $s(K)$ with similar properties was found by Rasmussen [54] using Khovanov homology [33]. For a link $L \subset S^3$, the Khovanov homology is a link invariant $Kh(L)$, defined as the cohomology of a bigraded chain complex $(CKh(D_L), d)$ associated to a diagram D_L of the link. The invariant $s(K)$ comes from Lee's deformation [36] of the Khovanov differential and the resulting spectral sequence. The invariants $s(K)$ and $2\tau(K)$ share a number of properties, given in

the next proposition, and coincide for many small knots but are known to differ in general [25].

Proposition 6 ([47, 54]) *The maps $\tau : \mathcal{C} \to \mathbb{Z}$ and $s : \mathcal{C} \to \mathbb{Z}$ are surjective homomorphisms on the knot concordance group $\mathcal{C}$. They satisfy the following properties:*

1. *The absolute values of $\tau(K)$ and $\frac{s(K)}{2}$ give lower bounds on the slice genus $g_4(K)$,*

$$|\tau(K)| \leq g_4(K), \qquad |s(K)| \leq 2g_4(K).$$

2. *If β is a positive n-braid of length k whose closure $\hat{\beta}$ is a knot, then $s(\hat{\beta}) = 2\tau(\hat{\beta}) = k - n + 1$. In particular, for the (p, q)-torus knot $T_{p,q}$, $p, q \geq 1$, we have*

$$s(T_{p,q}) = 2\tau(T_{p,q}) = (p-1)(q-1) = 2g(T_{p,q}).$$

3. *$\tau(-K) = -\tau(K)$ and $s(-K) = -s(K)$ where $-K$ denotes the concordance inverse of K. (In fact this follows directly from the fact that τ and $\frac{s}{2}$ are homomorphisms.)*
4. *Let K_+ be a knot, and K_- the new knot obtained by changing one positive crossing in K_+ to a negative crossing. Then*

$$\tau(K_+) - 1 \leq \tau(K_-) \leq \tau(K_+)$$

and

$$s(K_+) - 2 \leq s(K_-) \leq s(K_+).$$

Remark 1 It is worth remarking that *any* concordance homomorphism satisfying (1)–(3) will also satisfy a crossing change formula analogous to the one given in (4). Specifically, the argument given in Corollary 3 of [38] uses only the formal properties (1)–(3).

The bound on the slice genus is shown to be sharp for the classes of positive, quasipositive, and strongly quasipositive knots, [38, 47, 50]. In general, we have the following estimate for τ and s:

Lemma 1 *Let β be an n-braid with k positive crossings and ℓ negative crossings. If the closure of β is a knot $\hat{\beta}$ then*

$$\frac{1}{2}(k - \ell - n + 1) \leq \tau(\hat{\beta}) \leq \frac{1}{2}(k - \ell + n - 1)$$

and

$$k - \ell - n + 1 \leq s(\hat{\beta}) \leq k - \ell + n - 1.$$

Proof Let β_+ denote the braid obtained from β by changing all of the negative crossings to positive and β_- the braid obtained by changing all of the positive crossings to negative. Then by part (3) of Proposition 6, we have $\tau(\hat{\beta}_+) - \ell \leq \tau(\hat{\beta})$ and $\tau(\hat{\beta}) \leq \tau(\hat{\beta}_-) + k$. Now parts (2) and (4) imply that

$$\tau(\hat{\beta}_+) = \frac{1}{2}(k + \ell - n + 1)$$

since $\hat{\beta}_+$ is a positive braid, and

$$\tau(\hat{\beta}_-) = -\frac{1}{2}(k + \ell - n + 1)$$

since $\hat{\beta}_- = -\hat{\beta}_+$.

Putting this information together, we see that

$$\frac{1}{2}(k - \ell - n + 1) = \tau(\hat{\beta}_+) - \ell \leq \tau(\hat{\beta}) \leq \tau(\hat{\beta}_-) + k = \frac{1}{2}(k - \ell + n - 1).$$

Similarly, for s we have

$$s(\hat{\beta}_+) = k + \ell - n + 1$$

and

$$s(\hat{\beta}_-) = -s(\hat{\beta}_+) = -k - \ell + n - 1.$$

The crossing change formula now gives the same inequality we had for τ, except is it multiplied by the necessary factor of two,

$$s(\hat{\beta}_+) - 2\ell \leq s(\hat{\beta}) \leq s(\hat{\beta}_-) + 2k$$

Thus,

$$k - \ell - n + 1 \leq s(\hat{\beta}) \leq k - \ell + n - 1.$$

Remark 2 The lower bound for $\tau(\hat{\beta})$ above appears in [38, Corollary 11]. The lower bound for $s(\widehat{\beta})$ appears in [51, Proposition 4] and [58, Lemma 1.C]. Each of these bounds immediately implies the slice-Bennequin inequality due to Rudolph. The upper bound for $s(\widehat{\beta})$ above is no stronger than the bounds from [40, Theorem 1.10] and additionally can be obtained from [44, Theorem 3.5].

The full knot Floer complex yields further concordance invariants, such as ν^+ [28]. We will discuss their connection to monodromies as well. Since these

invariants give lower bounds for the slice genus, we can ask about their relation to the FDTC.

Using properties of τ and s, we can easily prove that if one starts with a fixed knot and adds many full positive twists, then indeed the genus, slice genus, τ and s invariants all grow. If Δ is the Garside element, the braid $\beta\Delta^{2n}$ is the concatenation of β with n positive full twists. Note that $BT(\beta\Delta^{2n}) = BT(\beta)+n$ by Proposition 4. We prove Proposition 1 from the Introduction:

Proposition 1 *Let $\beta \in B_m$ be a braid, and let K_n be the closure of the braid $\beta\Delta^{2n}$. When n becomes large, $g_3(K_n) = g_4(K_n)$ and this value grows roughly as $\frac{1}{2}nm^2$.*

Proof Suppose that β has k positive crossings and l negative crossings. We first show that if $n \geq l$ then the braid $\beta\Delta^{2n}$ is a positive braid. Using the braid group relations it is not hard to show that $\Delta^2\sigma_i^{-1}$ is a positive braid for each $1 \leq i \leq m-1$. Since Δ^2 is central, each negative crossing can be removed by a single full-twist.

Thus, when $n \geq l$, since K_n is the closure of a positive braid, $g_3(K_n) = g_4(K_n)$ [34]. In fact, we have $g_3(K_n) = g_4(K_n) = \tau(K_n)$ since τ is equal to the slice-genus for positive braids [50]. By Proposition 6(2), $\tau(K_n)$ is a simple function of length of the positive braid $\beta\Delta^{2n}$ and the number of strands:

$$g_3(K_n) = g_4(K_n) = \tau(K_n) = \frac{1}{2}(k + (m-1)mn - l - m + 1).$$

Since m, l and k are constant, we achieve the desired asymptotics.

The above proposition tells us that the FDTC, the 3-genus, and the slice genus have similar asymptotics when we add more and more full twists to a given braid, but unlike Ito's bound, it gives no information about the relation between the FDTC and genus of the original braid, before the twists are added. Note that Ito's bound is weaker if we add a large number of positive twists to a fixed braid: it only says that the genus will grow as $\frac{1}{4}nm$.

Adding a single full twist to a braid increases the FDTC by one. A natural question to ask is whether this move also necessarily changes the slice genus, or whether it changes the concordance class of the knot. Certainly the above proposition shows that this is true asymptotically, but we show in the following proposition that Theorem 1.6 of Sato's work in [56] implies that this is also true for any fixed number of full twists for braids with three or more strands.

Proposition 2 *Let $\beta \in B_m$ be a braid, $m \geq 3$, and $\beta_k = \beta\Delta^{2k}$, $k \in \mathbb{Z}$, $k > 0$. Then all braid closures $\widehat{\beta}$, $\widehat{\beta_k}$, $k > 0$ lie in pairwise distinct concordance classes.*

Proof Hom and Wu in [28] define a refinement of the τ-invariant called ν^+ arising from the knot Floer complex. This invariant is a knot concordance invariant. In Proposition 1.5 of [56], Sato defines a partial order $<_{\nu^+}$ on the concordance group mod ν^+-equivalence. Let β_1 be the braid obtained by adding a single positive full-twist to β, that is, $\beta_1 = \beta\Delta^2$. Theorem 1.6 (2) of [56] implies that $[\widehat{\beta}] <_{\nu^+} [\widehat{\beta_1}]$. In particular, $\widehat{\beta}$ and $\widehat{\beta_1}$ are not concordant. As we add more full twists, we see that

$[\widehat{\beta}] <_{\nu^+} [\widehat{\beta_1}] <_{\nu^+} [\widehat{\beta_2}] <_{\nu^+} \cdots <_{\nu^+} [\widehat{\beta_k}]$. So all of these knots live in different ν^+-equivalence classes and none of them can be concordant to each other.

We note that for braids with two strands, Proposition 2 is false, as the braids σ_1^{-1} and σ_1 differ from each other by a single full twist in B_2 but both close to the unknot.

We prove Proposition 3 from the Introduction. The argument is based on the formal properties shared by τ and s.

Proposition 3 *Let $\beta = (\Delta^2)^k \sigma_1^{-1} \sigma_2^{-(6k-1)}$ in B_3. Then $\hat{\beta}$ is a knot, $BT(\beta) \geq k-1$, $|\tau(\hat{\beta})| \leq 1$ and $|s(\hat{\beta})| \leq 2$.*

Proof Observe that if $\Delta^2\beta'$ is any braid with $\hat{\beta}'$ a knot then $\widehat{\Delta^2\beta'}$ is also a knot. Now one can check that $\beta' = \sigma_1^{-1}\sigma_2^{-(6k-1)}$ is a knot.

Lemma 1 implies

$$-1 = \frac{1}{2}(6k - (6k-1) - 1 - 2) \leq \tau(\hat{\beta}) \leq \frac{1}{2}(6k - (6k-1) - 1 + 2) = 1.$$

Thus $|\tau(\hat{\beta})| \leq 1$. Similarly Lemma 1 implies that $|s(\hat{\beta})| \leq 2$. To show that $FDTC(\beta) \geq k-1$, first note that using braid relations we can rewrite

$$\beta = (\Delta^2)^{k-1}\sigma_1\sigma_2\sigma_1^2\sigma_2^{-6k+2}$$

For brevity let us write $\sigma_1\sigma_2\sigma_1^2\sigma_2^{-6k+2} = \alpha$. Note also that by property (c) of Proposition 4, $FDTC(\beta) = k - 1 + FDTC(\alpha)$. Finally, by Proposition 5, $BT(\alpha) \geq 0$ since it contains copies of σ_1 but no copies of σ_1^{-1}.

Thus any bound relating the fractional Dehn twist coefficient $BT(\beta)$ and the slice genus will have to resort to different proof methods than directly using the τ or the s-invariant.

4 Quasipositive Braids and the FDTC Bounds

A naive question would be to ask whether Theorem 8 holds as stated if we replace the three-genus by the slice genus. For instance, slice genus and three-genus are equal for closures of positive braids [34], and so the slice genus version of Theorem 8 immediately holds for positive braid closures. The next natural class of braids to consider are quasipositive braids: braids that can be written as a product of conjugates of the positive Artin generators. For quasipositive braids, we will prove the Ito-like bound stated in Theorem 2. In this section we assume the braid index $n \geq 3$.

Definition 1 We say that an n-braid $\beta \in B_n$ is *quasipositive* if it is represented by a braid word of the form

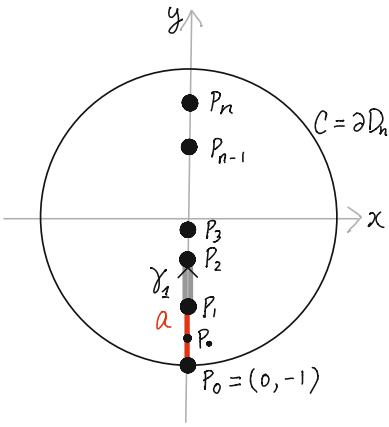


Fig. 1 Punctured disk D_n. Arcs a, point $p_\bullet$ and $\gamma_1 := \rho_1$ are used in the proof of Proposition 8

$$(w_1\sigma_{i_1}w_1^{-1})\,(w_2\sigma_{i_2}w_2^{-1})\cdots(w_m\sigma_{i_m}w_m^{-1})$$

where $\sigma_{i_1},\ldots,\sigma_{i_m} \in \{\sigma_1,\ldots,\sigma_{n-1}\}$ and w_i are some braid words in $\{\sigma_1^{\pm},\ldots,\sigma_{n-1}^{\pm}\}$. We say that the braid word has *quasi-positive-length* (qp-length) m.

Let us identify D_n with the unit disk in $\mathbb{R}^2$ equipped with (x, y) coordinates. Place n punctures on the y-axis and call them $p_1,\ldots,p_n$ so that the y-coordinate of p_i is less than that of p_{i+1} (see Fig. 1). Let ρ_i be the subarc of the y-axis joining p_i and p_{i+1}.

For a properly embedded arc γ in D_n joining two puncture points, let $H_\gamma \in \mathcal{MCG}(D_n)$ denote the positive half twist along the arc γ. The braid group B_n is isomorphic to the mapping class group $\mathcal{MCG}(D_n)$. With this isomorphism $\phi : B_n \to \mathcal{MCG}(D_n)$, a braid word $w \in B_n$ is identified with the mapping class $\phi(w)$ denoted by $\phi_w := \phi(w) \in \mathcal{MCG}(D_n)$. Let

$$\gamma_{w\sigma_i w^{-1}} := \phi_w(\rho_i)$$

be a properly embedded arc in D_n that joins two distinct points of $p_1,\ldots,p_n$. Then the braid word $w\sigma_i w^{-1} \in B_n$ is identified with the positive half twist along the arc $\gamma_{w\sigma_i w^{-1}}$, thus,

$$H_{\gamma_{w\sigma_i w^{-1}}} = \phi_{w\sigma_i w^{-1}}.$$

For more on this construction, see for instance [13] Example 3.5 or [20] Lemma 2.4.

Proposition 7 *Let β be a quasipositive n-braid. Let $m \geq 1$. The braid β has qp-length m if and only if there exist properly embedded arcs $\gamma_1,\ldots,\gamma_m$ joining the punctures in D_n as above, such that $\phi_\beta = H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}$.*

With Proposition 7 we may identify the braid $\beta \in B_n$ and the mapping class $\phi_\beta = H_{\gamma_m} \circ \cdots \circ H_{\gamma_1} \in \mathcal{MCG}(D_n)$ and we have

$$BT(\beta) = FT(H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}).$$

Here, we recall a useful lemma. The lemma applies for the FDTC with respect to any fixed boundary component C of a surface S (S may have one or several boundary components).

Lemma 2 ([30]) *Let $\phi \in \mathcal{MCG}(S)$. Fix a boundary component C of the surface S and let $FT(\phi)$ stand for the FDTC with respect to C. Let T_C denote the positive Dehn twist about C.*

If there exists an essential arc $\gamma \subset S$ that starts on C and satisfies $T_C^m(\gamma) \preceq \phi(\gamma) \preceq T_C^M(\gamma)$ for some $m, M \in \mathbb{Z}$ then

$$m \leq FT(\phi) \leq M.$$

In the above lemma, the symbol $\prec$ represents an ordering on the set of properly embedded arcs, see [31, Definition 3.2]. We write $\alpha \prec \beta$ if arcs α and β start at the same boundary point, realize the geometric intersection number, and β lies on the right side of α in a small neighborhood of the starting point. (The notion "β to the right of α" was already discussed in Sect. 2.2. Here, we write $\preceq$ to allow for the case $\alpha = \beta$.)

Lemma 3 *Let the braid index $n \geq 3$. Let $m \geq 1$. Let $\gamma_1, \ldots, \gamma_m$ be properly embedded arcs in D_n connecting two distinct punctures. Then we have $FT(H_{\gamma_1}) = 0$ and*

$$0 \leq FT(H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}) \leq m - 1.$$

Proof Since $n \geq 3$ there exists some essential arc that is fixed by the half twist H_{γ_1}. By Lemma 2 this means that $FT(H_{\gamma_1}) = 0$. (Note that if $n = 2$ then $FT(H_{\gamma_1}) = \frac{1}{2}$.) By the quasimorphism property of the FDTC in Proposition 4 and induction on m, we obtain $0 \leq FT(H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}) \leq m - 1$.

Let K be a link in $S^3 = \partial B^4$. Let $\chi_4(K)$ denote the maximal Euler characteristic of an oriented surface that is smoothly embedded in the 4-ball B^4 and bounded by K.

Lemma 4 ([55]) *Assume that $\beta \in B_n$ is a quasipositive braid of qp-length m. Then $\chi_4(\hat{\beta}) = n - m$.*

With the above two lemmas, we prove the following theorem.

Theorem 9 (Cf. Theorem 2) *Let $\beta \in B_n$ be a quasipositive n-braid of qp-length m. Then*

$$BT(\beta) \leq m - 1 = n - \chi_4(\hat{\beta}) - 1.$$

In particular, when the braid closure $\hat{\beta}$ is a knot then

$$BT(\beta) \leq 2g_4(\hat{\beta}) + (n-2).$$

Proof By Lemma 3, we have $BT(\beta) \leq m - 1$. By Lemma 4, we have $m - 1 = n - \chi_4(\hat{\beta}) - 1$.

The next proposition shows that the upper bound of $m - 1$ for $BT(\beta)$ is the best possible, as it guarantees the existence of a quasipositive braid β of quasi-positive length m with $BT = m - 1$.

Proposition 8 *For every $m = 2, 3, 4, \cdots$ there exist properly embedded arcs $\gamma_1, \cdots, \gamma_m$ in D_n joining distinct puncture points such that*

$$FT(H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}) = m - 1.$$

In the proof of Proposition 8 we use oriented train tracks. A train track is a graph. The edges are oriented and weighted. Each vertex has valence 3 or 4 where incoming edges and outgoing edges tangentially meet. As shown in the left picture in Fig. 2, at the vertex two outgoing edges (with weights b and c) are tangent to the incoming edge (with weight a) and the weights satisfy the equation $a = b + c$. (The same rule applies for a valence 4 vertex.) Replacing each edge of weight a with a parallel edges we obtain arcs. The right picture of Fig. 2 shows arcs carried by the train track with weights $a = 6, b = 4$ and $c = 2$. Orientations of the edges in the train track induce orientations of the carried arcs.

Proof of Proposition 8 Let $p_1, \cdots, p_n$ be the punctures of the disk D_n. Let $C = \partial D_n$ denote the boundary of the disk. Let a be an arc connecting p_1 to C as in Fig. 1. Let $\gamma_1 := \rho_1$, the line segment of the y-axis joining p_1 and p_2. We orient γ_1 upward (from p_1 to p_2).

To define a sequence of arcs $\{\gamma_m | m = 2, 3, \cdots\}$ we use two oriented train track templates $T_{\mathtt{1st}}(m)$ and $T_{\mathtt{2nd}}(m)$ depicted in Fig. 3.

In the train track $T_{\mathtt{1st}}(m)$, the orientation of edges near punctures are all outward. The bottom edge has weight 1 and ends at a point $p_\bullet$ on the y-axis between p_0 and p_1. Labels $j_m \in \mathbb{N}$ and $\alpha_{m,0}, \alpha_{m,1}, \ldots, \alpha_{m,2n-5} \in \{0, 1\}$ denote the weights of the edges. The edges $\alpha_{m,i}$ are labeled clockwise and $\alpha_{m,0}$ starts from the puncture

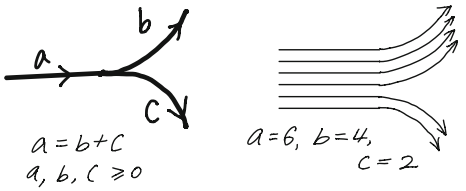


Fig. 2 An example of an oriented train track and its corresponding arcs

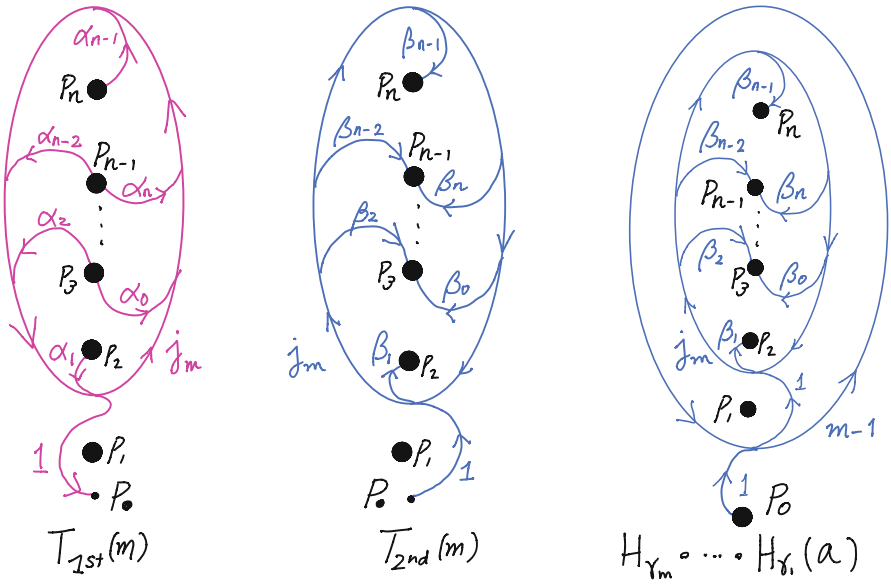


Fig. 3 Train tracks $T_{1\mathtt{st}}(m)$ and $T_{2\mathtt{nd}}(m)$ for γ_m, and a train track carrying the arc $H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}(a)$

p_3. (The same rule applies for the edges $\beta_{m,i}$ mentioned below.) For simplicity, the subscript m will be omitted in the following and $\alpha_i = \alpha_{m,i}$. We require only one of $\alpha_0, \ldots, \alpha_{2n-5}$ is 1 and the rest of them are all 0. The weights of the unlabeled edges can be easily computed.

The edges of $T_{2\mathtt{nd}}(m)$ are oriented and at each puncture the orientation points into the puncture. The weights $\beta_0, \cdots, \beta_{2n-5}$ are either 0 or 1 and only one of them is 1. The starting point of the bottom edge is $p_\bullet$.

Let $m = 2$. The arc γ_2 is defined to be an arc carried by the train track $T_{1\mathtt{st}}(2)$ first, and then $T_{2\mathtt{nd}}(2)$, where $j_2 = 1$ and $\alpha_1 = \beta_2 = 1$. We see that γ_2 starts at p_2 and ends at p_3 as in Fig. 4.

The image of the arc a under the diffeomorphism $H_{\gamma_2} \circ H_{\gamma_1}(a)$ is depicted in the bottom right picture in Fig. 4, which shows that

$$T_C(a) \prec H_{\gamma_2} \circ H_{\gamma_1}(a),$$

thus

$$FT(H_{\gamma_2} \circ H_{\gamma_1}) \geq 1.$$

For a general m, suppose that $m - 1 = k(2n-4) + r$ where $k \in \mathbb{Z}_{\geq 0}$ and $r \in \{0, 1, \ldots, 2n-5\}$ the remainder. The arc γ_m is defined to be an arc that is carried by $T_{1\mathtt{st}}(m)$ first and then $T_{2\mathtt{nd}}(m)$ with $\alpha_r = \beta_{r+1} = 1$ (here, the subscript

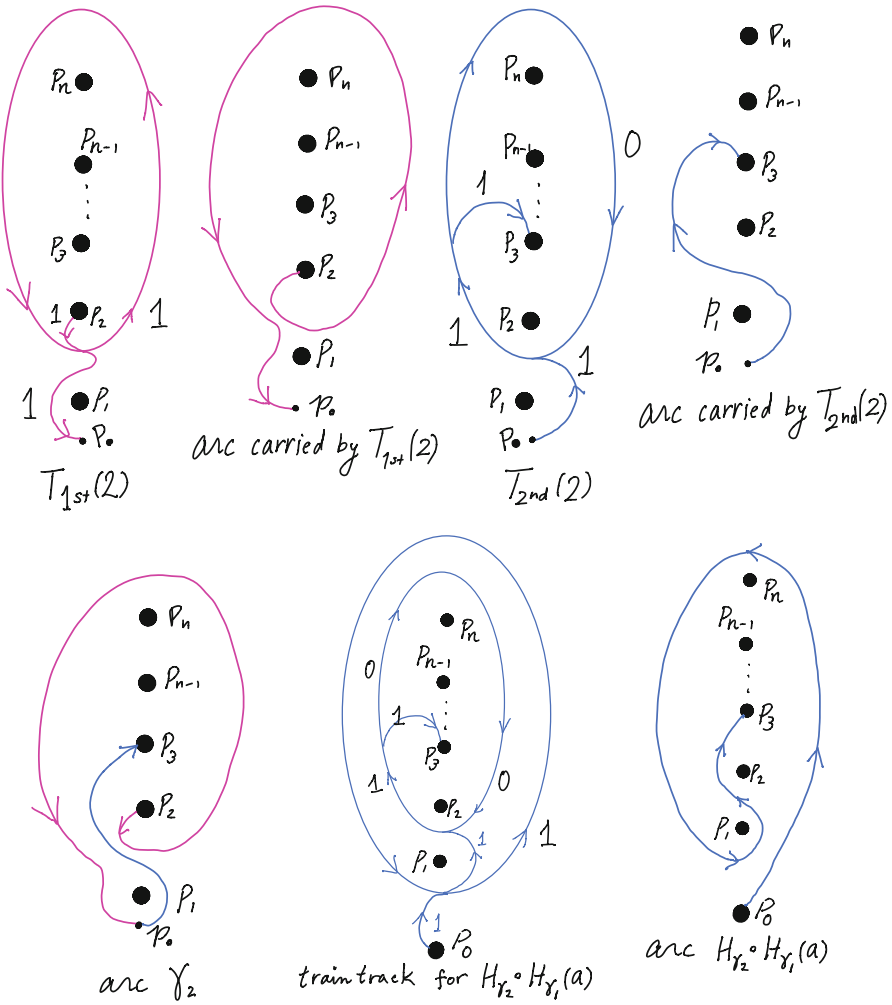


Fig. 4 ($m = 2$). The arcs γ_2 and $H_{\gamma_2} \circ H_{\gamma_1}(a)$ and their train tracks

$r + 1$ is considered modulo $2n - 4$) and

$$j_m = \left\lceil \frac{m-1}{2n-4} \right\rceil (= k \text{ or } k+1).$$

For example, when $m = 3$ the arcs γ_3 and $H_{\gamma_3} \circ H_{\gamma_2} \circ H_{\gamma_1}(a)$ are illustrated in Fig. 5.

See Fig. 3 again. We observe two things: (1) Since $\alpha_r = \beta_{r+1} = 1$ the templates $T_{\texttt{1st}}(m+1)$ and $T_{\texttt{2nd}}(m)$ are exactly the same except that they are oppositely

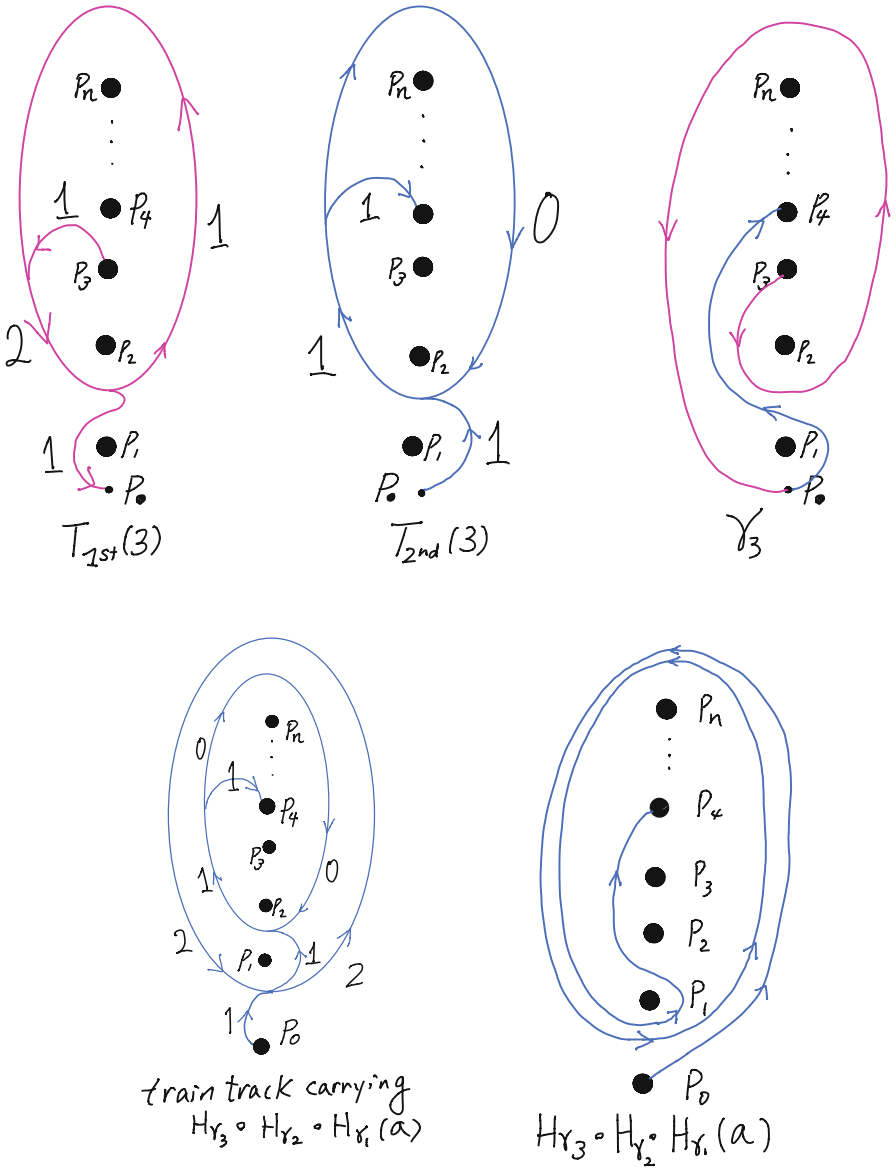


Fig. 5 ($m = 3$). Arcs γ_3 and $H_{\gamma_3} \circ H_{\gamma_2} \circ H_{\gamma_1}(a)$ and their train tracks

oriented and $T_{\texttt{1st}}(m+1)$ goes to the left side of p_1 whereas $T_{\texttt{2nd}}(m)$ goes to the right side of p_1. (2) Observe that $T_{\texttt{2nd}}(m)$ is embedded in $H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}(a)$.

With these observations in mind, applying the half-twist $H_{\gamma_{m+1}}$ to the arc $H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}(a)$, the embedded part $T_{\texttt{2nd}}(m)$ is erased by $T_{\texttt{1st}}(m+1)$ (which is the 1st part of γ_{m+1}) and then $T_{\texttt{2nd}}(m+1)$ (which is the 2nd part of γ_{m+1}) overwrites the erased part of $H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}(a)$. The resulting arc is exactly $H_{\gamma_{m+1}} \circ \cdots \circ H_{\gamma_1}(a)$.

Therefore, by induction on m we see that the image $H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}(a)$ is carried by the train track in the right sketch in Fig. 3.

The outer circle in the right picture in Fig. 3 is weighted $m-1$, which means that

$$T_C^{m-1}(a) \prec H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}(a).$$

By Theorem 9 and Lemma 2 we obtain $FT(H_{\gamma_m} \circ \cdots \circ H_{\gamma_1}) = m-1$.

Corollary 2 *There exist slice knots represented by quasipositive n-braids with* $BT = n-2$.

Proof For $m = n-1$, the construction of Proposition 8 produces a quasipositive braid β with $BT = n-2$. If we show that β is connected, then we get a slice knot, since by Lemma 4 we have $\chi_4(\beta) = 1$. To see connectedness, observe that the monodromy of β is the product of positive half-twists about the arcs $\gamma_1, \gamma_2, \ldots \gamma_{n-1}$ constructed in Proposition 8. The arc γ_k connects the k-th puncture to the $(k+1)$-st puncture, so the corresponding half-twist acts on the set of punctures by a transposition interchanging the k-th and $(k+1)$-st punctures. It follows that the composition of the half-twists about $\gamma_1, \gamma_2, \ldots \gamma_{n-1}$ gives a cyclic permutation of all the punctures: $n \mapsto n-1, \ldots, 3 \mapsto 2, 2 \mapsto 1, 1 \mapsto n$. So the resulting braid is connected.

In general, knots and links can be represented by a braid word of the form

$$(w_1 \sigma_{i_1} w_1^{-1})^{\epsilon_1} \, (w_2 \sigma_{i_2} w_2^{-1})^{\epsilon_2} \cdots (w_m \sigma_{i_m} w_m^{-1})^{\epsilon_m} \tag{2}$$

where the exponents $\epsilon_1, \ldots, \epsilon_m \in \{1, -1\}$. When $\epsilon_j = 1$ the word $w_j \sigma_{i_j} w_j^{-1}$ is called a positive syllable and when $\epsilon_j = -1$ the word $(w_j \sigma_{i_j} w_j^{-1})^{-1}$ is called a negative syllable. Let $p_+ \geq 0$ be the number of positive syllables and $p_- \geq 0$ be the number of negative syllables in the above braid representative (2). Thus $p_+ + p_- = m$. Note that when $p_+ = m$ and $p_- = 0$ the braid is quasipositive. Theorem 9 has another corollary.

Corollary 3 *Let* $\beta \in B_n$ *be the closure of an n-braid*

$$(w_1 \sigma_{i_1} w_1^{-1})^{\epsilon_1} \, (w_2 \sigma_{i_2} w_2^{-1})^{\epsilon_2} \cdots (w_m \sigma_{i_m} w_m^{-1})^{\epsilon_m}$$

that does not admit destabilizations and satisfies $\chi_4(\beta) = n - m$. *Let* p_+ *(resp.* p_-*) be the number of positive (resp. negative) syllables. Then the FDTC of* β *has the following upper and lower bounds:*

$$\min\{-p_- + 1, 0\} \leq BT(\beta) \leq \max\{p_+ - 1, 0\}$$

Proof Given a braid word $(w_1\sigma_{i_1}w_1^{-1})^{\epsilon_1}\ (w_2\sigma_{i_2}w_2^{-1})^{\epsilon_2}\cdots(w_m\sigma_{i_m}w_m^{-1})^{\epsilon_m}$ we define a quasipositive braid word as follows. For $j = 1, \ldots, m$ let

$$\delta_j = \begin{cases} 1 \text{ if } \epsilon_j = 1, \\ 0 \text{ if } \epsilon_j = -1. \end{cases}$$

Let

$$\beta_P = (w_1\sigma_{i_1}w_1^{-1})^{\delta_1}\ (w_2\sigma_{i_2}w_2^{-1})^{\delta_2}\cdots(w_m\sigma_{i_m}w_m^{-1})^{\delta_m}.$$

Note that β_P is a quasipositive (possibly trivial) braid of qp-length p_+. For every properly embedded arc γ in the n-punctured disk D_n the images under β and β_P satisfy

$$\beta(\gamma) \preceq \beta_P(\gamma)$$

thus, by Theorem 9 $BT(\beta) \leq BT(\beta_P) \leq \max\{p_+ - 1, 0\}$.

The lower bound can be obtained similarly.

Even though Proposition 8 claims sharpness, the estimates of the FDTC in both Theorem 9 and Corollary 3 are in general very rough for most braids as the next Proposition 9 implies.

Proposition 9 *Let $\beta \in B_n$. If $\hat{\beta}$ admits a positive (resp. negative) braid destabilization then $0 \leq BT(\beta) \leq 1$ (resp. $-1 \leq BT(\beta) \leq 0$).*

Proof Let $\hat{\beta}$ be a closed n-braid that admits a positive destabilization. Since the FDTC is invariant under conjugation, we may assume that β is represented by a braid word $w\sigma_{n-1}$ where $w \in B_{n-1}$. Define arcs δ and γ_{n-1} as in Fig. 6. We may identify the braids σ_{n-1} and w with the mapping classes $H_{\gamma_{n-1}}$ and $\phi_w \in \mathcal{MCG}(D_n)$. Since $w \in B_{n-1}$ we get $\phi_w(\delta) = \delta$. By Lemma 2 we get $BT(\phi_w) = 0$. Likewise, we get $BT(H_{\gamma_{n-1}}) = 0$ because we can find some properly embedded arc that is fixed by the half twist $H_{\gamma_{n-1}}$. By Proposition 4 we get $|BT(\beta)| \leq 1$.

Since $T_C^0(\delta) = \delta \prec \phi_\beta(\delta) = H_{\gamma_{n-1}}(\delta)$, Lemma 2 implies that $0 \leq BT(\beta)$. Thus we get $0 \leq BT(\beta) \leq 1$.

5 An Interesting Example

The examples in the last section show that the statement of Ito's Theorem 8 does not hold when we replace the genus with the slice genus, since Corollary 2 produces slice knots that are the closures of n-braids with BT equal to $n - 2$.

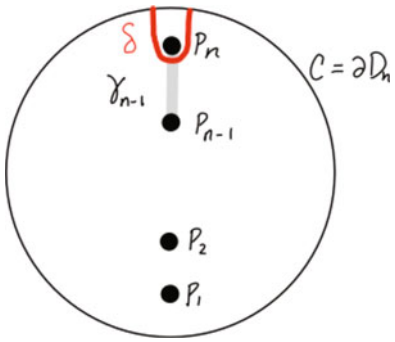


Fig. 6 Definition of the arcs δ and γ_{n-1}

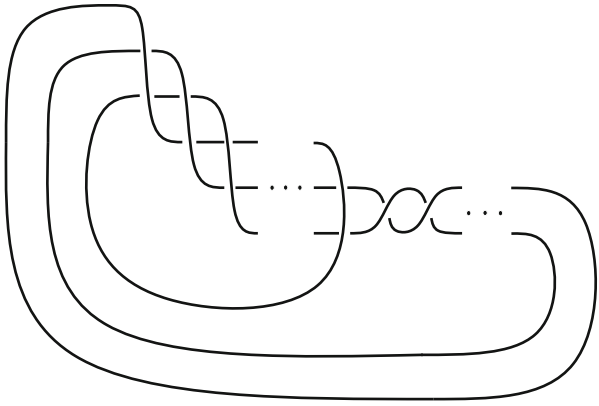


Fig. 7 $K_{m,k}$ is the closure of the braid $\beta = (\sigma_2\sigma_1)^{3m+1}\sigma_2^{-2k}$

In this section we give additional examples, fundamentally different from those in Corollary 2, that also show that Theorem 8 does not hold for the slice genus. Notice that for quasipositive 3-braids β whose closures are slice, Theorem 9 implies that $|BT(\beta)| \leq 1$. In Proposition 10, we produce examples of 3-braids that have arbitrarily large $|BT|$:

Proposition 10 *For each even $m > 0$, there exists a knot K_m and its 3-braid representative β_m such that $g_4(K_m) \leq m/2 + 1$ while $[\beta_m]_D = m$.*

We now describe the examples necessary to prove Proposition 10. Consider the braid word $\beta = (\sigma_2\sigma_1)^{3m+1}\sigma_2^{-2k}$ and let $K_{m,k} = \hat{\beta}$, see Fig. 7. Note that using the braid relations we can write $\beta = \Delta^{2m}\sigma_2\sigma_1\sigma_2^{-2k}$ where $\Delta = \sigma_1\sigma_2\sigma_1$ is the Garside element for braids of 3 strands.

Proposition 11 *Let $\beta = (\sigma_2\sigma_1)^{3m+1}\sigma_2^{-2k}$. Then we have $[\beta]_D = m$.*

Proof To show this, it is enough to show that $\Delta^{2m} \leq_D \beta <_D \Delta^{2m+2}$. To check that $\beta <_D \Delta^{2m+2}$ we must show that $\beta^{-1}\Delta^{2m+2}$ is σ_i-positive for $i = 1$ or 2. Simplifying, we have

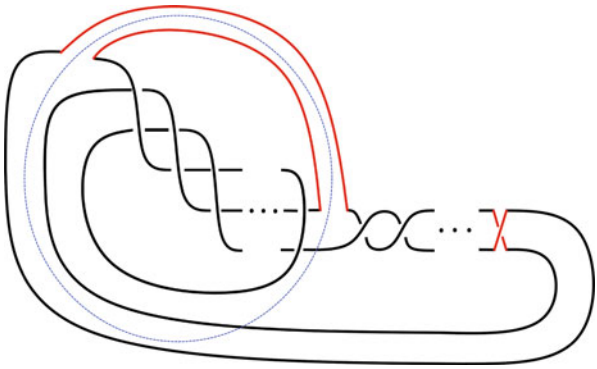


Fig. 8 $K_{m,k}$ after two band moves have been performed, with a decomposing sphere

$$\begin{aligned}\beta^{-1}\Delta^{2m+2} &= \sigma_2^{2k}\sigma_1^{-1}\sigma_2^{-1}\Delta^{-2m}\Delta^{2m+2}\\ &= \sigma_2^{2k}\sigma_1^{-1}\sigma_2^{-1}\Delta^2\\ &= \sigma_2^{2k}\sigma_1^{-1}\sigma_2^{-1}(\sigma_2\sigma_1)^3\\ &= \sigma_2^{2k}(\sigma_2\sigma_1)^2,\end{aligned}$$

which is in fact a σ_1-positive word. On the other hand

$$\begin{aligned}(\Delta^{2m})^{-1}\beta &= \Delta^{-2m}\Delta^{2m}\sigma_2\sigma_1\sigma_2^{-2k}\\ &= \sigma_2\sigma_1\sigma_2^{-2k},\end{aligned}$$

which is also a σ_1-positive word. The proposition follows.

We are now ready to prove Proposition 10.

Proof of Proposition 10 For any knot in the family $K_{m,k}$ there is a sequence of two band moves on K that yields the connected sum of torus knots $K' = T_{3,3m+1}\# -T_{2,2k+1}$, see Fig. 8.

By Feller [14, Corollary 3], we have

$$g_4(K') = \max\{|\tau(T_{3,3m+1}) - \tau(T_{2,2k+1})|, |\upsilon(T_{3,3m+1}) - \upsilon(T_{2,2k+1})|\}$$

where $\upsilon(K) := \Upsilon_K(1)$ is the Upsilon invariant evaluated at $t = 1$.

The values of τ and υ can be explicitly calculated for some torus knots. Let p and q be positive integers, then by Ozsváth and Szabó[47]

$$\tau(T_{p,q}) = \frac{(p-1)(q-1)}{2}.$$

For υ we have that

$$v(T_{2,2k+1}) = -k \text{ and } v(T_{3,3m+1}) = -2m$$

as computed by Feller [14].

Now for each even integer m, let K_m denote the closure of the braid $\beta_m = (\sigma_2\sigma_1)^{3m+1}\sigma_2^{-5m}$. Let $K'_m = T_{3,3m+1}\# - T_{2,5m+1}$ which is obtained from K_m from the two band moves illustrated in Fig. 8. Then

$$|\tau(T_{3,3m+1}) - \tau(T_{2,5m+1})| = |3m - 5m/2| = m/2$$

and

$$|v(T_{3,3m+1}) - v(T_{2,5m+1})| = | - 2m + 5m/2| = m/2.$$

Thus, $g_4(K'_m) = m/2$. Since K_m and K'_m are related by two band moves, $g_4(K_m)$ and $g_4(K'_m)$ differ by at most 1 and we obtain $g_4(K_m) \leq m/2 + 1$. Together with Proposition 11, this concludes the proof.

6 Potential Bounds on Slice Genus from the Braid Perspective

It is possible that the examples in Propositions 8 and 10 represent a worst-case scenario. Inspired by these examples and Theorem 2, in this section we further probe the following question.

Question 1 For any n-strand β representing a knot K, is it always true that

$$|BT(\beta)| \leq 2g_4(K) + n - 2?$$

If not, is there a bound of the order

$$|BT(\beta)| \leq C(n)g_4(K) + D(n),$$

where $C(n)$, $D(n)$ are constant for each fixed n?

Equivalently, similar questions can be asked about the Dehornoy's floor $[\beta]_D$, since $[\beta]_D$ and the fractional Dehn twist coefficient $BT(\beta)$ are related by (1).

One of the crucial observations necessary for the examples in Proposition 10 is that there exist cobordisms of controlled genus between stacked braids and connected sums of closed braids. We observe here that with a positive answer to Question 1 (or any similar bounds), this could produce an application: one could find bounds on the slice genus of connected sums of large torus knots. The slice genus of sums and differences of torus knots has been studied in several papers, see for instance [1, 2, 14, 37, 41].

In addition to the quasipositive braids studied in Sect. 4, we are able to answer Question 1 for certain slice knots. In fact, under some additional hypotheses, slice knots satisfy stronger bounds. First, note that if β closes to the unknot, then $|BT(\beta)| \leq 1$ by Theorem 8. (It is easy to find examples showing that even for the unknot, BT does not have to be zero.)

Further, recall that taking the connected sum of a knot and its mirror is a useful method for constructing slice knots: for any K, $K\# -K$ is a slice knot.

Proposition 12 (Follows Directly from [42]) *Choose an arbitrary oriented knot K in S^3. For any braid β that closes to $\hat{\beta} = K\# -K$, we have that $|BT(\beta)| \leq 1$.*

Proof Theorem 15.3 of [42] states that for any braid α such that $|BT(\alpha)| > 1$ and $\hat{\alpha}$ is a knot, $\hat{\alpha}$ is prime—that is, it cannot be expressed as a connected sum. The proposition directly follows by the contrapositive.

We provide evidence for a positive answer to Question 1 for special classes of three-braids in Theorem 3.

Theorem 3 *Let K be a knot that can be represented as the closure of a braid in B_3. Suppose further that K is slice, or more generally, that K has a finite order in the concordance group. Then any 3-braid representative β of K satisfies $|BT(\beta)| \leq 1$.*

A powerful tool at our disposal is a classification up to conjugation due to Murasugi. Before proving the theorem, we state the classification of 3-braids and note some properties of slice 3-braids.

Theorem 10 ([45]) *Let w be a braid word in B_3. Then w is conjugate to one of the following:*

1. $\Delta^{2d}\sigma_1\sigma_2^{-a_1}\cdots\sigma_1\sigma_2^{-a_n}$, *where the $a_i \geq 0$ with at least one $a_i \neq 0$*
2. $\Delta^{2d}\sigma_2^m$ *for $m \in \mathbb{Z}$*
3. $\Delta^{2d}\sigma_1^m\sigma_2^{-1}$ *where $m \in \{-1, -2, -3\}$.*

We will say a braid word in B_3 is in Murasugi normal form if it takes one of the above three forms. Baldwin classifies which closures of 3-braids are quasi-alternating in the following theorem.

Theorem 11 (Theorem 8.6 in [5]) *Suppose L is link with braid index at most 3 and is the closure of a braid β represented by $w \in B_3$ which is in Murasugi normal form. Then L is quasi-alternating if and only if one of the following holds:*

1. *w is in the first class and $d \in \{-1, 0, 1\}$;*
2. *w is in the second class and either $d = 1$ and $m \in \{-1, -2, -3\}$ or $d = -1$ and $m \in \{1, 2, 3\}$;*
3. *w is in the third class and $d \in \{0, 1\}$.*

Baldwin also obtains a result on 3-braids with finite concordance order.

Proposition 13 (Proposition 1.6 in [5]) *If K is a knot with braid index at most 3 and K has finite concordance order, then K can be represented as the closure of a*

braid of the form $\Delta^{2d}\sigma_1\sigma_2^{-a_1}\cdots\sigma_1\sigma_2^{-a_n}$ *where the* $a_i \geq 0$ *with at least one* $a_j \neq 0$, $d \in \{-1, 0, 1\}$, *and the* a_i *satisfy some further conditions.*

This implies the following immediate corollary.

Corollary 4 *A knot* K *with braid index at most* 3 *which has finite concordance order is quasi-alternating.*

Proposition 14 *Let* $w \in B_3$ *be a braid word in Murasugi normal form. Then*

1. $BT(w) = d$ *if* w *is in class* (1).
2. $BT(w) = d$ *if* w *is in class* (2).
3. $BT(w) = d - 1/3$, $BT(w) = d - 1/2$ *or* $BT(w) = d - 2/3$ *if* m *is* -1, -2 *or* -3 *respectively, and* w *is in class* (3).

Proof If w is in class (1) then $BT(w) = d + BT(\sigma_1\sigma_2^{-a_1}\cdots\sigma_1\sigma_2^{-a_n})$. If w is in class (2) then $BT(w) = d + BT(\sigma_2^m) = d$. It follows from Proposition 5 that the second summand in each case is zero, so we obtain that $BT(\beta) = BT(w) = d$ for w in class (1) or (2).

If w is in class (3), then $BT(w) = d + BT(\sigma_1^m\sigma_2^{-1})$. Now $BT(\sigma_1^m\sigma_2^{-1})$ is not zero. There are three cases to consider since $m \in \{-1, -2, -3\}$. We compute the following using the Artin relations, and properties of the FDTC enumerated in Proposition 4.

(a) $BT(\sigma_1^{-1}\sigma_2^{-1}) = (1/3)BT((\sigma_1^{-1}\sigma_2^{-1})^3) = (1/3)BT(\Delta^{-2}) = -1/3$
(b) $BT(\sigma_1^{-2}\sigma_2^{-1}) = (1/2)BT((\sigma_1^{-2}\sigma_2^{-1})^2) = (1/2)BT(\Delta^{-2}) = -1/2$
(c) $BT(\sigma_1^{-3}\sigma_2^{-1}) = (1/3)BT((\sigma_1^{-3}\sigma_2^{-1})^3) = (1/3)BT(\Delta^{-4}) = -2/3$

Thus for w in class (3) we have that $BT(w) = d - 1/3$, $BT(w) = d - 1/2$ or $BT(w) = d - 2/3$ if m is -1, -2 or -3 respectively.

Proof of Theorem 3 Corollary 4 shows that β closes to a quasi-alternating knot. We know that β is conjugate to a braid word w in one of the three classes of Theorem 10. To close to a quasi-alternating knot, w must satisfy the conditions on d in Theorem 11.

If w is in class (1) or (2) this means that $-1 \leq BT(w) \leq 1$ by Proposition 14. If w is in class (3) then $-2/3 \leq BT(w) \leq 1$. The theorem follows from the fact that BT is a conjugacy invariant.

We conclude the section with some computational evidence toward an affirmative answer to the first bound given in Question 1. We use the Hedden-Mark bound of Corollary 1. The quantity $\dim_{\mathbb{F}} \widehat{HF}(\Sigma(K))$ can be obtained from computations done in [60]. We set aside slice knots with 3-braids representatives, as for these Question 1 is already answered by Theorem 3. For all other knots with 12 crossings or fewer, we calculate that the inequality

$$\dim_{\mathbb{F}} \widehat{HF}(\Sigma(K)) - |H_1(\Sigma(K))| + 2 \leq 2g_4(K) + n - 2. \tag{3}$$

holds with 120 exceptions. As before, n stands for the braid index.

This means that the bound $|BT(\beta)| \leq 2g_4(K) + n - 2$ is valid for odd-strand braid representatives all but possibly 120 knots with 12 crossings or fewer, i.e. for at least 96% of such knots. (According to KnotInfo, there are 2977 knots with 12 crossings or fewer.)

Similarly, we can check that odd-strand braid representatives of all but 5 quasi-positive knots of 12 crossings or fewer satisfy inequality (3), and thus they satisfy the first inequality in Question 1. (For these knots, this experimental statement is stronger than Theorem 9, as it also applies to non-quasipositive braid representatives of quasipositive knots.)

7 Fibered Knots and Knot Floer Stable Equivalence

In Sects. 7 and 8, we collect some observations about the relationship between the fractional Dehn twist coefficient of fibered knots and their topological properties.

We say that two knots are *knot Floer stably equivalent* if their knot Floer complexes are isomorphic after possibly adding acyclic summands. Hom [26] shows that if two knots in S^3 are concordant, then they are knot Floer stably equivalent. Moreover, if two knots are knot Floer stably equivalent, then many invariants derived from knot Floer homology will coincide for the two knots. For example, the Ozsváth-Stipsicz-Szabó concordance invariant $\Upsilon_K(t)$ [49] will satisfy $\Upsilon_{K_1}(t) = \Upsilon_{K_2}(t)$ if K_1 and K_2 are stably equivalent.

He, Hubbard, and Truong showed in [21] that the $\Upsilon_K(t)$ invariant can detect right-veeringness. More precisely:

Theorem 12 ([21, Theorem 1.3]) *Suppose K is a fibered knot in S^3. Then associated to K is an open book decomposition (Σ, ϕ) of S^3. If $\Upsilon'_K(t) = -g$ for some $t \in [0, 1)$, where g is the genus of the fibered surface Σ, then $\phi : \Sigma \to \Sigma$ is right-veering.*

We note that Theorem 1.3 in [21] is slightly more general, as it is stated for null-homologous knots in rational homology three-spheres. Due to this theorem we are motivated to ask whether the knot Floer stable equivalence class of a fibered knot also can detect right-veeringness. However, it is easy to see that the answer is no.

Lemma 5 *The right-veering property and the FDTC are* not *invariants of knot Floer stable equivalence of a fibered knot.*

Proof Let K be the $(2, 1)$-cable of the figure-eight knot, which has positive FDTC by Kazez and Roberts [35] so it is right veering. Since $K\#P\#-P$ is concordant to K for any fibered knot P, the connected sum $K\#P\#-P$ is stably equivalent to K. The monodromy of $K\#P\#-P$ sends some arcs to the right and some to the left. Thus, $K\#P\#-P$ is not right-veering, and the fractional Dehn twist coefficient of $K\#P\#-P$ is zero. (See Proposition 5 and the discussion that follows it.)

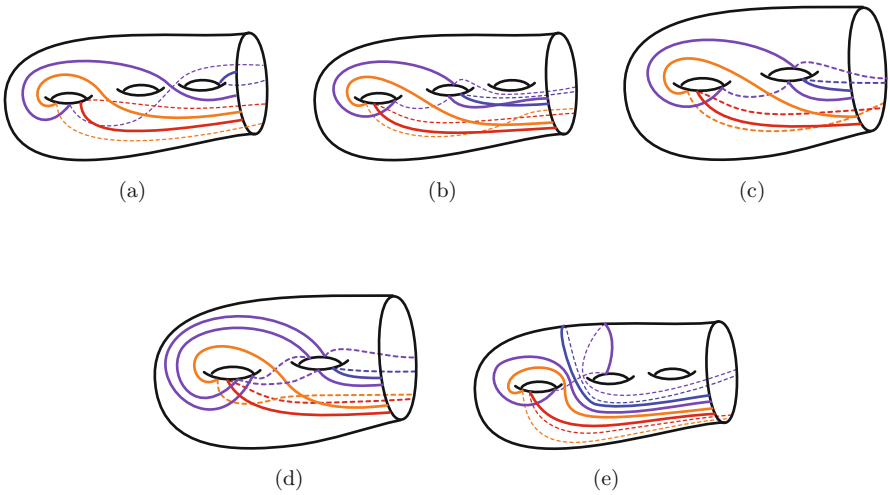


Fig. 9 Fibered slice knots with vanishing fractional Dehn twist coefficient. Each figure shows a pair of arcs x (in red) and y (in blue) such that $\phi(x)$ (in orange) is to the right of x and $\phi(y)$ (in purple) is to the left of y. The monodromies are presented using the conventions of Knot Info [39]; see Fig. 10. A lowercase letter means a right-handed Dehn twist and an uppercase letter means a left-handed Dehn twist. A word is read from right to left; thus, aB means first perform a left-handed twist around b then perform a right-handed twist around a. (**a**) The monodromy ϕ for 8_9 is $abcDEF$. (**b**) The monodromy ϕ for 9_{27} is $abCDEf$. (**c**) The monodromy ϕ for 10_{137} is $abCdF$. (**d**) The monodromy ϕ for 10_{140} is $abbCDf$. (**e**) The monodromy ϕ for $11a_{96}$ is $aBcDEf$

In the next section, we will study the fractional Dehn twist coefficient of a special class of fibered knots.

8 Fractional Dehn Twist Coefficient of Fibered Slice Knots

Many simple examples of fibered slice knots have monodromies with the fractional Dehn twist coefficient zero. For example, given any fibered knot K, the slice knot $K\# -m(K)$ has the FDTC zero. One way to see this is that we can assume there is some arc that the monodromy of K sends to the right, then the corresponding arc is sent to the left by the monodromy for $-m(K)$ and so $K\# -m(K)$ sends arcs to both the right and the left.

There are many more examples of fibered slice knots with fractional Dehn twist coefficient zero. For instance, the knot 8_{20} is slice and fibered, and has monodromy with fractional Dehn twist coefficient zero [35, Example 2.8]. Figure 9 gives several more examples of fibered slice knots with vanishing fractional Dehn twist coefficient $FT = 0$. For the monodromies the conventions of Knot Info are followed [39]; see Fig. 10.

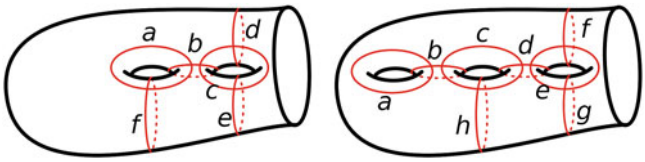


Fig. 10 This figure shows the curves whose Dehn twists are generators of the mapping class group of a fiber surface of genus two and three with one boundary component

Despite the numerous examples of fibered slice knots with vanishing fractional Dehn twist coefficient, the next proposition shows there exist many non-vanishing examples.

Proposition 15 *For every integer* $|p| \geq 2$*, there exist fibered, slice knots with fractional Dehn twist coefficient* $\frac{1}{p}$.

Proof Our examples are $(p, 1)$-cables of fibered, slice knots. Indeed, any cable of a fibered knot K is well-known to be fibered. This follows from [59] or from explicitly building a fibration from the fibration of the companion knot K and the fibration of the pattern torus knot. If K and K' are concordant, then their (p, q)-cables are also concordant (see, for example, [27]). Thus, if K is a slice knot, then the $(p, 1)$-cable of K is also slice. The fractional Dehn twist coefficient of the $(p, 1)$-cable of K is $1/p \neq 0$ by Kazez and Roberts [35, Proposition 4.2].

The examples in Proposition 15 are cable knots and thus they correspond to reducible monodromies. Restricting to fibered slice knots with pseudo-Anosov monodromies leads to the following question.

Question 2 If K is a fibered slice knot with pseudo-Anosov monodromy, then is the fractional Dehn twist coefficient zero?

Recall that a result of Gabai (see Theorem 4 stated in the Introduction) establishes a relation between the 3-genus of a fibered knot and the FDTC of the fibration. A positive answer to Question 2 would signify a connection between the fractional Dehn twist coefficient and the 4-ball genus.

Acknowledgments This article resulted from work completed at the Women in Symplectic and Contact Geometry and Topology Workshop (WiSCon) hosted by ICERM. J. Elisenda Grigsby, Diana Hubbard, Keiko Kawamuro, and Olga Plamenevskaya were mentors for Project 6, titled "Homological invariants, braids, transverse links, and surfaces". The participants were Elaina Aceves, Celeste Damiani, Feride Ceren Kose, Gage Martin, Juanita Pinzón Caicedo, Katherine Raoux, Linh Truong, and Hannah Turner. GM and LT thank Siddhi Krishna for helpful discussions. The results in this article reflect the observations of two sub-groups of Project 6. The mentors and participants of Project 6 would like to thank the organizers of WiSCon and are grateful to ICERM and to the AWM for supporting WiSCon via the AWM ADVANCE grant NSF-HRD 1500481. Finally, the authors thank the anonymous referee for their feedback; in particular they suggested an improvement to Proposition 1.

References

1. Paolo Aceto and Antonio Alfieri, *On sums of torus knots concordant to alternating knots*, Bull. Lond. Math. Soc. **2** (2019), no. 2, 327–343. MR 3937591
2. Samantha Allen, *Concordances from differences of torus knots to L-space knots*, Proc. Amer. Math. Soc. **148** (2020), no. 4, 1815–1827.
3. J. W. Alexander, *A Lemma on Systems of Knotted Curves*, Proc. of the Natl. Acad. of Sci. U.S.A **9** (1923), no. 3, 93–95.
4. Emil Artin, *Theorie der Zöpfe*, Abh. Math. Sem. Univ. Hamburg **4** (1925), no. 1, 47–72. MR 3069440
5. John A. Baldwin, *Heegaard Floer homology and genus one, one-boundary component open books*, J. Topol. **1** (2008), no. 4, 963–992. MR 2461862
6. Joan S. Birman and Tara E. Brendle, *Braids: a survey*, Handbook of knot theory, Elsevier, 2005, pp. 19–103.
7. John A. Baldwin and John B. Etnyre, *Admissible transverse surgery does not preserve tightness*, Math. Ann. **357** (2013), no. 2, 441–468. MR 3096514
8. John A. Baldwin and J. Elisenda Grigsby, *Categorified invariants and the braid group*, Proc. Amer. Math. Soc. **143** (2015), no. 7, 2801–2814.
9. John A. Baldwin and D. Shea Vela-Vick and Vera Vértesi, *On the equivalence of Legendrian and transverse invariants in knot Floer homology*, Geometry & Topology **17** (2013), no. 2, 925–974.
10. Joan S Birman, *Braids, links, and mapping class groups.(am-82)*, vol. 82, Princeton University Press, 2016.
11. Patrick Dehornoy, *Braid groups and left distributive operations*, Trans. Amer. Math. Soc. **345** (1994), no. 1, 115–150. MR 1214782
12. David Eisenbud and Walter Neumann, *Three-dimensional link theory and invariants of plane curve singularities*, Annals of Mathematics Studies, vol. 110, Princeton University Press, Princeton, NJ, 1985. MR 817982
13. John B. Etnyre and Jeremy Van Horn-Morris, *Monoids in the mapping class group*, Interactions between low-dimensional topology and mapping class groups, Geom. Topol. Monogr., vol. 19, Geom. Topol. Publ., Coventry, 2015, pp. 319–365. MR 3609913
14. Peter Feller, *Optimal cobordisms between torus knots*, Comm. Anal. Geom. **24** (2016), no. 5, 993–1025. MR 3622312
15. R. Fenn, M. T. Greene, D. Rolfsen, C. Rourke, and B. Wiest, *Ordering the braid groups*, Pacific J. Math. **191** (1999), no. 1, 49–74. MR 1725462
16. Peter Feller and Diana Hubbard, *Braids with as many full twists as strands realize the braid index*, J. Topol. **12** (2019), no. 4, 1069–1092. MR 3977871
17. David Gabai, *Problems in foliations and laminations*, Geometric topology (Athens, GA, 1993), AMS/IP Stud. Adv. Math., vol. 2, Amer. Math. Soc., Providence, RI, 1997, pp. 1–33. MR 1470750
18. Paolo Ghiggini, *Knot Floer homology detects genus-one fibred knots*, Amer. J. Math. **130** (2008), no. 5, 1151–1169. MR 2450204
19. David Gabai and Ulrich Oertel, *Essential laminations in* 3*-manifolds*, Ann. of Math. (2) **130** (1989), no. 1, 41–73. MR 1005607
20. Kyle Hayden, *Quasipositive links and Stein surfaces*, arXiv preprint arXiv:1703.10150 (2017).
21. Dongtai He, Diana Hubbard, and Linh Truong, *On the upsilon invariant of fibered knots and right-veering open books*, 2019.
22. Ko Honda, William H. Kazez, and Gordana Matić, *Right-veering diffeomorphisms of compact surfaces with boundary*, Invent. Math. **169** (2007), no. 2, 427–449. MR 2318562
23. _____, *Right-veering diffeomorphisms of compact surfaces with boundary. II*, Geom. Topol. **12** (2008), no. 4, 2057–2094. MR 2431016
24. Matthew Hedden and Thomas E. Mark, *Floer homology and fractional Dehn twists*, Adv. Math. **324** (2018), 1–39. MR 3733880

25. Matthew Hedden and Philip Ording, *The Ozsváth-Szabó and Rasmussen concordance invariants are not equal*, Amer. J. Math. **130** (2008), no. 2, 441–453. MR 2405163
26. Jennifer Hom, *A survey on Heegaard Floer homology and concordance*, J. Knot Theory Ramifications **26** (2017), no. 2, 1740015, 24. MR 3604497
27. Matthew Hedden and Juanita Pinzon-Caicedo, *Satellites of infinite rank in the smooth concordancegroup*, 2018.
28. Jennifer Hom and Zhongtao Wu, *Four-ball genus bounds and a refinement of the Ozváth-Szabó tau invariant*, J. Symplectic Geom. **14** (2016), no. 1, 305–323. MR 3523259
29. Tetsuya Ito and Keiko Kawamuro, *Overtwisted discs in planar open books*, Internat. J. Math. **26** (2015), no. 3, 1550027, 29.
30. ———, *Essential open book foliations and fractional Dehn twist coefficient*, Geom. Dedicata **187** (2017), 17–67. MR 3622682
31. ———, *Quasi-right-veering braids and nonloose links*, Algebr. Geom. Topol. **19** (2019), no. 6, 2989–3032. MR 4023334
32. Tetsuya Ito, *Braid ordering and knot genus*, J. Knot Theory Ramifications **20** (2011), no. 9, 1311–1323. MR 2844810
33. Mikhail Khovanov, *A categorification of the Jones polynomial*, Duke Math. J. **101** (2000), no. 3, 359–426. MR 1740682
34. P. B. Kronheimer and T. S. Mrowka, *The genus of embedded surfaces in the projective plane*, Math. Res. Lett. **1** (1994), no. 6, 797–808. MR 1306022
35. William H. Kazez and Rachel Roberts, *Fractional Dehn twists in knot theory and contact topology*, Algebr. Geom. Topol. **13** (2013), no. 6, 3603–3637. MR 3248742
36. Eun Soo Lee, *An endomorphism of the Khovanov invariant*, Adv. Math. **197** (2005), no. 2, 554–586. MR 2173845
37. Charles Livingston, *Concordances from connected sums of torus knots to L-space knots*, New York J. Math. **24** (2018), 233–239. MR 3778502
38. Charles Livingston, *Computations of the Ozsváth-Szabó knot concordance invariant*, Geom. Topol. **8** (2004), 735–742. MR 2057779
39. Charles Livingston and Allison H. Moore, *Knotinfo: Table of knots invariants*, Feb. 11 2020, http://www.indiana.edu/~knotinfo.
40. Andrew Lobb, *Computable bounds for Rasmussen's concordance invariant*, Compositio Mathematica **147** (2011), no. 2, 661–668. MR 2776617
41. Charles Livingston and Cornelia A. Van Cott, *The four-genus of connected sums of torus knots*, Math. Proc. Cambridge Philos. Soc. **164** (2018), no. 3, 531–550. MR 3784267
42. A. V. Malyutin, *Writhe of (closed) braids*, Algebra i Analiz **16** (2004), no. 5, 59–91. MR 2106667
43. Ciprian Manolescu, *An introduction to knot Floer homology*, Physics and mathematics of link homology, Contemp. Math., vol. 680, Amer. Math. Soc., Providence, RI, 2016, pp. 99–135. MR 3591644
44. Gage Martin, *Annular Rasmussen invariants: Properties and 3-braid classification*, arXiv:1909.09245 [math] (2019), arXiv: 1909.09245.
45. Kunio Murasugi, *On closed* 3*-braids*, American Mathematical Society, Providence, R.I., 1974, Memoirs of the American Mathmatical Society, No. 151. MR 0356023
46. Yi Ni, *Knot Floer homology detects fibred knots*, Invent. Math. **170** (2007), no. 3, 577–608. MR 2357503
47. Peter Ozsváth and Zoltán Szabó, *Knot Floer homology and the four-ball genus*, Geom. Topol. **7** (2003), 615–639. MR 2026543
48. ———, *Holomorphic disks and genus bounds*, Geom. Topol. **8** (2004), 311–334. MR 2023281
49. Peter S. Ozsváth, András I. Stipsicz, and Zoltán Szabó, *Concordance homomorphisms from knot Floer homology*, Adv. Math. **315** (2017), 366–426. MR 3667589
50. Olga Plamenevskaya, *Bounds for the Thurston-Bennequin number from Floer homology*, Algebr. Geom. Topol. **4** (2004), 399–406. MR 2077671
51. ———, *Transverse knots and Khovanov homology*, Mathematical Research Letters **13** (2006), no. 4, 571–586 (EN).

52. _____, *Braid monodromy, orderings and transverse invariants*, Algebr. Geom. Topol. **18** (2018), no. 6, 3691–3718. MR 3868232
53. Jacob Andrew Rasmussen, *Floer homology and knot complements*, ProQuest LLC, Ann Arbor, MI, 2003, Thesis (Ph.D.)–Harvard University. MR 2704683
54. Jacob Rasmussen, *Khovanov homology and the slice genus*, Invent. Math. **182** (2010), no. 2, 419–447. MR 2729272
55. Lee Rudolph, *Quasipositivity as an obstruction to sliceness*, Bull. Amer. Math. Soc. (N.S.) **29** (1993), no. 1, 51–59. MR 1193540
56. Kouki Sato, *A full-twist inequality for the v^+-invariant*, Topology Appl. **245** (2018), 113–130. MR 3823992
57. Nikolai Saveliev, *Lectures on the topology of 3-manifolds*, revised ed., De Gruyter Textbook, Walter de Gruyter & Co., Berlin, 2012, An introduction to the Casson invariant. MR 2893651
58. Alexander N. Shumakovitch, *Rasmussen invariant, slice-Bennequin inequality, and sliceness of knots*, Journal of Knot Theory and Its Ramifications **16** (2007), no. 10, 1403–1412.
59. John R. Stallings, *Constructions of fibred knots and links*, Algebraic and geometric topology (Proc. Sympos. Pure Math., Stanford Univ., Stanford, Calif., 1976), Part 2, Proc. Sympos. Pure Math., XXXII, Amer. Math. Soc., Providence, R.I., 1978, pp. 55–60.
60. Bohua Zhan, *Computations in bordered Heegaard Floer homology*, 2013, available online.

Photographs

B. Acu et al. (eds.), *Research Directions in Symplectic and Contact Geometry and Topology*, Association for Women in Mathematics Series 27,
https://doi.org/10.1007/978-3-030-80979-9


GEMS


WORKSHOP FOR
WOMEN IN SYMPLECTIC
CONTACT GEOMETRY AN
TOPOLOGY